ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Bachelor Thesis International Econometrics and Operational Research

Title of the thesis: Goodness-of-Fit Tests for a Heavy Tailed Distribution

Name of the student: Paul Hendriks
Student ID number: 379345

Name of supervisor: A.J. Koning
Name of second assessor: P.J.F. Groenen

Date final version: June 30, 2017

Abstract

In this article the goodness-of-fit tests for heavy tailed distributions are evaluated regarding their relation to the underlying distribution, power, type I error and the sample size. The Kolmogorov-Smirnov test, Berk-Jones test, estimated score test, their corresponding quadratic tests and four different tests based on test measures of Rényi are evaluated in simulations and on Realized Variance data from the DAX stock index between 2000 and 2017. Samples with different sample sizes are simulated from different heavy tailed and non-heavy tailed distributions to compare the relation to the underlying distributions of the described test measures. The results indicate that the critical values of the test measures have a low relation to the underlying distribution if around 5% of the ordered observations are used. Furthermore, critical values should be simulated for different sample sizes for all test measures as critical values are dependent on both the proportion of the ordered distribution used and the number of observations. If this is impossible or impractical the quadratic Berk-Jones test can be used, as it is the least dependent on the sample size. Moreover, the first Rényi measure has the Type I error closest to the corresponding critical value for different distributions and percentages of the ordered observations used. On the other hand, the quadratic estimated score test has the highest power. If the relation between the power and the type I error is the focus then the Berk-Jones test has the most added value. Of the measures based on Rényi, the fourth one performs generally the best due to lower type I errors and higher powers.

Contents

1 Introduction

On Monday, 28 October 1929 the US stock market fell more than 13%. Similarly, on Sunday, 1 February 1953 the Dutch were surprised by a spring tide and sea storm, which led to the North Sea Flood. Both are examples of unfortunate extreme events having significant impact on those involved. In order to protect themselves, banks have been trying to maintain adequate capital buffers against potential losses, while the Dutch government has been trying to improve its water protection. In both cases parties tried to defend themselves against extreme events, without focusing on 'normal' days. To have an adequate policy against extreme events there needs to be an understanding about the distribution of these events. To find which distribution to fit to the observations, testing if there are relatively frequent extreme observations is useful. In this paper tests if samples show a heavy tail, that is, they have relatively frequent extreme observations, are evaluated. With these tests, policies can be made to adequately prepare for the possibility of extreme events. This seems a necessary contribution to the current literature, as recently heavy tails have been observed to exist in more phenomena than initially thought. Applications of heavy tails could be found in internet traffic, fire losses, hydrology and finance (Hjort and Koning, 2002).

Several test described by Kolmogorov (1933), Berk and Jones (1978) and Hjort and Koning (2002) exist to determine if the observations might have a heavy tail. These tests, their quadratic tests, and tests based on the academic work of Rényi (1973) will be evaluated. Firstly, the influence of the underlying distribution on the critical values will be compared for the different test measures. Further attention will be spend on the influence of two-tailed heavy tails and one-tailed heavy tail distributions on the test measures. Past research mainly focused on one-tailed heavy tails, while in several areas both frequent extreme positive observations and extreme negative observations are observed. An example concerns the water levels in rivers, which both needs to be at a certain level for ships to travel safely and below a certain level such that civilians are safe from floods. Currently, the water levels in rivers have been observed to have more frequent lower level and higher levels due to the influence of climate change (Katz et al., 2002). Other areas are only interested in either frequent positive or negative extreme returns. For example, in finance most of the risk is in extreme negative returns or periods of extreme volatility. Therefore, different tests could be better regarding two-tailed or one-tailed heavy tails. In section 2 the different goodness-of-fit tests for a heavy tailed distribution are explained.

Secondly, the type I error of the test measures related to different underlying distributions will be compared between test measures. Thirdly, the power of the test measures related to different underlying distributions will be compared between test measures. Lastly, the relation of the power and type I errors related to the sample size will be evaluated.

There is relatively limited past research on the topic of goodness-of-fit test for heavy tailed distributions. Initial literature about inferences on the tail of a distribution was done by Hill (1975) and the corresponding Hill estimator of the tail index. Initial work on goodness-of-fit was done by Davison and Smith (1990) and essential work was done by Koning and Peng (2008) by comparing various tests. This article mainly builds on the work of Koning and Peng (2008) and extends it with a further differentiation between one-tailed and two-tailed distributions, using the type I error as additional evaluation and using different sample sizes. In order to get a good understanding which test is the most powerful and rejects the least amount of 'true' heavy tails, a comparison is made in this article, i.e. the research question is:

"Amongst the described tests, which is the most powerful and has a type I error closest to the corresponding critical value in testing if a one-tailed or two-tailed population has heavy tails"

In further sections the tests will be described and evaluated. Furthermore, the tests will be applied to a series of observations regarding the Realized Volatility of the DAX stock index between 2000 and 2017. Finally, conclusions will be drawn together with potential further research ideas and limitations.

2 Methodology

2.1 Derivation

Generally, distributions can be classified in two groups: distribution with heavy tails and without heavy tails. Distributions with a heavy tail are characterized by tails that are not exponentially in their limits, as:

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha} \tag{1}$$

$$\text{or } P[X > x] \sim x^{-\alpha} \text{ as } x \to \infty \tag{2}$$

with the tail index $\alpha > 0$ as a real numbers. Distributions without a heavy tail are exponentially in their limits, as:

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = m\exp(kx) \tag{3}$$

with $m > 0$ and both m and k real numbers. Essentially, this means that the probability of 'extreme' observations is far more likely under distributions with heavy tails than under non-heavy tailed distributions. In order to test if a population has a heavy tail, there is the following null hypothesis and alternative hypothesis:

H_o: 'Distribution F is heavy tailed'
H_a: 'Distribution F is not heavy tailed'

The first aspect under the null hypothesis is that the tail index, α, is unknown and should be estimated. The most used estimator for α is the Hill estimator. Let $X_1, X_2, ..., X_n$ be ordered i.i.d observations from a distribution F. Then the Hill estimator is described as:

$$\hat{\alpha} = \big[k^{-1}\sum_{i=1}^{k} \ln(\frac{X_{n-k+i}}{X_{n-k}})\big]^{-1} \tag{4}$$

Secondly, there is no full description of the underlying distribution under the null hypothesis. The null hypothesis only describes the behavior of the tails. Therefore, there is a limited assumption on the underlying data distribution and only the potential tail of the distribution is tested. As a result, testing if a population has a heavy tail can be used to determine a fitting underlying distribution for the data.

In order to test for heavy tails the corresponding Kolmogorov-Smirnov test, Berk-Jones test, estimated-score test, their quadratic variants and four tests based on Rényi are used. To derive these tests, the equations Equation 5 and Equation 11 need to be defined. Therefore, let $X_1, X_2, ..., X_n$ be ordered i.i.d observations from a distribution F. Then $G_k(r)$ is defined as:

$$G_k(r) = k^{-1}\sum_{i=1}^{k} I(\frac{X_{n-k+i}}{X_{n-k}} \leq r) \tag{5}$$

with $r > 1$. In other words, $G_k(r)$ is a measure to determine how much percent of the most extreme observations are larger than a specific ranked observations by a proportion r. Important is that this measure focuses on the right tail for lower values of k, and includes the left tail for larger values of k. Notice that $G_k(r)$ accurately describes the empirical distribution of the sample $\frac{X_{n-k+1}}{X_{n-k}}, \frac{X_{n-k+2}}{X_{n-k}}, ..., \frac{X_n}{X_{n-k}}$. This is due to the fact that the empirical distribution is defined as:

$$\hat{F_n}(t) = n^{-1}\sum_{i=1}^{n} I(x \leq t) \tag{6}$$

To use $G_k(r)$ to test for heavy tails, the following assumption is made. If a distribution F is heavy tailed then let $U(x)$ be a function described as follows:

$$U(x) = \frac{1}{1 - F(x)} \tag{7}$$

$$\lim_{t \to \infty} \frac{\frac{U(tx)}{U(t)} - x^{\frac{1}{\alpha}}}{A(t)} = x^{\frac{1}{\alpha}} \frac{x^{\rho} - 1}{\rho} \tag{8}$$

with an existing function $A(t)$ that $A(t) \to 0$ as $t \to \infty$ and $\rho \leq 0$ (de Haan and Stadmüller (1996) and de Haan and Ferreira (2007)). Then let $\frac{k}{n} \to 0$ as $n \to \infty$. In other words, as the number of observations goes up to infinity, k can go to infinity as well, but it will go slower than n. Now, if it is the case that

$$\lim_{n \to \infty} \sqrt{k} A(\frac{n}{k}) = 0 \tag{9}$$

then $G_k(r)$ becomes the empirical distribution of a sample from the following Pareto distribution ((de Haan and Resnich, 1998) and (de Haan and Resnich, 1993)):

$$F(x, \alpha) = 1 - x^{-\alpha} \tag{10}$$

for $x \geq 1$. Unfortunately, a function $A(t)$ has not been found, so there is no proof this assumption holds in practice (de Haan and Ferreira, 2007). However, the following test measures based on this assumption appear to work well for lower values of k, indicating that it approximately holds for lower values of k (Koning and Peng, 2008).

In order to test that the null hypothesis of heavy tails the comparison is made between $G_k(r)$ and the cumulative Pareto distribution, $G(r, \alpha)$, defined as:

$$G(r, \alpha) = 1 - r^{-\alpha} \tag{11}$$

and the Pareto maximum likelihood estimator of α as [1]:

$$\hat{\alpha} = [k^{-1} \sum_{i=1}^{k} \ln(\frac{X_{n-k+i}}{X_{n-k}})]^{-1} \tag{12}$$

It can be observed that under this assumption the maximum likelihood estimator of the tail index is the Hill estimator. The comparison between $G(r, \alpha)$ and $G_k(r)$ is a comparison between the assumption that the population has a heavy tail ($G(r, \alpha)$) and no assumption on the tail ($G_k(r)$). Therefore, if the difference between the two measures is small enough, there will be less reason to reject heavy tails and vice versa.

2.2 Test Measures

Using the difference between $G(r, \alpha)$ and $G_k(r)$, the following test measures are designed. As r is a variable above one, all test measures use an optimization to find the highest differences between the two estimates. This is done in order to have an 'omnibus' test that for any r the hypothesis is not rejected.

Firstly, the Kolmogorov-Smirnov test (referred to as KS) uses the difference between the two measures, defined as(Kolmogorov, 1933):

$$\begin{aligned} & sup_{r>1} |\sqrt{k} KS(r, \hat{\alpha})| \\ =& sup_{r>1} |\sqrt{k}(1 - G_k(r) - (1 - G(r, \alpha)))| \\ =& sup_{r>1} |\sqrt{k}(1 - G_k(r) - r^{-\hat{\alpha}})| \end{aligned} \tag{13}$$

[1] derivation of the maximum likelihood is given in Appendix 1

Important is to note that not the difference between $G(r,\alpha)$ and $G_k(r)$ is used, but the difference between $1-G(r,\alpha)$ and $1-G_k(r)$. Both $G(r,\alpha)$ and $G_k(r)$ are CDF measures. Therefore, if there is interest in the tail, the values of $1-G(r,\alpha)$ and $1-G_k(r)$ should be investigated. Additionally, the value is scaled by the term $\sqrt{k}$ to adjust for potential differences in k and in order to find the limiting distribution later on.

Next, the Berk-Jones test (referred to as BJ) is defined as (Berk and Jones, 1978):

$$
\begin{aligned}
&sup_{r>1}(k * BJ(r,\hat{\alpha})) && (14)\\
=&sup_{r>1}(k * 2(G_k(r)\ln(\frac{G_k(r)}{G(r,\hat{\alpha})}) + (1-G_k(r))\ln(\frac{1-G_k(r)}{1-G(r,\hat{\alpha})}))) && (15)
\end{aligned}
$$

Important to consider here is that if $G(r,\alpha)$ and $G_k(r)$ are close together BJ will go towards zero and when they differ BJ will increase. Since the test measure has a supremum it will automatically look for the value of r with the highest difference between the two values.

Additionally, the estimated-score test (referred to as SC) is defined as (Hjort and Koning, 2002):

$$
\begin{aligned}
&sup_{r>1}|\sqrt{k}SC(r,\hat{\alpha})| && (16)\\
=&sup_{r>1}|\sqrt{k}(G_k(r) - \int_1^r (1-G_k(s))d\Lambda(s,\hat{\alpha}))| &&\\
=&sup_{r>1}|\sqrt{k}(G_k(r) - \hat{\alpha}\int_1^r \frac{1-G_k(s)}{s}ds)| &&\\
\text{as } &\Lambda(s,\hat{\alpha}) = -\ln(1-G(r,\hat{\alpha})) = \hat{\alpha}\ln(r) && (17)
\end{aligned}
$$

This test measure takes the integral of a discontinuous function $G_k(s)$. Therefore, the estimated-score test is not ready for use. In order for this test measure to be evaluated, the integral has to be rewritten. Unfortunately, the estimated-score test is non-continuous and thus the integrals cannot be calculated using classical Rienman Integrals. To calculate the integrals the Lebesgue-Stieltjes Integrals for a known finite number of steps, stepwise discontinuous function should be used. These can be described as follows:

$$
\int_A L(r)dGk(r) = \sum_{i:r_i\in A} L(r_i)m_i \qquad (18)
$$

In which r_i are the potential values of r and m_i the corresponding probabilities from $Gk(r)$. Rienman Integrals are special scenarios of Lebesgue-Stieltjes Integrals due to the fact that continuous functions have the same limit when a function evaluation is approached from both sides. Using Lebesque-Stieltjes Integrals, the integral can then be rewritten in the following equation:

$$\begin{aligned}
\int_1^r (1-G_k(s))d\Lambda(s,\hat{\alpha})) &= \int_1^r (1-k^{-1}\sum_{i=1}^k I(\frac{X_{n-k+i}}{X_{n-k}} \leq s))d\Lambda(s,\hat{\alpha})) \\
&= \int_1^r k^{-1}\sum_{i=1}^k I(\frac{X_{n-k+i}}{X_{n-k}} > s))d\Lambda(s,\hat{\alpha})) \\
&= k^{-1}\sum_{i=1}^k \int_1^r I(\frac{X_{n-k+i}}{X_{n-k}} > s))d\Lambda(s,\hat{\alpha})) \\
&= k^{-1}\sum_{i=1}^k \int_1^{min(r,\frac{X_{n-k+i}}{X_{n-k}})} d\Lambda(s,\hat{\alpha})) \\
&= k^{-1}\sum_{i=1}^k [\Lambda(s,\hat{\alpha}))]_1^{min(r,\frac{X_{n-k+i}}{X_{n-k}})} \\
&= k^{-1}\hat{\alpha}\sum_{i=1}^k \ln[min(r,\frac{X_{n-k+i}}{X_{n-k}})]
\end{aligned} \tag{19}$$

Similarly, this test measure looks for the largest difference by integrating $G_k(r)$ with respect to the cumulative hazard function of $G(r,\alpha)$. This hazard function illustrates the frequency of observing an extreme return according to $G(r,\alpha)$. Therefore, this test measure looks at the difference between frequencies between the values. Additionally, the following test measures based on Rényi (1953) and Rényi (1973) will be considered:

$$R_1 = sup_{r>\frac{X_{n-k+1}}{X_{n-k}}}(\frac{G_k(r)}{G(r,\hat{\alpha})}) \tag{20}$$

$$R_2 = sup_{r>1}(\frac{G(r,\hat{\alpha})}{G_k(r)}) \tag{21}$$

$$R_3 = sup_{r>1}(\frac{1-G_k(r)}{1-G(r,\hat{\alpha})}) \tag{22}$$

$$R_4 = sup_{r>1}(\frac{1-G(r,\hat{\alpha})}{1-G_k(r)}) \tag{23}$$

(24)

In this case the difference between $G(r,\alpha)$ and $G_k(r)$ is also used instead of only the difference between $1-G(r,\alpha)$ and $1-G_k(r)$. This would help showing the different characteristics between test measures based on $1-G_k(r)$ and $G_k(r)$. In all cases, if $G(r,\alpha)$ and $G_k(r)$ have similar values R_i will go to one and if they differ R_i goes either to zero or infinity. For R_1 the function is bounded by $\frac{X_{n-k+1}}{X_{n-k}}$ as otherwise the function evaluation for values of r between one and $\frac{X_{n-k+1}}{X_{n-k}}$ would be infinity.

