South African Statistical Journal
Vol. 54, No. 2, 255–265
https://doi.org/10.37920/sasj.2020.54.2.9

AN EMPIRICAL STUDY OF THE BEHAVIOUR OF THE SAMPLE KURTOSIS IN SAMPLES FROM SYMMETRIC STABLE DISTRIBUTIONS

J. Martin van Zyl
University of the Free State, South Africa
e-mail: *wwjvz@ufs.ac.za*

Sean van der Merwe[1]
University of the Free State, South Africa
e-mail: *vandermerwes@ufs.ac.za*

Kurtosis is seen as a measure of the discrepancy between the observed data and a Gaussian distribution and is defined when the fourth moment is finite. In this work an empirical study is conducted to investigate the behaviour of the sample estimate of kurtosis with respect to sample size and the tail index when applied to heavy-tailed data where the fourth moment does not exist. The study will focus on samples from the symmetric stable distributions. It was found that the expected value of excess kurtosis divided by the sample size is finite for any value of the tail index and the sample estimate of excess kurtosis increases as a linear function of sample size and it is approximately equal to $n(1 - \alpha/2)$.

Key words: Kurtosis, Sample size, Stable distribution, Tail index.

1. Introduction

The theoretical kurtosis is defined and finite when the fourth moment is finite. For regularly varying heavy-tailed distributions (Mikosch, 1999), in terms of the tail index (α), theoretical kurtosis is defined where $\alpha > 4$. In practice, data is observed with an unknown distribution and kurtosis is used to measure how leptokurtic the sample is. In financial data, especially when looking at the distribution of log-returns to estimate volatility, it is often observed that $\alpha < 2$ (Bielinskyi et al., 2019) and the estimated kurtosis is used to get an indication of how leptokurtic the data is. Estimates of kurtosis in asset returns range from 4 to 50 (Engle and Patton, 2001). Heavy-tailed distributions with $\alpha < 4$ are fitted to log-returns, see for example Xu et al. (2011). This work will largely focus on examples relating to financial data. McCulloch (1996) explain that the accumulation of a large number of i.i.d. random asset returns (possibly based on separate pieces of information) has a limiting distribution in the stable distribution family.

In this work an empirical study is conducted to check the behaviour and usefulness of sample kurtosis for the symmetric stable distribution with $\alpha < 2$, and specifically where $\alpha \geq 1$ which is mostly found when applied to real data.

[1] Corresponding author.
MSC2010 subject classifications. 60E07; 62F10.

The main result found was that the expected value of the sample kurtosis increases as a linear function of the sample size and it was found that for symmetric samples from stable distributions, the approximate sample estimate of excess kurtosis for a sample size n and tail index α is $n(1-\alpha/2)$.

More than one method has been suggested to estimate kurtosis, but in this work the Pearson kurtosis as discussed by Fiori and Zenga (2009) is used, as is common in finance and risk analysis. Kurtosis is defined as

$$\beta_2(x) = \frac{E(X-\mu)^4}{[E(X-\mu)^2]^2} = \frac{\mu_4}{\sigma^4},$$

with μ_4 the fourth central moment and σ^2 the variance of X. $\beta_2(x)$ is location-scale invariant and all data simulated will be for a location parameter $\mu = 0$ and scale parameter $\sigma = 1$. For a regular distribution, $\mu_4 \geq \mu_2^2$ (Ord and Stuart, 1994, p. 109). The excess kurtosis is

$$\gamma_2 = \beta_2 - 3,$$

which is also equal to $\gamma_2 = \kappa_4/\kappa_2^2$ when expressed in terms of cumulants. For the normal distribution, the excess kurtosis is zero.

The sample kurtosis is denoted by b_2 and the excess kurtosis by $g_2 = b_2 - 3$. Algebraic inequalities which do not depend on distributional properties exist for the sample kurtosis, including that the sample estimate of kurtosis is less than the sample size (Johnson and Lowe, 1979; Cox, 2010), thus

$$b_2 = \frac{\frac{1}{n}\sum_{j=1}^n \left(x_j - \bar{x}\right)^4}{\left[\frac{1}{n}\sum_{j=1}^n \left(x_j - \bar{x}\right)^2