All previous test measures are obtained using the supremum. An alternative to supremums are quadratic tests. In this case, quadratic tests are based on integration with respect to $G_{r,\hat{\alpha}}$ (Koning and Peng, 2008). Corresponding quadratic forms from Einmahl and McKeague (2003) and Wellner and Koltchinskii (2003)tests are:

$$KSI = \int_1^\infty (\sqrt{k}KS(r,\hat{\alpha})^2 dG(r,\hat{\alpha}) \tag{25}$$

$$BJI = \int_1^\infty kBJ(r,\hat{\alpha})dG(r,\hat{\alpha}) \tag{26}$$

$$SCI = \int_1^\infty (\sqrt{k}SC(r,\hat{\alpha}))^2 dG(r,\hat{\alpha}) \tag{27}$$

These tests are similar to the Kolmogorov-Smirnov test, the Berk-Jones test and the estimated-score test. Quadratic tests for the test measures based on Rényi are not available. These quadratic tests are very similar (not equivalent) to expected values of the KS, BJ and SCI test measures to the CDF of $G(r, \alpha)$. Furthermore, as r should be higher than 1, the integrals are only defined on the range of 1 to infinity. Similarly to the estimated score-test, the integral evaluations should be calculated as they are not readily available. The KSI, BJI and SCI have discontinuous parts and thus should Lebesgue-Stieltjes Integrals for a known finite number of steps, stepwise discontinuous function be used. Important is to realize that the quadratic tests are less prone to extreme values, due to taking the integral (expectation) and thereby smoothing the function evaluations. Using Lebesque-Stieltjes Integrals the tests are rewritten to be evaluated [2]. The KSI becomes:

$$
\begin{aligned}
Y_i &= \frac{X_{n-k+i}}{X_{n-k}} \\
KSI &= \frac{1}{3}k + k^{-1}\sum_{i=1}^{k}\sum_{j=1}^{k}(max(Y_i, Y_j))^{-\alpha} - 2\sum_{i=1}^{k}(Y_i^{-\alpha}) + \sum_{i=1}^{k}(Y_i^{-2\alpha})
\end{aligned} \tag{28}
$$

Important to notice here is that the previous test measures had a linear computation time, while the KSI has a quadratic computation time related to number of k. Especially for higher values of k the calculation time of critical values and test measures could explode. For the BJI the test measure becomes:

$$
\begin{aligned}
BJI = {} & 2\sum_{j=1}^{k-1}((j)\ln(k^{-1}j) + (k-j)\ln(k^{-1}(k-j)))(Y_j^{-\alpha} - Y_{j+1}^{-\alpha}) \\
& - 2\sum_{i=1}^{k}((-1 + Y_i^{-\alpha})\ln(1 - Y_i^{-\alpha}) - Y_i^{-\alpha}\ln(Y_i^{-\alpha}) - 1)
\end{aligned} \tag{29}
$$

The computation time of BJI is lower than the KSI, as it is linear with respect to k. Lastly, the SCI test measure becomes:

$$
\begin{aligned}
SCI = {} & k^{-1}\sum_{i=1}^{k}\sum_{j=1}^{k}(max(Y_i, Y_j))^{-\alpha} + \alpha^2 k^{-1}\sum_{i=1}^{k}\sum_{j=1}^{k}(-\frac{min(Y_i, Y_j)^{-\alpha}\ln(min(Y_i, Y_j))}{\alpha} \\
& - min(Y_i, Y_j)^{-\alpha}\ln^2(min(Y_i, Y_j)) - \frac{\alpha * min(Y_i, Y_j)^{-\alpha}\ln(min(Y_i, Y_j)) + 2 * min(Y_i, Y_j)^{-\alpha}}{\alpha^2} \\
& + \frac{2}{\alpha^2} + \ln(min(Y_i, Y_j))(-max(Y_i, Y_j)^{-\alpha}\ln(max(Y_i, Y_j)) - \frac{max(Y_i, Y_j)^{-\alpha}}{\alpha}) \\
& + \ln(min(Y_i, Y_j))(min(Y_i, Y_j)^{-\alpha}\ln(min(Y_i, Y_j)) + \frac{min(Y_i, Y_j)^{-\alpha}}{\alpha}) + \ln(Y_i)\ln(Y_j)max(Y_i, Y_j)^{-\alpha}) \\
& - 2\alpha k^{-1}\sum_{i=1}^{k}\sum_{j=1}^{k}([-max(Y_i, Y_j)^{-\alpha}\ln(max(Y_i, Y_j)) - \frac{max(Y_i, Y_j)^{-\alpha}}{\alpha} - (-Y_i^{-\alpha}\ln(Y_i) - \frac{Y_i^{-\alpha}}{\alpha})] \\
& + \ln(Y_j)max(Y_i, Y_j)^{-\alpha})
\end{aligned} \tag{30}
$$

The SCI test measure is extensive and has a quadratic computation time related to k. Furthermore, due to the extensive formula, mistakes can be made easily in calculating values of SCI. Practitioners using the SCI should pay close attention to make sure they calculate the test measure correctly.

[2] The entire derivations are stated in the appendix Equation 59, Equation 60, Equation 61

2.3 Optimization

To optimize the KS, BJ and SC test measures a linear search or other optimization algorithm cannot be used as the KS, BJ and SC test measures are discontinuous. Optimization algorithms assume functions are continues and can therefore skip pivotal points of the function evaluation.

2.3.1 KS

To optimize the test measures the derivate of $-(1-G(r,\alpha))$ with respect r is needed:

$$\frac{\partial-(1-G(r,\alpha))}{\partial r}=\frac{\partial-r^{-\alpha}}{\partial r}=\alpha r^{-\alpha-1} \tag{31}$$

Now two scenarios are looked at: firstly, the derivative of KS when it is between two points of r for which $G_k(r)$ changes value such that $G_k(r)$ does not change value. Secondly, the derivative of KS between two points of r for which $G_k(r)$ does change value.

Firstly, it is clear from the derivate in Equation 31 that KS generally increases for higher value of r for which $G_k(r)$ does not change. Secondly, when r is increased such that $G_k(r)$ does change value it is observed that $k^{-1}\sum_{i=1}^{k} I(Y_i \leq r)$ must increase, as the term $Y_i \leq r$ becomes less restrictive. Therefore, $1-G_k(r)$ will decrease by the value $\frac{1}{k}$. These scenarios indicate that the optimal value for r of KS is just before a value of $\frac{X_{n-k+i}}{X_{n-k}}$ or at infinity.

2.3.2 BJ

For the optimization of BJ only one case needs to be analyzed: the derivate of BJ when it is between two points of r for which $G_k(r)$ does not change value. BJ can be written as follows:

$$\begin{aligned}
BJ(r,\hat{\alpha}) &= 2(G_k(r)\ln(\frac{G_k(r)}{G(r,\hat{\alpha})})+(1-G_k(r))\ln(\frac{1-G_k(r)}{1-G(r,\hat{\alpha})}) \\
&= 2((k^{-1}\sum_{i=1}^{k} I(Y_i\leq r))\ln(\frac{(k^{-1}\sum_{i=1}^{k} I(Y_i\leq r)}{1-r^{-\alpha}}) \\
&+ (k^{-1}\sum_{i=1}^{k} I(Y_i> r))\ln(\frac{(k^{-1}\sum_{i=1}^{k} I(Y_i> r)}{r^{-\alpha}})) \\
&= 2((k^{-1}\sum_{i=1}^{k} I(Y_i\leq r))(\ln(k^{-1}\sum_{i=1}^{k} I(Y_i\leq r))-\ln(1-r^{-\alpha})) \\
&+ (k^{-1}\sum_{i=1}^{k} I(Y_i> r))(\ln(k^{-1}\sum_{i=1}^{k} I(Y_i> r))-\ln(r^{-\alpha}))) \\
&= 2((k^{-1}\sum_{i=1}^{k} I(Y_i\leq r))\ln(k^{-1}\sum_{i=1}^{k} I(Y_i\leq r))-(k^{-1}\sum_{i=1}^{k} I(Y_i\leq r))\ln(1-r^{-\alpha}) \\
&+ (k^{-1}\sum_{i=1}^{k} I(Y_i> r))\ln(k^{-1}\sum_{i=1}^{k} I(Y_i> r))-(k^{-1}\sum_{i=1}^{k} I(Y_i> r))\ln(r^{-\alpha}))
\end{aligned}$$

To get the derivate of BJ when it is between two points of r for which $G_k(r)$ does not change value the following equation is used:

$$\begin{aligned}
&\frac{\partial 2((k^{-1}\sum_{i=1}^{k} I(Y_i \leq r))\ln(k^{-1}\sum_{i=1}^{k} I(Y_i \leq r)) - (k^{-1}\sum_{i=1}^{k} I(Y_i \leq r)))\ln(1-r^{-\alpha})}{\partial r}\\
&+\frac{\partial 2((k^{-1}\sum_{i=1}^{k} I(Y_i > r))\ln(k^{-1}\sum_{i=1}^{k} I(Y_i > r)) - (k^{-1}\sum_{i=1}^{k} I(Y_i > r))\ln(r^{-\alpha}))}{\partial r}\\
&=\frac{\partial - 2(k^{-1}\sum_{i=1}^{k} I(Y_i \leq r))\ln(1-r^{-\alpha})}{\partial r} - \frac{\partial 2(k^{-1}\sum_{i=1}^{k} I(Y_i > r))\ln(r^{-\alpha}))}{\partial r}\\
&= -2(k^{-1}\sum_{i=1}^{k} I(Y_i \leq r))\frac{\alpha r^{-\alpha-1}}{1-r^{-\alpha}} + 2(k^{-1}\sum_{i=1}^{k} I(Y_i > r))\frac{\alpha r^{-\alpha-1}}{r^{-\alpha}}\\
&= 2\alpha r^{-\alpha-1}(\frac{1-G_k(r)}{1-G(r,\alpha)} - \frac{G_k(r)}{G(r,\alpha)})
\end{aligned} \tag{32}$$

This results in the fact that the change in r can have a positive or negative effect on the value of BJ and is mainly dependent on the value of both $G_k(r)$ and $G(r,\alpha)$. It is clear that for values of r close to one, the derivative is positive, due to $\frac{G_k(r)}{G(r,\alpha)}$ being close to zero. Therefore, the BJ test measure will be evaluated at the values for r from one onwards with step size of 0.001 until $\frac{X_n}{X_{n-k}}$ or the point at which an increase of r does not increase the value of BJ. From $\frac{X_n}{X_{n-k}}$ onwards it is clear that the derivative is negative as $\frac{1-G_k(r)}{1-G(r,\alpha)}$ becomes zero.

2.3.3 SC

Here one scenario is again evaluated: The derivative of SC when it is between two points of r for which $G_k(r)$ changes value. This derivative equals:

$$\begin{aligned}
&\frac{\partial G_k(r) - k^{-1}\hat{\alpha}\sum_{i=1}^{k}\ln[min(r, \frac{X_{n-k+i}}{X_{n-k}})]}{\partial r}\\
&= \frac{\partial - k^{-1}\hat{\alpha}(k_1\ln[r] + \sum_{i=k_1+1}^{k}\ln[\frac{X_{n-k+i}}{X_{n-k}}])}{\partial r}\\
&= -k^{-1}\hat{\alpha}\frac{k_1}{r}\\
&\text{with } 1 \leq k_1 \leq k
\end{aligned} \tag{33}$$

Therefore, as r increases SC will decrease. As a result, to calculate the optimal value for this test measure only the evaluations of r equal to $\frac{X_{n-k+i}}{X_{n-k}}$ and a value just above one needs to evaluated.

2.3.4 Rényi

Similar to the previous cases, the Rényi test measurements will be evaluated. Firstly, the derivatives are determined for changes in r that do not affect the value of $G_k(r)$:

$$R_1\text{:} \frac{\partial \frac{k^{-1}\sum_{i=1}^{k} I(Y_i \leq r)}{1-r^{-\alpha}}}{\partial r} = \frac{0 - k^{-1}\sum_{i=1}^{k} I(Y_i \leq r)\alpha r^{-\alpha-1}}{(1-r^{-\alpha})^2} \leq 0 \tag{34}$$

$$R_2\text{:} \frac{\partial \frac{1-r^{-\alpha}}{k^{-1}\sum_{i=1}^{k} I(Y_i \leq r)}}{\partial r} = \frac{k^{-1}\sum_{i=1}^{k} I(Y_i \leq r)\alpha r^{-\alpha-1} - 0}{(k^{-1}\sum_{i=1}^{k} I(Y_i \leq r))^2} \geq 0 \tag{35}$$

$$R_3\text{:} \frac{\partial \frac{k^{-1}\sum_{i=1}^{k} I(Y_i > r)}{r^{-\alpha}}}{\partial r} = \frac{0 + k^{-1}\sum_{i=1}^{k} I(Y_i > r)\alpha r^{-\alpha-1}}{(r^{-\alpha})^2} \geq 0 \tag{36}$$

$$R_4\text{:} \frac{\partial \frac{r^{-\alpha}}{k^{-1}\sum_{i=1}^{k} I(Y_i > r)}}{\partial r} = \frac{-\alpha r^{-\alpha-1} k^{-1}\sum_{i=1}^{k} I(Y_i > r) - 0}{(k^{-1}\sum_{i=1}^{k} I(Y_i > r))^2} \leq 0 \tag{37}$$

$$\tag{38}$$

These test measures indicate that R_1 and R_4 are optimal for values of r either at a point of $\frac{X_{n-k+i}}{X_{n-k}}$ or close to one (or $\frac{X_{n-k+1}}{X_{n-k}}$ for R_1) as the derivative is negative. Furthermore, as the derivative of R_2 and R_3 is positive, their optimal values are reached for r equal to either at a point of $\frac{X_{n-k+i}}{X_{n-k}}$, just before a point of $\frac{X_{n-k+i}}{X_{n-k}}$ or at infinity.

2.4 Evaluation

In order to reject the hypothesis the critical values of these test measures need to be calculated. Corresponding critical values need to be simulated under the null hypothesis for an adequate sample. All previously discussed tests will be evaluated with respect to their power, i.e.:

$$\text{Power} = \mathbf{P}(\text{reject } H_0 | H_1 \text{ is true}) \tag{39}$$

and their corresponding type I errors, i.e.:

$$\lambda = \mathbf{P}(\text{reject } H_0 | H_0 \text{ is true}) \tag{40}$$

With

$$H_0 = \text{The underlying distribution of the observations is heavy tailed}$$
$$H_1 = \text{The underlying distribution of the observations is not heavy tailed}$$

Important is to consider that these two measures are related by Bayes law in the following way:

$$\mathbf{P}(\text{reject } H_0) = \mathbf{P}(\text{reject } H_0 | H_1 \text{ is true})\mathbf{P}(H_1 \text{ is true}) + \mathbf{P}(\text{reject } H_0 | H_0 \text{ is true})\mathbf{P}(H_0 \text{ is true}) \tag{41}$$

Therefore, assuming that $\mathbf{P}(H_1 \text{ is true})$ and $\mathbf{P}(H_0 \text{ is true})$ remain fixed in a real world application, an increase in type I error would mean a decrease in power and vice versa. Both measures are then negatively related, i.e. measures with a higher power will have a lower probability of a type I error and vice versa. Both evaluations are important to consider regarding the reliability and validity of the test measures. The corresponding power will give an idea if the test measures are able to reject the a heavy-tailed solution given the number of observations. Therefore, different number of simulated observations might be helpful to determine the differences in increased power of the different test measures. Furthermore, the type I error will function as a unconditional coverage for the created critical values. For a confidence level of 95% there should approximately be 5% of wrongly rejected hypotheses, regardless of the type of heavy-tailed underlying distribution. The test measures are supposed to only differentiate between a distribution with and without heavy tails. In other words, if the observations come from a distribution with a heavy tail, no matter what distribution, it should be able to not-reject a heavy tail with equal power and type I error for all heavy-tailed distributions. Similarly, if the observations are obtained from a non-heavy tailed distribution, it should be able to reject a heavy tail, no matter the underlying non-heavy tailed distribution. Therefore, a good test measure has a correct type one error and high power.

The test measures will be evaluated regarding to the two types of heavy tails: one-tailed and two-tailed. For the one-tailed evaluation, critical values will be empirically calculated using 100.000 random samples with a sample size of $n = 1000$ with $k = 20, 30, ..., 200$ from a Lévy distribution for both cases of 95% and 99% confidence levels. The Lévy distribution is defined as follows:

$$f_{Levy}(x|\mu = 0; c = 1) = \sqrt{\frac{c}{2\pi}} \frac{\exp(-\frac{c}{2(x-\mu)})}{(x-\mu)^{3/2}} \tag{42}$$

Next, the corresponding type I error will be evaluated against 10.000 simulations with $n = 100, 200, ..., 1000$ and $k = 20, 30, ..., 200$ of the alternative one-tailed heavy tailed distributions of the Weibull distribution, Pareto distribution and log-normal distribution:

$$f_{Weibull}(x|\lambda = 1; \kappa = 1.5) = \frac{\kappa}{\lambda}(\frac{x}{\lambda})^{k-1} \exp((-x/\lambda)^{\kappa})) \tag{43}$$

$$f_{Pareto}(x|x_m = 1; \alpha = 1) = \frac{\alpha x_m^{\alpha}}{x^{\alpha+1}} \qquad x \geq x_m \tag{44}$$

$$f_{LogNormal}(x|\mu = 0; \sigma = 1) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(\ln(x) - \mu)^2}{2\sigma^2}) \tag{45}$$

Furthermore, the corresponding power will be evaluated against 10.000 simulations with $n = 100, 200, ..., 1000$ of the alternative one-tailed non-heavy tailed distributions of the exponential distribution, Chi-squared distribution and F-distribution:

$$f_{exp}(x|\lambda = 1) = \lambda^{-\lambda x} \tag{46}$$

$$f_{Chi-Squared}(x|k = 3) = \frac{1}{2^{\frac{k}{2}}\Gamma\frac{k}{2}} x^{\frac{k}{2}-1} \exp(-\frac{x}{2}) \tag{47}$$

$$f_F(x|d_1 = 5; d_2 = 1) = \frac{\sqrt{\frac{(d_1 x)_1^d d_2^{d_2}}{(d_1 x + d_2)^{d_1+d_2}}}}{x\mathbf{B}(\frac{d_1}{2)}, \frac{d_2}{2})} \tag{48}$$

$$\text{with } \mathbf{B}(y, z) = \int_0^1 t^{x-1}(1-t)^{y-1} dt \tag{49}$$

For the two-sided heavy tails, critical values will be empirically calculated using 100.000 random samples with a sample size of $n = 1000$ for $k = 20, 30, ..., 200$ for both 95% and 99% confidence levels from a Cauchy distribution:

$$f_{Cauchy}(x|x_0 = 0; \gamma = 1) = \frac{1}{\pi * \gamma[1 + (\frac{x-x_0}{\gamma})^2]} \tag{50}$$

Next, the corresponding type I error will be evaluated against 10.000 simulations with $n = 100, 200, ..., 1000$ and $k = 20, 30, ..., 200$ form the alternative two-tailed heavy tailed distributions of the t-distribution:

$$f_t(x|v = 3) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{v\pi}\Gamma(\frac{v}{2})}(1 + \frac{x^2}{v})^{\frac{-v+1}{2}} \tag{51}$$

$$f_t(x|v = 5) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{v\pi}\Gamma(\frac{v}{2})}(1 + \frac{x^2}{v})^{\frac{-v+1}{2}} \tag{52}$$

$$\tag{53}$$

Furthermore, the corresponding power will be evaluated against 10.000 simulations with $n = 100, 200, ..., 1000$ and $k = 20, 30, ..., 200$ of the non-heavy tailed distributions of the normal distribution and Laplace distribution:

$$f_{normal}(x|\mu = 0; \sigma^2 = 2) = \frac{1}{\sqrt{2\sigma^2\pi}} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) \tag{54}$$

$$f_{Laplace}(x|\mu = 0; b = 2) = \frac{1}{2b} \exp(-\frac{|x-\mu|}{b}) \tag{55}$$

By simulating samples with different sample sizes the effect of the sample size on the type I error and power of the test can be observed.

3 Data

Next to a simulation the test measures will be evaluated using daily Realized Variance (RV) data of the German DAX stock index between January 1, 2000 and March 2, 2017 obtained from the Oxford-Man Institute of Quantitative Finance Realized Library (4344 observations) and defined as:

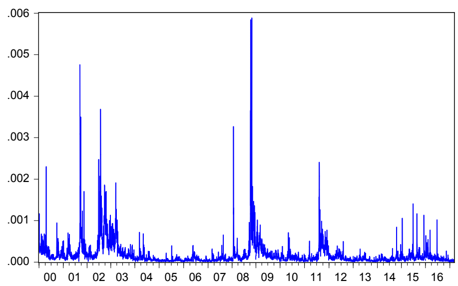
$$RV_t = \sum_{i=1}^{n} r_{t,i}^2 \tag{56}$$

In which t is a period of a day and each sub-period i is a five minute interval. It can be observed in Figure 1 that there are exceptional cases of extreme observations in the RV. Especially in the periods 2001-2002, 2008-2009, 2011 and around 2015 there are higher observations of the RV.

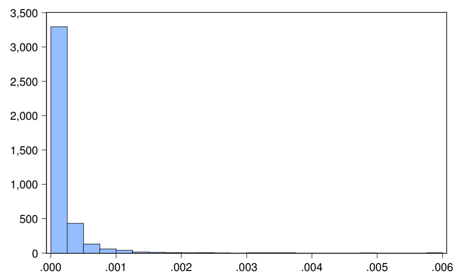
In order to test for heavy-tailes the test describes in section 2 will be applied with k ranging between 1 and 4343. For the values of k for which critical values are available, the measurement values will be compared and the hypothesis for a heavy-tail will be evaluated. For higher values of k it can be seen if the assumption that $\lim_{n\to\infty} \sqrt{k}A(\frac{n}{k}) = 0$ holds for the data.

Figure 1

(a) Realized Variance Graph



(b) Realized Variance Historgram



4 Results

4.1 Critical Values

The corresponding critical values for 95% and 99% confidence levels are calculated using a sample size of $n = 1000$ for both the Lévy and Cauchy distributions. The critical values for 99% confidence levels are shown in Table 7 and Table 8 in the appendix. The corresponding critical values for 95% confidence levels with the Lévy distribution as underlying are the following:

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0,983	0,220	6,249	1,410	1,192	0,446	18,809	4,152	1,950	2,988
30	1,001	0,221	6,584	1,385	1,227	0,454	19,287	4,334	2,158	3,208
40	1,008	0,220	6,797	1,373	1,242	0,453	19,256	4,454	2,310	3,379
50	1,018	0,221	6,986	1,358	1,260	0,452	19,783	4,486	2,413	3,501
60	1,021	0,222	7,064	1,352	1,267	0,453	19,371	4,541	2,512	3,608
70	1,027	0,221	7,229	1,350	1,269	0,456	19,369	4,565	2,551	3,686
80	1,030	0,220	7,367	1,355	1,280	0,462	20,352	4,616	2,618	3,770
90	1,035	0,221	7,410	1,349	1,282	0,455	19,694	4,621	2,662	3,823
100	1,036	0,222	7,505	1,348	1,287	0,459	19,516	4,636	2,709	3,880
110	1,038	0,221	7,563	1,351	1,292	0,458	19,493	4,651	2,776	3,919
120	1,049	0,220	7,674	1,340	1,297	0,458	19,889	4,677	2,754	3,957
130	1,046	0,221	7,673	1,344	1,295	0,458	19,448	4,691	2,785	3,998
140	1,047	0,221	7,760	1,353	1,295	0,456	20,015	4,717	2,823	4,042
150	1,050	0,220	7,822	1,343	1,305	0,456	20,149	4,722	2,847	4,075
160	1,050	0,221	7,811	1,342	1,307	0,457	19,599	4,739	2,825	4,090
170	1,051	0,224	7,858	1,346	1,310	0,458	19,887	4,707	2,843	4,144
180	1,055	0,222	7,884	1,342	1,307	0,458	19,699	4,728	2,852	4,163
190	1,057	0,221	7,971	1,358	1,309	0,460	19,652	4,742	2,843	4,216
200	1,060	0,222	7,969	1,356	1,303	0,462	19,367	4,779	2,834	4,240

Table 1: Critical Values Lévy for 95% Confidence Level

If these critical values are compared to the simulated critical values in Koning and Peng (2008) then these are minor differences in the values, due to a different underlying distribution. The underlying distribution used in Koning and Peng (2008) is the Frechet distribution: $F(x) = \exp -x^{-1}$. Therefore, critical values can be different and dependent on the underlying distribution. Generally, the critical values increase for higher values of k for both the Lévy distribution and the Frechet distribution.

To test if the critical values become different for one-tailed and two-tailed distributions, the critical values are also calculated with a underlying Cauchy Distribution. The critical values are as follows:

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0,994	0,219	6,250	1,414	1,189	0,441	19,580	4,159	2,082	2,987
30	1,006	0,220	6,556	1,390	1,231	0,451	19,314	4,348	2,165	3,226
40	1,016	0,220	6,763	1,371	1,245	0,452	19,124	4,480	2,301	3,411
50	1,026	0,221	6,937	1,375	1,252	0,450	19,178	4,556	2,362	3,536
60	1,032	0,222	7,102	1,377	1,268	0,462	19,141	4,614	2,422	3,645
70	1,040	0,221	7,147	1,373	1,273	0,458	19,677	4,649	2,472	3,741
80	1,044	0,224	7,257	1,386	1,281	0,458	19,135	4,736	2,528	3,841
90	1,057	0,224	7,282	1,401	1,296	0,465	19,430	4,778	2,501	3,949
100	1,066	0,228	7,358	1,404	1,297	0,470	19,107	4,817	2,518	4,042
110	1,072	0,233	7,439	1,438	1,310	0,483	18,774	4,867	2,504	4,109
120	1,084	0,237	7,511	1,462	1,321	0,491	18,145	4,912	2,487	4,212
130	1,103	0,244	7,655	1,507	1,347	0,510	18,142	5,000	2,425	4,298
140	1,116	0,254	7,752	1,573	1,362	0,530	18,220	5,078	2,410	4,431
150	1,144	0,269	7,876	1,637	1,392	0,564	17,870	5,193	2,349	4,504
160	1,169	0,281	8,072	1,731	1,425	0,599	17,706	5,254	2,314	4,617
170	1,197	0,300	8,323	1,842	1,471	0,648	17,392	5,297	2,253	4,764
180	1,232	0,327	8,639	1,986	1,515	0,696	17,117	5,399	2,184	4,905
190	1,271	0,349	8,944	2,147	1,570	0,754	16,801	5,521	2,105	5,024
200	1,308	0,383	9,407	2,347	1,626	0,835	16,158	5,572	2,041	5,191

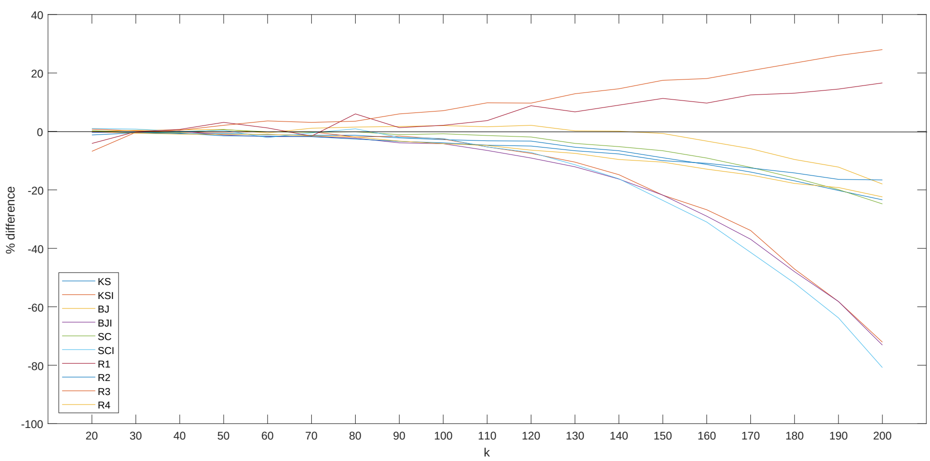
Table 2: Critical Values Cauchy for a 95% Confidence Level

When the differences between the critical values of a one-tailed and two-tailed distributions are evaluated it is noted that most supremum measures slowly increase if k is increased for a Cauchy distribution. However, this is not always the case with a Lévy distribution. Furthermore, most test measures seem to have either a stable increase or remain stable, whereas the first Rényi test measure is less stable. The decreased stability is due to the lower values of $G(r, \hat{\alpha})$. Therefore, the test measure is more dependent on extreme observations and can deviate more.

If the relative difference between the critical values over all k of the distributions is observed for then it is noted that there are some differences. While test measures, such as the third Rényi measure, hardly differ between the distributions, other test measures, such as the BJI, differ more.

If the distinction between the test measures for different values of k is examined the following differences are shown for the critical values at the 95% confidence level Table 9 and Figure 2[3]:

Figure 2: Relative Differences Critical Values at 95% Confidence Level



[3]The differences for the critical values at the 99% confidence level are shown in Table 10 and Figure 9

Here it can be seen that for bigger values of k the test measures differ more. This is potentially due to the reason that for larger values of k more of the empirical distribution is used in the test measure. Therefore, the test measures start to compare distributions instead of tails. As a result, it is expected that test measures with lower values of k are better able to differentiate between heavy tails and non-heavy tails. As most researchers are mostly interested in the tail behavior around 5% and 1% of the distribution, recommended is that k is at most 5% of the number of observations. It can be seen in Table 9 that these values of k have a low difference between different underlying distributions.

4.2 Type I Error

When the type I errors of each test measure for a sample size of $n = 1000$ are observed over an average of the discussed distributions in subsection 2.4, the following patterns are shown using the critical values at the 95% confidence level[4]:

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0,079	0,060	0,041	0,069	0,079	0,059	0,046	0,072	0,238	0,062
30	0,094	0,074	0,039	0,086	0,091	0,077	0,044	0,075	0,150	0,088
40	0,110	0,094	0,045	0,104	0,114	0,097	0,045	0,082	0,092	0,113
50	0,131	0,116	0,049	0,127	0,139	0,124	0,042	0,084	0,060	0,146
60	0,152	0,145	0,056	0,158	0,173	0,161	0,043	0,087	0,038	0,175
70	0,184	0,180	0,068	0,193	0,211	0,204	0,043	0,095	0,026	0,217
80	0,217	0,220	0,082	0,229	0,253	0,247	0,041	0,095	0,017	0,250
90	0,253	0,263	0,106	0,273	0,296	0,302	0,042	0,099	0,013	0,286
100	0,290	0,315	0,133	0,324	0,349	0,362	0,041	0,103	0,009	0,329
110	0,334	0,359	0,164	0,367	0,406	0,419	0,039	0,107	0,008	0,369
120	0,372	0,414	0,204	0,420	0,458	0,479	0,041	0,108	0,006	0,406
130	0,415	0,463	0,251	0,468	0,507	0,531	0,041	0,110	0,006	0,447
140	0,456	0,514	0,301	0,516	0,559	0,583	0,040	0,110	0,005	0,476
150	0,494	0,554	0,353	0,559	0,598	0,631	0,040	0,112	0,005	0,515
160	0,533	0,597	0,401	0,597	0,644	0,666	0,040	0,113	0,005	0,550
170	0,564	0,628	0,452	0,633	0,673	0,701	0,039	0,120	0,004	0,577
180	0,597	0,661	0,495	0,663	0,703	0,727	0,040	0,122	0,004	0,606
190	0,625	0,689	0,534	0,688	0,729	0,747	0,039	0,123	0,004	0,626
200	0,653	0,711	0,578	0,710	0,750	0,765	0,039	0,124	0,003	0,649

Table 3: Average Type I errors

Here it can clearly be seen that the third Rényi test measure does not give the intended type I error for any value of k. Furthermore, most other test measures give good type I errors for lower values of k, but have too high values for higher values of k. As discussed in the previous section, this could be due to distribution differences. For lower values of k the best test measure tends to be the BJ. However, after k passed a value of 70, the best test measure related to the type one error is the first Rényi test measure, which remain stable around 4%, while other climb to values around 55%. Overall, the first Rényi measure seems to remain closer to its intended type I error. Of the test measures not related to Rényi, the BJ measure performs the best as the type one error is the lowest for higher values of k. Furthermore, the fourth Rényi measure performs similarly to the non-Rényi measures.

[4]The type I errors of the critical values at the 99% confidence level are shown in Table 11

4.3 Power

When the power of each test measure for a sample size of $n = 1000$ are observed over the discussed distributions in subsection 2.4, the following patterns are shown using the critical values of the 95% confidence level [5]:

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0,106	0,083	0,037	0,095	0,076	0,087	0,042	0,087	0,181	0,076
30	0,134	0,118	0,040	0,131	0,107	0,124	0,042	0,095	0,093	0,127
40	0,173	0,160	0,053	0,175	0,154	0,178	0,041	0,100	0,049	0,189
50	0,213	0,213	0,067	0,226	0,205	0,239	0,041	0,106	0,029	0,257
60	0,263	0,271	0,087	0,284	0,261	0,310	0,039	0,112	0,020	0,329
70	0,312	0,336	0,121	0,353	0,330	0,384	0,040	0,121	0,014	0,395
80	0,372	0,402	0,162	0,409	0,401	0,455	0,040	0,115	0,012	0,462
90	0,420	0,466	0,219	0,478	0,459	0,527	0,041	0,121	0,011	0,522
100	0,474	0,525	0,285	0,536	0,522	0,593	0,040	0,128	0,010	0,578
110	0,523	0,574	0,350	0,588	0,574	0,633	0,039	0,129	0,010	0,624
120	0,563	0,623	0,418	0,636	0,620	0,676	0,039	0,131	0,010	0,666
130	0,605	0,664	0,479	0,674	0,664	0,708	0,039	0,131	0,009	0,696
140	0,640	0,696	0,536	0,702	0,698	0,738	0,037	0,135	0,010	0,721
150	0,671	0,725	0,589	0,729	0,721	0,756	0,038	0,133	0,009	0,743
160	0,696	0,744	0,632	0,751	0,744	0,773	0,039	0,139	0,010	0,762
170	0,720	0,759	0,668	0,767	0,761	0,785	0,038	0,142	0,010	0,770
180	0,738	0,776	0,702	0,780	0,774	0,793	0,039	0,143	0,009	0,783
190	0,755	0,785	0,724	0,788	0,786	0,799	0,037	0,142	0,010	0,789
200	0,768	0,792	0,747	0,795	0,794	0,804	0,037	0,146	0,010	0,795

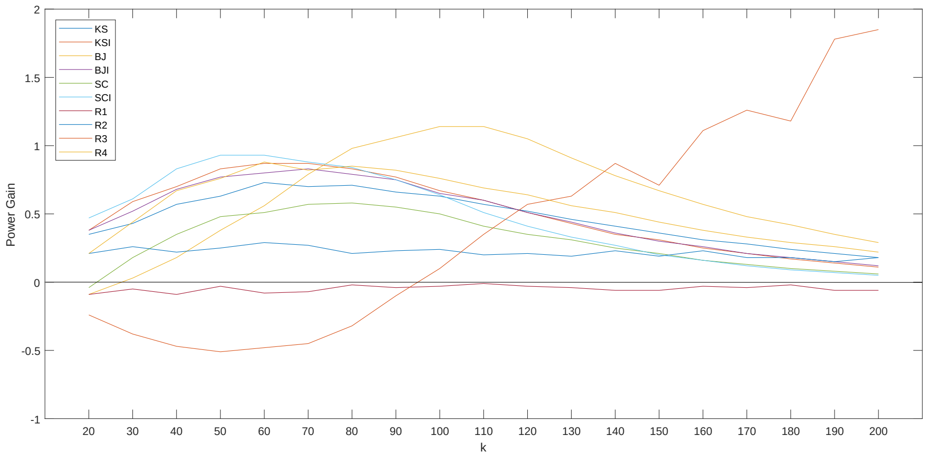
Table 4: Average Power at critical values at 95% Confidence Level for $n = 1000$

Taking into account the type I error, it can be seen that no test measure is clearly able to distinguish between heavy-tailed and non-heavy tailed distributions. In fact, for most test measures the power is only slightly higher than the type I error, indicating that all test measures are having trouble differentiating between heavy and non-heavy tailed distributions. Furthermore, it is seen that the same pattern as with the type I error; for higher values of k the power increases for most test measures and test measures that performed well for the type I error perform worse for the power. Therefore, a trade-off exists; increase k and improve the power of the test measures or decrease k to improve the type I error. In Figure 3 the difference between the Power and the type I error divided by the type I error is shown[6]. A higher value means that the test measure has more added value to distinguish between a heavy tailed and non-heavy tailed distribution. It can be seen that most test measures have the greatest relative difference between the two for values of k around 60, or 6% of the number of observations. Furthermore, it is shown that the first and third Rényi test measures have almost no added value, while the second has a constant added value over the values of k. Of the other test measures, BJ has the highest added value for values of k near 100, while the KS, KSI, BJI, SC (and SCI) have more added value for lower values of k. While, the first Rényi measure was good for the type I errors, it is not good for the power. The fourth Rényi measure performs well in both as it has a relatively low type I error and a high power with a high power gain.

[5] The power with the critical values of the 99% confidence levels are shown in Table 12

[6] In Table 13, Table 14 and Figure 10 the power gain data and graphs using 95% and 99% confidence level's critical values are shown

Figure 3: Power Gain with critical values for 95% Confidence Level



4.4 Relative Sample Size

In the previous sections the test measures were evaluated for a sample with the same size as simulated to calculate the critical values. However, it could be the case the critical values could change for different sample sizes or relative sample sizes compared to k. By using sample sizes ranging from $n = 100, 200, ..., 1000$ and the critical values described before, the following type I errors and powers are found using the critical values of the 95% confidence levels [7]:

n	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
100	0,727	0,733	0,588	0,581	0,721	0,746	0,055	0,540	0,160	0,717
200	0,760	0,768	0,667	0,618	0,760	0,777	0,052	0,525	0,142	0,754
300	0,693	0,702	0,595	0,679	0,696	0,717	0,071	0,388	0,137	0,709
400	0,651	0,664	0,550	0,673	0,659	0,682	0,049	0,284	0,035	0,679
500	0,621	0,638	0,514	0,646	0,632	0,658	0,034	0,209	0,019	0,656
600	0,594	0,614	0,484	0,623	0,609	0,637	0,035	0,175	0,020	0,636
700	0,572	0,594	0,458	0,604	0,590	0,621	0,036	0,156	0,022	0,618
800	0,550	0,575	0,434	0,585	0,571	0,605	0,037	0,144	0,023	0,602
900	0,531	0,559	0,413	0,569	0,555	0,591	0,038	0,136	0,024	0,588
1000	0,512	0,544	0,395	0,553	0,540	0,578	0,039	0,129	0,025	0,573

Table 5: Average Power with Critical Values for 95% Confidence Level with Different Sample Sizes

[7] The type I errors and powers using the critical values of the 99% confidence levels are found in Table 15 and Table 16

n	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
100	0,731	0,735	0,585	0,576	0,743	0,746	0,053	0,539	0,148	0,717
200	0,763	0,769	0,664	0,612	0,778	0,781	0,051	0,523	0,126	0,755
300	0,685	0,695	0,570	0,670	0,708	0,711	0,070	0,375	0,118	0,674
400	0,609	0,623	0,487	0,630	0,641	0,643	0,050	0,265	0,038	0,606
500	0,551	0,568	0,428	0,575	0,589	0,591	0,035	0,186	0,023	0,553
600	0,504	0,523	0,381	0,530	0,547	0,549	0,037	0,150	0,027	0,507
700	0,463	0,484	0,340	0,491	0,510	0,515	0,039	0,131	0,030	0,466
800	0,424	0,450	0,302	0,455	0,476	0,484	0,040	0,119	0,033	0,430
900	0,389	0,417	0,269	0,423	0,445	0,458	0,040	0,109	0,036	0,398
1000	0,357	0,386	0,241	0,392	0,414	0,433	0,041	0,104	0,038	0,369

Table 6: Average Type I Error with Critical Values for 95% Confidence Level with Different Sample Sizes

Here the power and type I error are averaged over the values of k. Important is to keep in mind that for $n = 100$ only values of k below 100 can be used. This also results in the fact that the power and type I error are lower for $n = 100$ than $n = 200$. From these tables it can be concluded that all test measures are dependent on the number of observations. Furthermore, there is no stable relationship between 'n' and k for which the critical values are stable. Therefore, critical values should be simulated for every different sample size. In case that this is not possible or impractical, the BJI measure seems the most robust to sample size as its type I error is relatively low while its power is relatively high.

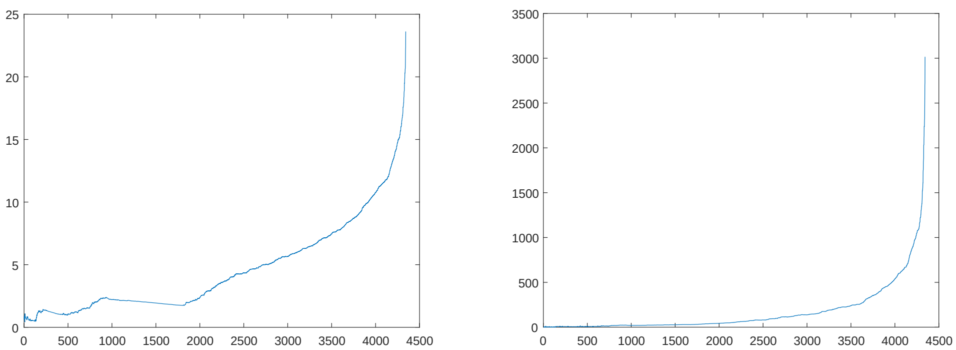
4.5 Realized Volatility

In order to evaluate the test measures using the Realized Volatility, their stability with respect to k will be inspected and their corresponding rejection of heavy tails.

Firstly, the KS test measure appears to be relatively stable for lower values of k. It seems to continuously move between subsequent values of k, but there are several 'phases'. The test measure repeats a cycle of an increase and consequent decrease two times, after which it seems to slowly go to infinity. The BJ test measure is initially very stable, but appears to explode more than the KS measure for higher values of k. Therefore, $G(r, \hat{\alpha})$ will explode. Both test measures do not reject the heavy-tails null hypothesis for lower values of k of which critical values are calculated (see Table 17 and Table 18). However, once k reaches 160 the KS test measure rejects it, while the BJ test measure does not. This could be due to the relatively higher type one error for the KS test measure for higher values of k in relation to the BJ test measure.

Figure 4

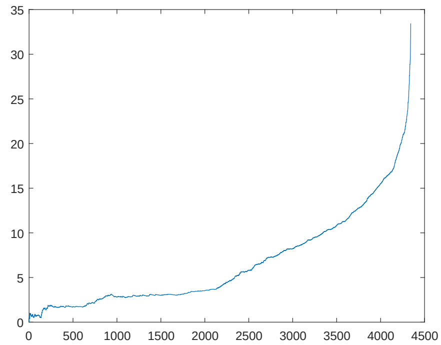
(a) KS (b) BJ



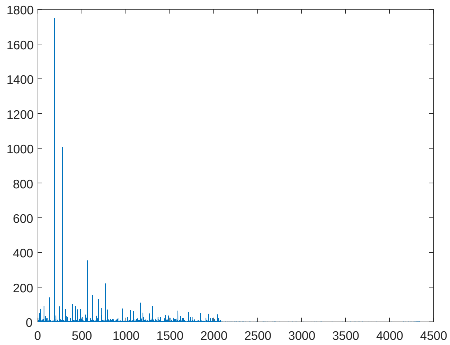
The SC test measure seems to have similar problems as the KS test measures, as it slowly goes to infinity for higher values of k. There can be several reasons why test measures could explode to infinity. Firstly, there could be a dependence in the data (Koning and Peng, 2008). Another reason could be that the assumptions $\lim_{n\to\infty}\sqrt{k}A(\frac{n}{k}) = 0$ does not hold for higher values of k. The first Rényi test measure seems to be much less stable. Although, it does go towards a stable point for higher values of k, it can abruptly increase for lower values of k. Therefore, this test measure seems unreliable. Both test measures do not reject the heavy-tails null hypothesis for lower values of k. However, once k reaches 160 the SC test measure rejects it, similarly as the KS test measure. Similarly, this could be due to the higher type one error for the SC test measure for higher values of k.

Figure 5

(a) SC



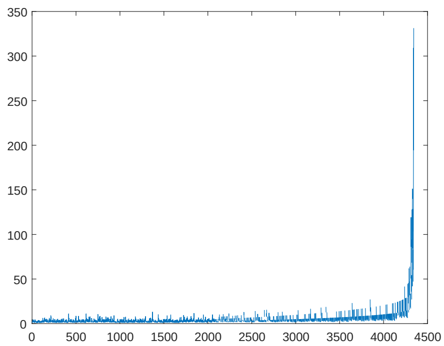
(b) R1



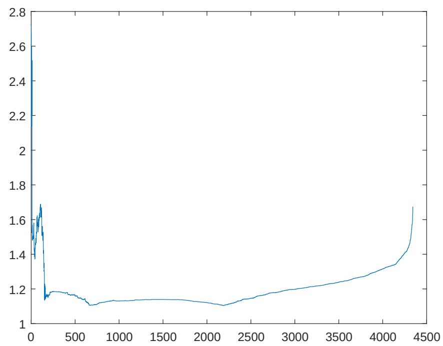
Similar to the first Rényi test measure, the second test measure appear to abruptly increase for lower values of k. Furthermore, the test measure tends towards infinity for higher values of k. The third Rényi test measure seems much stable and reliable. Although, it has a spike towards lower and higher values of k, it is almost fixed for most values of k. Both test measures do not reject the heavy-tails null hypothesis for any values of k of which the critical values are calculated.

Figure 6

(a) R2



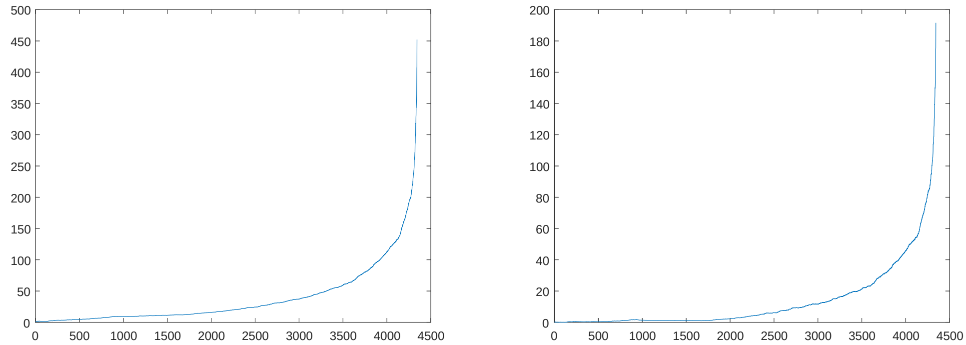
(b) R3



The last test measure of Rényi appears to be very similar to the BJ test measure. It is very stable for lower values of k, but explodes for the higher values. The KSI looks very similar with the same characteristics. Both test measures do not reject the heavy-tails null hypothesis for lower values of k. However, once k reaches 150 the KSI test measure rejects it, earlier than the KS and SC measure. This could be due to the higher type one error for the KSI test measure for higher values of k compared to other test measures.

Figure 7

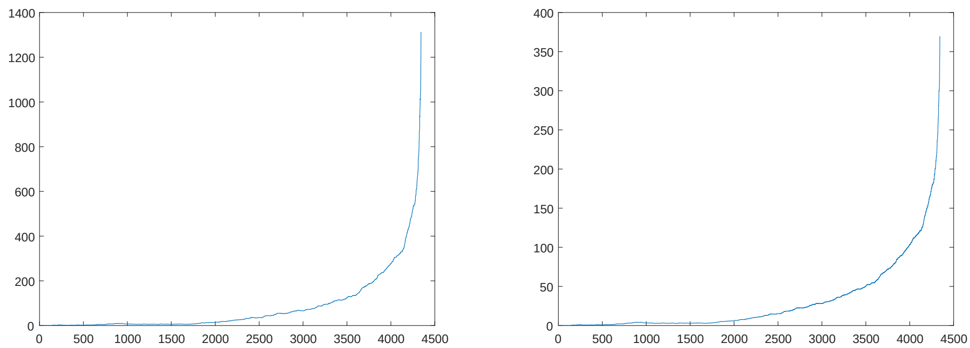
(a) R4 (b) KSI



Both the BJI and SCI are very similar, as the explode for higher values of k. Both test measures show similar rejection patterns to the RV data as the KS, KSI and SC test measures. They do not reject heavy-tails for lower values of k, but do reject it for values higher than 150. The test measure values of BJI, SC and SCI for the RV data do not reject heavy-tails for 99% confidence. This could indicate that these test measures might be better than KS and KSI test measures, as they do not change their conclusion on different levels of k.

Figure 8

(a) BJI (b) SCI



The test measures show the same pattern as with the analysis of the Danish fire loss data and internet traffic data described in Koning and Peng (2008). Initially, the measures seem to be relatively stable and react only to changes in the data. However, once k starts to increase above a certain bound around 1500-2000 the measures all explode. This is most likely due to either dependence in the data, which is available in the RV data, or because the assumption of $\lim_{n\to\infty} \sqrt{k}A(\frac{n}{k}) = 0$ does not hold. Similarly to Koning and Peng (2008) there are starting to emerge strange patterns in the data for values of k above 150, as here the measures start to reject the heavy-tail (see Table 17 and Table 18). Interesting is to see that the KS and SC measure do not explode as fast as other measures. Furthermore, the explosion of the test measure happens around the 3500th observation, while the explosion in the article of Koning and Peng (2008) happened around the 1500th observation. This could indicate that the type of data might play a role in the stability of the measure.

5 Discussion

The results indicate that the critical values of the test measures have a low relation to the underlying distribution if around 5% of the ordered observations are used for k. For higher values of k the type I error becomes too large for different distributions and reliability and validity problems could exist. Furthermore, critical values should be simulated for different sample sizes for all test measures as critical values are dependent on both the proportion of the ordered distribution used and the number of observations. If this is impossible or impractical the quadratic Berk-Jones test can be used, as it is the least dependent on the sample size. Moreover, the first Rényi measure has the Type I error closest to the corresponding critical value for different distributions and percentage of the ordered observations used. However, it lacks on power and stability. On the other hand, the quadratic estimated score test has the highest power. If the relation between the power and the type I error is the focus than the Berk-Jones test has the most added value. Of the measures based on Rényi, the fourth one performs generally the best due to lower type I errors and higher powers. Lastly, due to symmetry in the two-tailed distributions the critical values appear to be similar for lower values of k. Therefore, the current test measures are able to be applied to both two-tailed and one-tailed distributions.

5.1 Limitations

In this thesis the critical values are calculated using specific underlying distributions and compared to other specific distributions. As it appears that distribution can have an influence on the critical values for higher values of k, these critical values can be wrong and dependent on other distributions. Therefore, critical values are dependent on the researchers' ability to specify a very similar distribution with a heavy tail to the distribution of the sample in order to simulate the most useful critical values. Furthermore, no asymmetric two-tailed distribution (such as a skewed T-distribution) was used in this paper. Values for the different test measures could differ from values reported here and influence the reliability and validity of the test measures. Additionally, it could be possible that there are populations that only have a heavy-tail on one tail of the distribution. Important is to realize that the critical values are simulated, and not calculated using a mathematical distribution or asymptotic distributions. Therefore, minor differences can exist between the 'omnibus' critical values, which are independent on the underlying distribution for the simulation, and the simulated critical values. This could result in minor distortions in the results and conclusion.

5.2 Further Research

Firstly, in this thesis and past research critical values have been simulated from specific distributions. As these critical values appear to be dependent on this distribution, the critical values could be made from a variety of different distributions. This could potentially improve the type I error for larger values of k and improve the Power. Secondly, as indicated with the RV data, the type of data could influence the stability of the test measures. How this influences the measures and how the measures could be improved to decrease this effect could be investigated. Lastly, as the difference between the type I error and power is relatively low for the test measures related to the use of $G_k(r)$, it could be that test measures related to $G_k(r)$ are not able to differentiate adequately between heavy-tailed and non-heavy tailed distributions. Therefore, another test measure could be developed which does not use $G_k(r)$ and improves the ability to distinguish between types of distributions.

References

R.H. Berk and D.H. Jones. Relatively optimal combinations of test statistics. *Statist*, 5(3):158–162, 1978.

A.C. Davison and R.L. Smith. Models for exceedances over high thresholds. *Roy. Statist. Soc. Ser.*, 52(3):393–442, 1990.

L. de Haan and A. Ferreira. Extreme value theory: An introduction. *Pure Math.Statist*, 1(1):all, 2007.

L. de Haan and S. Resnich. Estimating the limit distribuion of multivariate extremes. *Comm. Statist. Stoachstic Models*, 9(2):275–309, 1993.

L. de Haan and S. Resnich. On asymptotic normality of the hill estimator. *Comm. Statist. Stoachstic Models*, 14(4):849–866, 1998.

L. de Haan and U. Stadmüller. Generalized regular variation of second order. *Pure Math.Statist*, 61(3):381–395, 1996.

J.H.J. Einmahl and I.W. McKeague. Empirical likelihood based hypothesis testing. *Bernoulli*, 9 (2):279–290, 2003.

B.M. Hill. A simple general approach to inference about the tail of a distribution. *Ann. Statist*, 3 (5):1163–1174, 1975.

N.L. Hjort and A. Koning. Tests for constancy of model parameters over time. *Nonparametric Statistics*, 14(2):113–132, 2002.

R.W. Katz, M.B. Parlange, and P. Naveau. Statistics of extremes in hydrology. *Advanced Water Resources*, 25(1):1287–1304, 2002.

A. Kolmogorov. Sulla determinazione empirica di una legge di distribuziona. *Giorn. Ist. Ital. Attuari.*, 4(1):83–91, 1933.

A.J. Koning and L. Peng. Goodness-of-fit tests for a heavy tailed distribution. *Statistical Planning and Inference*, 138(1):3960–3981, 2008.

A. Rényi. On the theory of order statistics. *Ann. Math. Statist.*, 39(1):755–771, 1953.

A. Rényi. On a group of problems in the theory of order statistics. *Selected Translation in Math. Statist. and Probab. Amer. Math. Soc.*, 13(1):289–298, 1973.

J.A. Wellner and V. Koltchinskii. A note on the asymptotic distribution of berk-jones type statistics under the null hypothesis. *High dimensional Probability, Progress of Probability*, 55(1):321–332, 2003.

A Appendix

Derivation I:
Maximum Likelihood Estimator of α in a Pareto Distribution

$$f(x,\alpha) = \frac{\alpha}{x_i^{\alpha+1}} \tag{57}$$

$$L(x,\alpha) = \prod_{i=1}^{n} \frac{\alpha}{x_i^{\alpha+1}}$$

$$\ln(L(x,\alpha)) = n\ln(\alpha) - (\alpha+1)\sum_{i=1}^{n}\ln(x_i)$$

$$\frac{\partial \ln(L(x,\alpha))}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n}\ln(x_i) = 0$$

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n}\ln(x_i)} \tag{58}$$

In this particular case, x_i is $\frac{X_{n-k+i}}{X_{n-k}}$.

$$Y_i = \frac{X_{n-k+i}}{X_{n-k}}$$

$$\begin{aligned}
KSI &= \int_1^\infty (\sqrt{k}(1 - G_k(r) - r^{-\hat{\alpha}}))^2 dG(r, \hat{\alpha}) \\
&= \int_1^\infty (\sqrt{k} KS(r, \hat{\alpha}))^2 dG(r, \hat{\alpha}) \\
&= \int_1^\infty k(1 - G_k(r) - r^{-\alpha})^2 \alpha r^{-\alpha-1} dr \\
&= \int_1^\infty k(1 + G_k(r)^2 + r^{-2\alpha} - 2G_k(r) - 2r^{-\alpha} + 2G_k(r) r^{-\alpha}) \alpha r^{-\alpha-1} dr \\
&= \int_1^\infty k(1 + r^{-2\alpha} - 2r^{-\alpha}) \alpha r^{-\alpha-1} dr + \int_1^\infty k(G_k(r)^2 - 2G_k(r) + 2G_k(r) r^{-\alpha}) \alpha r^{-\alpha-1} dr \\
&= [-r^{-\alpha} k - \frac{1}{3} k r^{-3\alpha} + k r^{-2\alpha}]_1^\infty \\
&+ \int_1^\infty [k((k^{-1} \sum_{i=1}^k I(Y_i \le r))^2)(\alpha r^{-\alpha-1}) - 2 \sum_{i=1}^k I(Y_i \le r)(\alpha r^{-\alpha-1}) + 2 \sum_{i=1}^k I(Y_i \le r)(\alpha r^{-2\alpha-1})] dr \\
&= [0 - 0 + 0] - [-k - \frac{1}{3} k + k] + \int_1^\infty k^{-1} (\sum_{i=1}^k I(Y_i \le r))^2 \alpha r^{-\alpha-1} dr \\
&- 2 \int_1^\infty \sum_{i=1}^k I(Y_i \le r) \alpha r^{-\alpha-1} dr + 2 \int_1^\infty \sum_{i=1}^k I(Y_i \le r) \alpha r^{-2\alpha-1} \\
&= \frac{1}{3} k + k^{-1} \sum_{i=1}^k \int_1^\infty I(Y_i \le r) \sum_{j=1}^k I(Y_j \le r) \alpha r^{-\alpha-1} dr \\
&- 2 \sum_{i=1}^k \int_1^\infty I(Y_i \le r) \alpha r^{-\alpha-1} dr + 2 \sum_{i=1}^k \int_1^\infty I(Y_i \le r) \alpha r^{-2\alpha-1} dr \\
&= \frac{1}{3} k + k^{-1} \sum_{i=1}^k \int_{Y_i}^\infty \sum_{j=1}^k I(Y_j \le r) \alpha r^{-\alpha-1} dr - 2 \sum_{i=1}^k \int_{Y_i}^\infty \alpha r^{-\alpha-1} dr \\
&+ 2 \sum_{i=1}^k \int_{Y_i}^\infty \alpha r^{-2\alpha-1} dr \\
&= \frac{1}{3} k + k^{-1} \sum_{i=1}^k \sum_{j=1}^k \int_{Y_i}^\infty I(Y_j \le r) \alpha r^{-\alpha-1} dr - 2 \sum_{i=1}^k [-r^{-\alpha}]_{Y_i}^\infty + 2 \sum_{i=1}^k [-\frac{1}{2} r^{-2\alpha}]_{Y_i}^\infty \\
&= \frac{1}{3} k + k^{-1} \sum_{i=1}^k \sum_{j=1}^k \int_{max(Y_i, Y_j)}^\infty \alpha r^{-\alpha-1} dr - 2 \sum_{i=1}^k (Y_i^{-\alpha}) + \sum_{i=1}^k (Y_i^{-2\alpha}) \\
&= \frac{1}{3} k + k^{-1} \sum_{i=1}^k \sum_{j=1}^k [-r^{-\alpha}]_{max(Y_i, Y_j)}^\infty - 2 \sum_{i=1}^k (Y_i^{-\alpha}) + \sum_{i=1}^k (Y_i^{-2\alpha}) \\
&= \frac{1}{3} k + k^{-1} \sum_{i=1}^k \sum_{j=1}^k (max(Y_i, Y_j))^{-\alpha} - 2 \sum_{i=1}^k (Y_i^{-\alpha}) + \sum_{i=1}^k (Y_i^{-2\alpha}) \qquad (59)
\end{aligned}$$

$$
\begin{aligned}
BJI &= \int_1^\infty kBJ(r,\hat{\alpha})dG(r,\hat{\alpha}) \\
&= \int_1^\infty k2(G_k(r)\ln(\frac{G_k(r))}{G(r,\hat{\alpha}}) + (1-G_k(r))\ln(\frac{1-G_k(r)}{1-G(r,\hat{\alpha})}))\alpha r^{-\alpha-1}dr \\
&= 2k\int_1^\infty (k^{-1}\sum_{i=1}^k I(Y_i \le r))\ln(\frac{(k^{-1}\sum_{i=1}^k I(Y_i \le r)}{1-r^{-\alpha}}) \\
&+ (k^{-1}\sum_{i=1}^k I(Y_i > r))\ln(\frac{(k^{-1}\sum_{i=1}^k I(Y_i > r)}{r^{-\alpha}}))\alpha r^{-\alpha-1}dr \\
&= 2\int_1^\infty (\sum_{i=1}^k I(Y_i \le r))(\ln(k^{-1}\sum_{i=1}^k I(Y_i \le r)) - \ln(1-r^{-\alpha}))\alpha r^{-\alpha-1}dr \\
&+ 2\int_1^\infty (\sum_{i=1}^k I(Y_i > r))(\ln(k^{-1}\sum_{i=1}^k I(Y_i > r)) - \ln(r^{-\alpha}))\alpha r^{-\alpha-1}dr \\
&= 2\int_1^\infty (\sum_{i=1}^k I(Y_i \le r))\ln(k^{-1}\sum_{i=1}^k I(Y_i \le r))\alpha r^{-\alpha-1}dr \\
&- 2\int_1^\infty (\sum_{i=1}^k I(Y_i \le r))\ln(1-r^{-\alpha})\alpha r^{-\alpha-1}dr \\
&+ 2\int_1^\infty (\sum_{i=1}^k I(Y_i > r))\ln(k^{-1}\sum_{i=1}^k I(Y_i > r))\alpha r^{-\alpha-1}dr \\
&- 2\int_1^\infty (\sum_{i=1}^k I(Y_i > r))\ln(r^{-\alpha})\alpha r^{-\alpha-1}dr
\end{aligned}
$$

These four sections of the integral are calculate as follows:

Section I:

$$
\begin{aligned}
&2\int_1^\infty (\sum_{i=1}^k I(Y_i \le r))\ln(k^{-1}\sum_{i=1}^k I(Y_i \le r))\alpha r^{-\alpha-1}dr \\
&= 2\int_1^{Y_1} (\sum_{i=1}^k I(Y_i \le r))\ln(k^{-1}\sum_{i=1}^k I(Y_i \le r))\alpha r^{-\alpha-1}dr \\
&+ 2\sum_{j=1}^{k-1}\int_{Y_j}^{Y_{j+1}} (\sum_{i=1}^k I(Y_i \le r))\ln(k^{-1}\sum_{i=1}^k I(Y_i \le r))\alpha r^{-\alpha-1}dr \\
&+ 2\int_{Y_k}^\infty (\sum_{i=1}^k I(Y_i \le r))\ln(k^{-1}\sum_{i=1}^k I(Y_i \le r))\alpha r^{-\alpha-1}drv \\
&= 2\int_1^{Y_1}(0)\ln(0)\alpha r^{-\alpha-1}dr + 2\sum_{j=1}^{k-1}\int_{Y_j}^{Y_{j+1}}(j)\ln(k^{-1}j)\alpha r^{-\alpha-1}dr + 2\int_{Y_k}^\infty (k)\ln(1)\alpha r^{-\alpha-1}dr \\
&= 0 + 2\sum_{j=1}^{k-1}(j)\ln(k^{-1}j)\int_{Y_j}^{Y_{j+1}} \alpha r^{-\alpha-1}dr + 0 \\
&= 2\sum_{j=1}^{k-1}(j)\ln(k^{-1}j)[-r^{-\alpha}]_{Y_j}^{Y_{j+1}} = 2\sum_{j=1}^{k-1}(j)\ln(k^{-1}j)(Y_j^{-\alpha} - Y_{j+1}^{-\alpha})
\end{aligned}
$$

Section II:

$$-2\int_1^{\infty}(\sum_{i=1}^{k} I(Y_i \le r))\ln(1-r^{-\alpha})\alpha r^{-\alpha-1}dr = -2\sum_{i=1}^{k}\int_{Y_i}^{\infty}\ln(1-r^{-\alpha})\alpha r^{-\alpha-1}dr$$

using Integration by parts and the Integral substitution with:

$$u = \ln(1-r^{-\alpha}) \; u' = \frac{\alpha r^{-\alpha-1}}{1-r^{-\alpha}} \; v = 1-r^{-\alpha} \; v' = \alpha r^{-\alpha-1}$$

$$-2\sum_{i=1}^{k}[\ln(1-r^{-\alpha})(1-r^{-\alpha}) - \int \frac{\alpha r^{-\alpha-1}}{1-r^{-\alpha}}(1-r^{-\alpha})dr]_{Y_i}^{\infty}$$

$$= -2\sum_{i=1}^{k}[(1-r^{-\alpha})\ln(1-r^{-\alpha}) + r^{-\alpha})]_{Y_i}^{\infty}$$

$$= -2\sum_{i=1}^{k}(1*\ln(1)+0-((1-Y_i^{-\alpha})\ln(1-Y_i^{-\alpha})+Y_i^{-\alpha}))$$

$$= -2\sum_{i=1}^{k}(-1+Y_i^{-\alpha})\ln(1-Y_i^{-\alpha}) - Y_i^{-\alpha}$$

Section III:

$$2\int_1^{\infty}(\sum_{i=1}^{k} I(Y_i > r))\ln(k^{-1}\sum_{i=1}^{k} I(Y_i > r))\alpha r^{-\alpha-1}dr$$

$$= 2\int_1^{Y_1}(\sum_{i=1}^{k} I(Y_i > r))\ln(k^{-1}\sum_{i=1}^{k} I(Y_i > r))\alpha r^{-\alpha-1}dr$$

$$+ 2\sum_{j=1}^{k-1}\int_{Y_j}^{Y_{j+1}}(\sum_{i=1}^{k} I(Y_i > r))\ln(k^{-1}\sum_{i=1}^{k} I(Y_i > r))\alpha r^{-\alpha-1}dr$$

$$+ 2\int_{Y_k}^{\infty}(\sum_{i=1}^{k} I(Y_i > r))\ln(k^{-1}\sum_{i=1}^{k} I(Y_i > r))\alpha r^{-\alpha-1}dr$$

$$= 2\int_1^{Y_1}(k)\ln(1)\alpha r^{-\alpha-1}dr + 2\sum_{j=1}^{k-1}\int_{Y_j}^{Y_{j+1}}(k-j)\ln(k^{-1}(k-j))\alpha r^{-\alpha-1}dr$$

$$+ 2\int_{Y_k}^{\infty}(0)\ln(0)\alpha r^{-\alpha-1}dr$$

$$= 0 + 2\sum_{j=1}^{k-1}(k-j)\ln(k^{-1}(k-j))[-r^{-\alpha}]_{Y_j}^{Y_{j+1}}dr + 0$$

$$= 2\sum_{j=1}^{k-1}(k-j)\ln(k^{-1}(k-j))(Y_j^{-\alpha} - Y_{j+1}^{-\alpha})$$

Section IV:

$$-2\int_1^\infty (\sum_{i=1}^{k} I(Y_i > r))\ln(r^{-\alpha})\alpha r^{-\alpha-1}dr$$

$$= -2\sum_{i=1}^{k}\int_1^{Y_i}\ln(r^{-\alpha})\alpha r^{-\alpha-1}dr$$

using Integration by Parts with:

$$u = \ln(r^{-\alpha})\ u' = -\frac{\alpha}{r}\ v = -r^{-\alpha}\ v' = \alpha r^{-\alpha-1}$$

$$-2\sum_{i=1}^{k}[-r^{-\alpha}\ln(r^{-\alpha}) - \alpha\int r^{-\alpha-1}dr]_1^{Y_i}$$

$$= -2\sum_{i=1}^{k}[-r^{-\alpha}\ln(r^{-\alpha}) + r^{-\alpha}]_1^{Y_i} = -2\sum_{i=1}^{k}(-Y_i^{-\alpha}\ln(Y_i^{-\alpha}) + Y_i^{-\alpha} - 1)$$

All these sections make BJI become:

$$BJI = 2\sum_{j=1}^{k-1}(j)\ln(k^{-1}j)(Y_j^{-\alpha} - Y_{j+1}^{-\alpha}) - 2\sum_{i=1}^{k}((-1 + Y_i^{-\alpha})\ln(1 - Y_i^{-\alpha}) - Y_i^{-\alpha})$$

$$+ 2\sum_{j=1}^{k-1}(k-j)\ln(k^{-1}(k-j))(Y_j^{-\alpha} - Y_{j+1}^{-\alpha}) - 2\sum_{i=1}^{k}(-Y_i^{-\alpha}\ln(Y_i^{-\alpha}) + Y_i^{-\alpha} - 1)$$

$$= 2\sum_{j=1}^{k-1}((j)\ln(k^{-1}j) + (k-j)\ln(k^{-1}(k-j)))(Y_j^{-\alpha} - Y_{j+1}^{-\alpha})$$

$$-2\sum_{i=1}^{k}((-1 + Y_i^{-\alpha})\ln(1 - Y_i^{-\alpha}) - Y_i^{-\alpha}\ln(Y_i^{-\alpha}) - 1) \quad (60)$$

$$
\begin{aligned}
SCI &= \int_1^\infty (\sqrt{k} SC(r,\hat{\alpha}))^2 dG(r,\hat{\alpha}) \\
&= \int_1^\infty kSC^2 \alpha r^{-\alpha-1} dr = \int_1^\infty k(G_k(r) - \alpha k^{-1}\alpha \sum_{i=1}^k \ln(min(r,Y_i)))^2 \alpha r^{-\alpha-1} dr \\
&= \int_1^\infty (kG_k(r)^2 + \alpha^2 k^{-1}(\sum_{i=1}^k \ln(min(r,Y_i)))^2 - 2G_k(r)\alpha \sum_{i=1}^k \ln(min(r,Y_i)))\alpha r^{-\alpha-1} dr \\
&= \int_1^\infty k^{-1}(\sum_{i=1}^k I(Y_i \le r))^2 \alpha r^{-\alpha-1} dr + \alpha^2 k^{-1} \int_1^\infty (\sum_{i=1}^k \ln(min(r,Y_i)))^2 \alpha r^{-\alpha-1} dr \\
&- 2\int_1^\infty k^{-1} \sum_{i=1}^k I(Y_i \le r)\alpha \sum_{j=1}^k \ln(min(r,Y_j))\alpha r^{-\alpha-1} dr \\
&= k^{-1} \sum_{i=1}^k \int_{Y_i}^\infty \sum_{j=1}^k I(Y_j \le r)\alpha r^{-\alpha-1} dr \\
&+ \alpha^2 k^{-1} \sum_{i=1}^k \int_1^\infty \ln(min(r,Y_i)) \sum_{j=1}^k \ln(min(r,Y_j))\alpha r^{-\alpha-1} dr \\
&- 2\alpha k^{-1} \sum_{i=1}^k \int_{Y_i}^\infty \sum_{j=1}^k \ln(min(r,Y_j))\alpha r^{-\alpha-1} dr \\
&= k^{-1} \sum_{i=1}^k \sum_{j=1}^k \int_{max(Y_i,Y_j)}^\infty \alpha r^{-\alpha-1} dr + \alpha^2 k^{-1} \sum_{i=1}^k \sum_{j=1}^k \int_1^\infty \ln(min(r,Y_i)) \ln(min(r,Y_j))\alpha r^{-\alpha-1} dr \\
&- 2\alpha k^{-1} \sum_{i=1}^k \sum_{j=1}^k \int_{Y_i}^\infty \ln(min(r,Y_j))\alpha r^{-\alpha-1} dr
\end{aligned}
$$

Part I

$$
\begin{aligned}
&k^{-1} \sum_{i=1}^k \sum_{j=1}^k \int_{max(Y_i,Y_j)}^\infty \alpha r^{-\alpha-1} dr \\
&= k^{-1} \sum_{i=1}^k \sum_{j=1}^k (max(Y_i,Y_j))^{-\alpha}
\end{aligned}
$$

Part II and using the sum rule together with integration by parts with

$$\alpha^2 k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} \int_1^\infty \ln(min(r, Y_i)) \ln(min(r, Y_j)) \alpha r^{-\alpha-1} dr$$

$$= \alpha^2 k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} (\int_1^{min(Y_i,Y_j)} \ln(r) \ln(r) \alpha r^{-\alpha-1} dr + \int_{min(Y_i,Y_j)}^{max(Y_i,Y_j)} \ln(min(Y_i, Y_j)) \ln(r) \alpha r^{-\alpha-1} dr$$

$$+ \int_{max(Y_i,Y_j)}^{\infty} \ln(Y_i) \ln(Y_j) \alpha r^{-\alpha-1} dr)$$

Section I:

$$u = \ln(r) \ u' = \frac{1}{r} \ v = -r^{-\alpha} \ln(r) - \frac{r^{-\alpha}}{\alpha} \ v' = \ln(r) \alpha r^{-\alpha-1}$$

$$\int_1^{min(Y_i,Y_j)} \ln(r) \ln(r) \alpha r^{-\alpha-1} dr = [\ln(r)(-\frac{-r^{-\alpha}}{\alpha} - r^{-\alpha} \ln(r)) + \int r^{-\alpha-1} \ln(r) + \frac{r^{-\alpha-1}}{\alpha} dr]_1^{min(Y_i,Y_j)}$$

$$= [\ln(r)(-\frac{-r^{-\alpha}}{\alpha} - r^{-\alpha} \ln(r)) + \frac{-\alpha r^{-\alpha} \ln(r) - r^{-\alpha}}{\alpha^2} - \frac{r^{-\alpha}}{\alpha^2}]_1^{min(Y_i,Y_j)}$$

$$= [(\ln(min(Y_i, Y_j))(-\frac{-min(Y_i, Y_j)^{-\alpha}}{\alpha} - min(Y_i, Y_j)^{-\alpha} \ln(min(Y_i, Y_j)))$$

$$+ \frac{-\alpha * min(Y_i, Y_j)^{-\alpha} \ln(min(Y_i, Y_j)) - min(Y_i, Y_j)^{-\alpha}}{\alpha^2} - \frac{min(Y_i, Y_j)^{-\alpha}}{\alpha^2}) - (0 + \frac{-1}{\alpha^2} - \frac{1}{\alpha^2}]$$

$$= -\frac{min(Y_i, Y_j)^{-\alpha} \ln(min(Y_i, Y_j))}{\alpha} - min(Y_i, Y_j)^{-\alpha} \ln^2(min(Y_i, Y_j))$$

$$- \frac{\alpha * min(Y_i, Y_j)^{-\alpha} \ln(min(Y_i, Y_j)) + 2 * min(Y_i, Y_j)^{-\alpha}}{\alpha^2} + \frac{2}{\alpha^2}$$

Section II:

$$u = \ln(r) \ u' = \frac{1}{r} \ v = -r^{-\alpha} \ v' = \alpha r^{-\alpha-1}$$

$$\int_{min(Y_i,Y_j)}^{max(Y_i,Y_j)} \ln(min(Y_i, Y_j)) \ln(r) \alpha r^{-\alpha-1} dr = \ln(min(Y_i, Y_j))[-r^{-\alpha} \ln(r) + \int r^{-\alpha-1} dr]_{min(Y_i,Y_j)}^{max(Y_i,Y_j)}$$

$$= \ln(min(Y_i, Y_j))[-r^{-\alpha} \ln(r) - \frac{r^{-\alpha}}{\alpha}]_{min(Y_i,Y_j)}^{max(Y_i,Y_j)}$$

$$= \ln(min(Y_i, Y_j))(-max(Y_i, Y_j)^{-\alpha} \ln(max(Y_i, Y_j)) - \frac{max(Y_i, Y_j)^{-\alpha}}{\alpha})$$

$$+ \ln(min(Y_i, Y_j))(min(Y_i, Y_j)^{-\alpha} \ln(min(Y_i, Y_j)) + \frac{min(Y_i, Y_j)^{-\alpha}}{\alpha})$$

Section III:

$$\int_{max(Y_i,Y_j)}^{\infty} \ln(Y_i) \ln(Y_j) \alpha r^{-\alpha-1} dr = [-\ln(Y_i) \ln(Y_j) r^{-\alpha}]_{max(Y_i,Y_j)}^{\infty}$$

$$= \ln(Y_i) \ln(Y_j) max(Y_i, Y_j)^{-\alpha}$$

Part III:

$$- 2\alpha k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} \int_{Y_i}^{\infty} \ln(min(r, Y_j)) \alpha r^{-\alpha-1} dr$$
$$= -2\alpha k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} (\int_{Y_i}^{max(Y_i,Y_j)} \ln(r) \alpha r^{-\alpha-1} dr + \int_{max(Y_i,Y_j)}^{\infty} \ln(Y_j) \alpha r^{-\alpha-1} dr)$$
$$= -2\alpha k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} ([-r^{-\alpha} \ln(r) - \frac{r^{-\alpha}}{\alpha}]_{Y_i}^{max(Y_i,Y_j)} + [-\ln(Y_j) r^{-\alpha}]_{max(Y_i,Y_j)}^{\infty}$$
$$= -2\alpha k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} ([-max(Y_i, Y_j)^{-\alpha} \ln(max(Y_i, Y_j)) - \frac{max(Y_i, Y_j)^{-\alpha}}{\alpha} - (-Y_i^{-\alpha} \ln(Y_i) - \frac{Y_i^{-\alpha}}{\alpha})]$$
$$+ \ln(Y_j) max(Y_i, Y_j)^{-\alpha})$$

This results in the following combined integral:

$$SCI = k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} (max(Y_i, Y_j))^{-\alpha} + \alpha^2 k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} (- \frac{min(Y_i, Y_j)^{-\alpha} \ln(min(Y_i, Y_j))}{\alpha}$$
$$- min(Y_i, Y_j)^{-\alpha} \ln^2(min(Y_i, Y_j)) - \frac{\alpha * min(Y_i, Y_j)^{-\alpha} \ln(min(Y_i, Y_j)) + 2 * min(Y_i, Y_j)^{-\alpha}}{\alpha^2}$$
$$+ \frac{2}{\alpha^2} + \ln(min(Y_i, Y_j))(-max(Y_i, Y_j)^{-\alpha} \ln(max(Y_i, Y_j)) - \frac{max(Y_i, Y_j)^{-\alpha}}{\alpha})$$
$$+ \ln(min(Y_i, Y_j))(min(Y_i, Y_j)^{-\alpha} \ln(min(Y_i, Y_j)) + \frac{min(Y_i, Y_j)^{-\alpha}}{\alpha}) + \ln(Y_i) \ln(Y_j) max(Y_i, Y_j)^{-\alpha})$$
$$- 2\alpha k^{-1} \sum_{i=1}^{k} \sum_{j=1}^{k} ([-max(Y_i, Y_j)^{-\alpha} \ln(max(Y_i, Y_j)) - \frac{max(Y_i, Y_j)^{-\alpha}}{\alpha} - (-Y_i^{-\alpha} \ln(Y_i) - \frac{Y_i^{-\alpha}}{\alpha})]$$
$$+ \ln(Y_j) max(Y_i, Y_j)^{-\alpha}) \quad (61)$$

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	1,172	0,330	9,466	2,075	1,447	0,714	90,944	5,352	2,961	3,633
30	1,202	0,333	9,833	2,055	1,487	0,716	95,429	5,732	3,568	3,947
40	1,209	0,330	10,014	2,044	1,506	0,722	95,237	5,945	3,969	4,226
50	1,216	0,335	10,171	2,023	1,521	0,728	93,991	6,002	4,309	4,418
60	1,217	0,340	10,264	2,000	1,537	0,720	96,310	6,229	4,571	4,557
70	1,226	0,332	10,410	1,990	1,541	0,738	96,946	6,220	4,817	4,715
80	1,230	0,335	10,502	1,975	1,543	0,732	104,161	6,244	4,947	4,816
90	1,231	0,336	10,600	1,998	1,551	0,737	99,406	6,267	5,085	4,933
100	1,236	0,341	10,637	2,009	1,557	0,737	93,108	6,386	5,170	5,019
110	1,238	0,335	10,635	1,995	1,569	0,739	98,392	6,363	5,370	5,064
120	1,246	0,334	10,728	1,985	1,562	0,741	100,458	6,384	5,427	5,170
130	1,251	0,336	10,682	1,995	1,562	0,742	96,588	6,464	5,424	5,195
140	1,238	0,333	10,912	1,985	1,554	0,732	108,452	6,437	5,569	5,312
150	1,249	0,335	10,933	1,984	1,569	0,735	101,279	6,527	5,446	5,359
160	1,248	0,337	10,950	1,978	1,577	0,733	93,507	6,489	5,617	5,363
170	1,251	0,339	11,067	1,991	1,578	0,723	101,185	6,500	5,573	5,427
180	1,252	0,337	10,983	2,004	1,571	0,729	97,086	6,490	5,650	5,504
190	1,252	0,337	11,089	2,029	1,574	0,736	99,884	6,545	5,626	5,580
200	1,254	0,337	10,976	1,996	1,579	0,742	95,508	6,614	5,561	5,593

Table 7: Critical Values for 99% Confidence Level Lévy

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	1,192	0,329	9,397	2,090	1,462	0,699	96,625	5,311	3,055	3,665
30	1,199	0,332	9,840	2,077	1,498	0,721	101,233	5,769	3,534	3,984
40	1,217	0,335	9,846	2,022	1,502	0,724	92,249	5,998	3,930	4,208
50	1,222	0,338	9,971	2,043	1,513	0,733	96,087	6,124	4,162	4,465
60	1,231	0,337	10,199	2,045	1,540	0,734	96,428	6,246	4,353	4,639
70	1,240	0,335	10,273	2,044	1,536	0,731	95,096	6,288	4,674	4,783
80	1,249	0,337	10,305	2,070	1,552	0,737	93,403	6,465	4,715	4,940
90	1,257	0,340	10,368	2,074	1,560	0,727	98,906	6,504	4,599	5,057
100	1,265	0,345	10,372	2,071	1,570	0,731	94,923	6,603	4,780	5,207
110	1,283	0,350	10,460	2,165	1,582	0,738	90,720	6,731	4,668	5,287
120	1,290	0,356	10,572	2,175	1,597	0,739	90,183	6,728	4,663	5,449
130	1,308	0,369	10,687	2,264	1,625	0,735	91,666	6,902	4,514	5,577
140	1,325	0,384	10,798	2,359	1,653	0,737	90,534	6,919	4,484	5,707
150	1,363	0,405	10,895	2,475	1,680	0,742	90,409	7,151	4,365	5,854
160	1,383	0,431	11,075	2,592	1,715	0,741	88,201	7,128	4,214	5,988
170	1,413	0,451	11,310	2,753	1,760	0,732	87,297	7,350	4,164	6,156
180	1,460	0,483	11,702	2,945	1,812	0,734	84,959	7,502	4,017	6,369
190	1,495	0,512	12,107	3,094	1,867	0,741	87,095	7,638	3,866	6,500
200	1,527	0,548	12,617	3,358	1,935	0,731	79,751	7,640	3,739	6,691

Table 8: Critical Values for 99% Confidence Level Cauchy

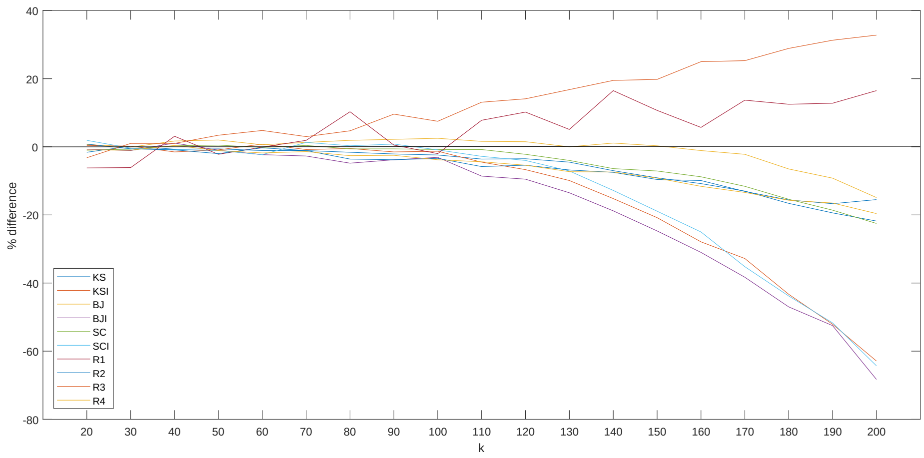
k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	-1,2 %	0,8 %	0,0 %	-0,3 %	0,3 %	1,0 %	-4,1 %	-0,2 %	-6,8 %	0,0 %
30	-0,5 %	0,2 %	0,4 %	-0,4 %	-0,3 %	0,8 %	-0,1 %	-0,3 %	-0,3 %	-0,6 %
40	-0,8 %	-0,3 %	0,5 %	0,1 %	-0,3 %	0,0 %	0,7 %	-0,6 %	0,4 %	-0,9 %
50	-0,7 %	-0,2 %	0,7 %	-1,2 %	0,6 %	0,4 %	3,1 %	-1,5 %	2,1 %	-1,0 %
60	-1,1 %	-0,1 %	-0,5 %	-1,8 %	-0,1 %	-2,1 %	1,2 %	-1,6 %	3,6 %	-1,0 %
70	-1,3 %	-0,1 %	1,1 %	-1,7 %	-0,3 %	-0,5 %	-1,6 %	-1,8 %	3,1 %	-1,5 %
80	-1,3 %	-1,9 %	1,5 %	-2,3 %	-0,1 %	0,9 %	6,0 %	-2,6 %	3,5 %	-1,9 %
90	-2,2 %	-1,6 %	1,7 %	-3,9 %	-1,1 %	-2,0 %	1,3 %	-3,4 %	6,0 %	-3,3 %
100	-2,8 %	-2,6 %	2,0 %	-4,2 %	-0,8 %	-2,4 %	2,1 %	-3,9 %	7,1 %	-4,2 %
110	-3,2 %	-5,3 %	1,6 %	-6,5 %	-1,4 %	-5,3 %	3,7 %	-4,7 %	9,8 %	-4,8 %
120	-3,3 %	-7,5 %	2,1 %	-9,1 %	-1,9 %	-7,2 %	8,8 %	-5,0 %	9,7 %	-6,4 %
130	-5,4 %	-10,5 %	0,2 %	-12,1 %	-4,1 %	-11,4 %	6,7 %	-6,6 %	12,9 %	-7,5 %
140	-6,6 %	-14,8 %	0,1 %	-16,3 %	-5,2 %	-16,2 %	9,0 %	-7,7 %	14,6 %	-9,6 %
150	-9,0 %	-21,8 %	-0,7 %	-21,8 %	-6,6 %	-23,6 %	11,3 %	-10,0 %	17,5 %	-10,5 %
160	-11,3 %	-26,8 %	-3,3 %	-29,0 %	-9,1 %	-31,0 %	9,7 %	-10,9 %	18,1 %	-12,9 %
170	-13,9 %	-33,9 %	-5,9 %	-36,9 %	-12,3 %	-41,4 %	12,5 %	-12,5 %	20,8 %	-14,9 %
180	-16,9 %	-47,1 %	-9,6 %	-48,0 %	-15,9 %	-51,9 %	13,1 %	-14,2 %	23,4 %	-17,8 %
190	-20,2 %	-58,2 %	-12,2 %	-58,2 %	-19,9 %	-63,8 %	14,5 %	-16,4 %	26,0 %	-19,2 %
200	-23,4 %	-72,1 %	-18,0 %	-73,1 %	-24,8 %	-80,8 %	16,6 %	-16,6 %	28,0 %	-22,4 %

Table 9: Relativie Differences Critical Values for 95% Confidence Level

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	-1,6 %	0,4 %	0,7 %	-0,7 %	-1,0 %	1,9 %	-6,2 %	0,8 %	-3,2 %	-0,9 %
30	0,2 %	0,2 %	-0,1 %	-1,1 %	-0,7 %	-0,4 %	-6,1 %	-0,6 %	1,0 %	-1,0 %
40	-0,7 %	-1,5 %	1,7 %	1,1 %	0,3 %	-0,8 %	3,1 %	-0,9 %	1,0 %	0,4 %
50	-0,6 %	-0,9 %	2,0 %	-1,0 %	0,5 %	-1,1 %	-2,2 %	-2,0 %	3,4 %	-1,1 %
60	-1,1 %	0,8 %	0,6 %	-2,3 %	-0,2 %	-2,3 %	-0,1 %	-0,3 %	4,8 %	-1,8 %
70	-1,1 %	-0,8 %	1,3 %	-2,7 %	0,4 %	1,3 %	1,9 %	-1,1 %	3,0 %	-1,4 %
80	-1,6 %	-0,5 %	1,9 %	-4,8 %	-0,6 %	0,3 %	10,3 %	-3,6 %	4,7 %	-2,6 %
90	-2,1 %	-1,5 %	2,2 %	-3,8 %	-0,6 %	0,8 %	0,5 %	-3,8 %	9,6 %	-2,5 %
100	-2,4 %	-1,3 %	2,5 %	-3,1 %	-0,8 %	-0,9 %	-2,0 %	-3,4 %	7,5 %	-3,8 %
110	-3,6 %	-4,5 %	1,6 %	-8,6 %	-0,8 %	-2,8 %	7,8 %	-5,8 %	13,1 %	-4,4 %
120	-3,5 %	-6,7 %	1,5 %	-9,5 %	-2,2 %	-4,0 %	10,2 %	-5,4 %	14,1 %	-5,4 %
130	-4,5 %	-9,9 %	0,0 %	-13,5 %	-4,0 %	-7,1 %	5,1 %	-6,8 %	16,8 %	-7,3 %
140	-7,0 %	-15,2 %	1,1 %	-18,8 %	-6,4 %	-12,8 %	16,5 %	-7,5 %	19,5 %	-7,4 %
150	-9,1 %	-20,8 %	0,3 %	-24,7 %	-7,1 %	-18,9 %	10,7 %	-9,6 %	19,8 %	-9,2 %
160	-10,8 %	-27,9 %	-1,1 %	-31,0 %	-8,8 %	-25,0 %	5,7 %	-9,9 %	25,0 %	-11,6 %
170	-13,0 %	-32,8 %	-2,2 %	-38,3 %	-11,6 %	-35,2 %	13,7 %	-13,1 %	25,3 %	-13,4 %
180	-16,6 %	-43,3 %	-6,5 %	-47,0 %	-15,4 %	-43,8 %	12,5 %	-15,6 %	28,9 %	-15,7 %
190	-19,4 %	-52,1 %	-9,2 %	-52,5 %	-18,6 %	-51,7 %	12,8 %	-16,7 %	31,3 %	-16,5 %
200	-21,8 %	-62,9 %	-14,9 %	-68,3 %	-22,5 %	-64,3 %	16,5 %	-15,5 %	32,8 %	-19,6 %

Table 10: Relative Differences Critical Values at 99% Confidence Level

Figure 9: Relative Differences Critical Values at 99% Confidence Level



k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0,019	0,012	0,008	0,016	0,022	0,010	0,010	0,016	0,095	0,011
30	0,023	0,016	0,007	0,022	0,027	0,015	0,008	0,018	0,045	0,019
40	0,029	0,024	0,008	0,029	0,038	0,023	0,009	0,020	0,023	0,028
50	0,038	0,031	0,008	0,038	0,048	0,030	0,009	0,022	0,012	0,040
60	0,047	0,042	0,009	0,053	0,064	0,046	0,009	0,022	0,006	0,052
70	0,061	0,062	0,012	0,069	0,082	0,065	0,009	0,025	0,004	0,074
80	0,076	0,081	0,015	0,092	0,108	0,089	0,008	0,026	0,002	0,091
90	0,096	0,107	0,019	0,118	0,134	0,126	0,009	0,028	0,002	0,112
100	0,121	0,136	0,027	0,152	0,167	0,167	0,008	0,028	0,001	0,140
110	0,144	0,175	0,036	0,180	0,209	0,208	0,007	0,031	0,001	0,172
120	0,174	0,219	0,051	0,224	0,253	0,263	0,008	0,033	0,001	0,196
130	0,204	0,262	0,073	0,258	0,296	0,322	0,009	0,032	0,001	0,235
140	0,245	0,307	0,095	0,312	0,351	0,383	0,008	0,034	0,001	0,262
150	0,270	0,354	0,126	0,350	0,398	0,442	0,008	0,034	0,001	0,299
160	0,311	0,391	0,166	0,400	0,450	0,497	0,009	0,035	0,001	0,341
170	0,345	0,439	0,206	0,441	0,493	0,555	0,007	0,034	0,001	0,374
180	0,376	0,480	0,250	0,475	0,533	0,598	0,008	0,035	0,001	0,405
190	0,418	0,524	0,289	0,518	0,577	0,635	0,007	0,037	0,001	0,437
200	0,459	0,559	0,343	0,556	0,609	0,670	0,008	0,039	0,001	0,470

Table 11: Type I errors n=1000 with 99% Confidence Level Critical values

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0,026	0,018	0,006	0,023	0,017	0,017	0,009	0,022	0,061	0,013
30	0,037	0,031	0,007	0,038	0,028	0,028	0,008	0,024	0,024	0,030
40	0,050	0,049	0,009	0,060	0,048	0,050	0,009	0,027	0,010	0,054
50	0,073	0,073	0,011	0,083	0,072	0,079	0,009	0,030	0,005	0,084
60	0,098	0,106	0,015	0,119	0,100	0,115	0,008	0,030	0,003	0,126
70	0,129	0,153	0,022	0,167	0,146	0,173	0,009	0,036	0,002	0,170
80	0,163	0,201	0,034	0,210	0,196	0,231	0,009	0,035	0,002	0,227
90	0,205	0,255	0,050	0,271	0,249	0,305	0,008	0,037	0,002	0,284
100	0,253	0,316	0,081	0,330	0,305	0,375	0,009	0,040	0,002	0,347
110	0,295	0,375	0,115	0,383	0,365	0,436	0,008	0,041	0,002	0,413
120	0,346	0,440	0,167	0,446	0,428	0,503	0,008	0,042	0,002	0,461
130	0,386	0,491	0,223	0,501	0,485	0,562	0,008	0,041	0,002	0,517
140	0,441	0,542	0,280	0,550	0,537	0,609	0,007	0,045	0,002	0,558
150	0,478	0,585	0,348	0,594	0,579	0,648	0,008	0,042	0,002	0,599
160	0,519	0,620	0,408	0,634	0,617	0,682	0,008	0,047	0,002	0,639
170	0,561	0,654	0,456	0,667	0,652	0,714	0,007	0,046	0,002	0,666
180	0,587	0,686	0,506	0,695	0,684	0,736	0,008	0,047	0,002	0,690
190	0,626	0,712	0,549	0,716	0,711	0,752	0,007	0,046	0,002	0,712
200	0,655	0,733	0,594	0,740	0,732	0,768	0,007	0,050	0,002	0,729

Table 12: Power at n=1000 with 99% Confidence Level Critical Values

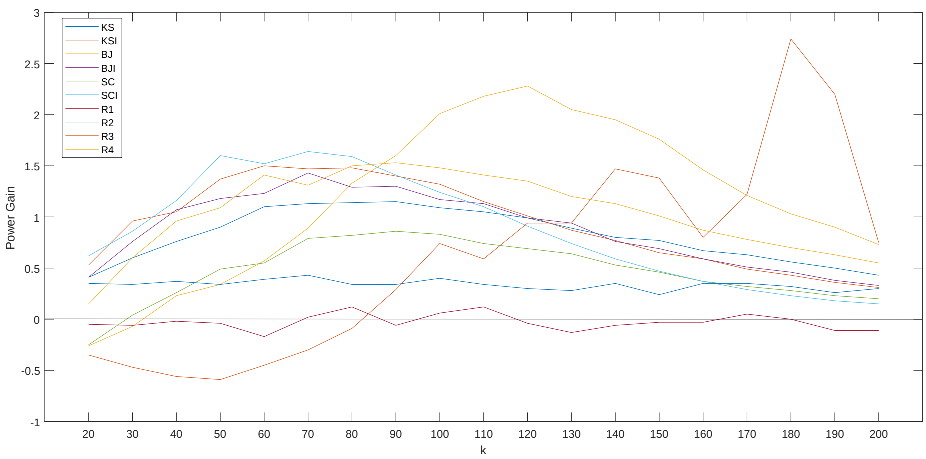
k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0,35	0,38	-0,09	0,38	-0,04	0,47	-0,09	0,21	-0,24	0,21
30	0,43	0,59	0,03	0,52	0,18	0,61	-0,05	0,26	-0,38	0,44
40	0,57	0,70	0,18	0,68	0,35	0,83	-0,09	0,22	-0,47	0,67
50	0,63	0,83	0,38	0,77	0,48	0,93	-0,03	0,25	-0,51	0,76
60	0,73	0,87	0,56	0,80	0,51	0,93	-0,08	0,29	-0,48	0,88
70	0,70	0,87	0,79	0,83	0,57	0,88	-0,07	0,27	-0,45	0,82
80	0,71	0,83	0,98	0,79	0,58	0,84	-0,02	0,21	-0,32	0,85
90	0,66	0,77	1,06	0,75	0,55	0,75	-0,04	0,23	-0,10	0,82
100	0,63	0,67	1,14	0,65	0,50	0,64	-0,03	0,24	0,10	0,76
110	0,57	0,60	1,14	0,60	0,41	0,51	-0,01	0,20	0,35	0,69
120	0,52	0,51	1,05	0,51	0,35	0,41	-0,03	0,21	0,57	0,64
130	0,46	0,43	0,91	0,44	0,31	0,33	-0,04	0,19	0,63	0,56
140	0,41	0,35	0,78	0,36	0,25	0,27	-0,06	0,23	0,87	0,51
150	0,36	0,31	0,67	0,30	0,21	0,20	-0,06	0,19	0,71	0,44
160	0,31	0,25	0,57	0,26	0,16	0,16	-0,03	0,23	1,11	0,38
170	0,28	0,21	0,48	0,21	0,13	0,12	-0,04	0,18	1,26	0,33
180	0,24	0,17	0,42	0,18	0,10	0,09	-0,02	0,18	1,18	0,29
190	0,21	0,14	0,35	0,15	0,08	0,07	-0,06	0,15	1,78	0,26
200	0,18	0,11	0,29	0,12	0,06	0,05	-0,06	0,18	1,85	0,22

Table 13: Relative Power Gain with critical values at 95% Confidence Level

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0,41	0,53	-0,26	0,41	-0,25	0,62	-0,05	0,35	-0,35	0,15
30	0,60	0,96	-0,07	0,76	0,04	0,86	-0,06	0,34	-0,47	0,60
40	0,76	1,05	0,23	1,07	0,26	1,16	-0,02	0,37	-0,56	0,96
50	0,90	1,37	0,34	1,18	0,49	1,60	-0,04	0,34	-0,59	1,09
60	1,10	1,50	0,57	1,23	0,55	1,52	-0,17	0,39	-0,45	1,41
70	1,13	1,47	0,89	1,43	0,79	1,64	0,02	0,43	-0,30	1,31
80	1,14	1,48	1,33	1,29	0,82	1,59	0,12	0,34	-0,09	1,50
90	1,15	1,40	1,60	1,30	0,86	1,41	-0,06	0,34	0,29	1,53
100	1,09	1,32	2,01	1,17	0,83	1,24	0,06	0,40	0,74	1,48
110	1,05	1,15	2,18	1,13	0,74	1,10	0,12	0,34	0,59	1,41
120	0,99	1,01	2,28	0,99	0,69	0,91	-0,04	0,30	0,94	1,35
130	0,89	0,87	2,05	0,94	0,64	0,74	-0,13	0,28	0,94	1,20
140	0,80	0,77	1,95	0,76	0,53	0,59	-0,06	0,35	1,47	1,13
150	0,77	0,65	1,76	0,69	0,46	0,47	-0,03	0,24	1,38	1,01
160	0,67	0,59	1,46	0,59	0,37	0,37	-0,03	0,35	0,80	0,87
170	0,63	0,49	1,21	0,51	0,32	0,29	0,05	0,35	1,22	0,78
180	0,56	0,43	1,03	0,46	0,28	0,23	0,00	0,32	2,74	0,70
190	0,50	0,36	0,90	0,38	0,23	0,18	-0,11	0,26	2,20	0,63
200	0,43	0,31	0,73	0,33	0,20	0,15	-0,11	0,30	0,75	0,55

Table 14: Relative Power Gain with 99% Confidence Level Critical Values

Figure 10: Power Gain with 99% Confidence Level Critical Values



n	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
100	0,625	0,646	0,486	0,488	0,633	0,651	0,040	0,433	0,010	0,622
200	0,679	0,698	0,582	0,545	0,689	0,705	0,036	0,420	0,015	0,678
300	0,597	0,621	0,498	0,597	0,615	0,634	0,050	0,276	0,003	0,624
400	0,543	0,575	0,437	0,585	0,569	0,591	0,024	0,160	0,006	0,582
500	0,499	0,538	0,387	0,548	0,532	0,558	0,007	0,089	0,004	0,546
600	0,460	0,505	0,346	0,515	0,499	0,529	0,007	0,066	0,005	0,514
700	0,425	0,475	0,311	0,486	0,470	0,504	0,007	0,055	0,005	0,486
800	0,394	0,449	0,281	0,460	0,443	0,481	0,007	0,048	0,006	0,461
900	0,365	0,425	0,255	0,435	0,419	0,462	0,008	0,044	0,006	0,439
1000	0,339	0,402	0,232	0,412	0,397	0,444	0,008	0,040	0,006	0,417

Table 15: Average Power with Different Sample Sizes at Critical Values for 99% Confidence Level

n	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
100	0,625	0,644	0,484	0,478	0,655	0,648	0,038	0,431	0,010	0,610
200	0,678	0,696	0,577	0,536	0,709	0,704	0,035	0,418	0,016	0,665
300	0,575	0,602	0,457	0,577	0,619	0,614	0,049	0,266	0,004	0,560
400	0,482	0,515	0,364	0,522	0,537	0,531	0,025	0,148	0,010	0,478
500	0,414	0,450	0,300	0,457	0,476	0,471	0,007	0,075	0,006	0,412
600	0,357	0,398	0,245	0,404	0,425	0,423	0,007	0,052	0,008	0,355
700	0,307	0,351	0,198	0,356	0,378	0,383	0,008	0,042	0,009	0,304
800	0,260	0,307	0,156	0,311	0,334	0,348	0,008	0,036	0,009	0,262
900	0,221	0,268	0,123	0,271	0,294	0,316	0,008	0,032	0,010	0,228
1000	0,189	0,233	0,099	0,236	0,260	0,288	0,008	0,029	0,011	0,201

Table 16: Average Type I error with Different Sample Sizes at Critical Values for 99% Confidence Level

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0
70	0	0	0	0	0	0	0	0	0	0
80	0	0	0	0	0	0	0	0	0	0
90	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0
110	0	0	0	0	0	0	0	0	0	0
120	0	0	0	0	0	0	0	0	0	0
130	0	0	0	0	0	0	0	0	0	0
140	0	0	0	0	0	0	0	0	0	0
150	0	1	0	1	0	0	0	0	0	0
160	1	1	0	1	1	1	0	0	0	0
170	1	1	0	1	1	1	0	0	0	0
180	1	1	0	1	1	1	0	0	0	0
190	1	1	0	0	1	1	0	0	0	0
200	1	1	0	1	1	1	0	0	0	0

Table 17: 95% Confidence Level Rejection of Heavy Tails for RV

k	KS	KSI	BJ	BJI	SC	SCI	R_1	R_2	R_3	R_4
20	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0
70	0	0	0	0	0	0	0	0	0	0
80	0	0	0	0	0	0	0	0	0	0
90	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0
110	0	0	0	0	0	0	0	0	0	0
120	0	0	0	0	0	0	0	0	0	0
130	0	0	0	0	0	0	0	0	0	0
140	0	0	0	0	0	0	0	0	0	0
150	0	0	0	0	0	0	0	0	0	0
160	0	0	0	0	0	0	0	0	0	0
170	1	1	0	0	0	0	0	0	0	0
180	1	1	0	0	1	0	0	0	0	0
190	0	0	0	0	0	0	0	0	0	0
200	0	0	0	0	0	0	0	0	0	0

Table 18: 99% Confidence Level Rejection of Heavy Tails for RV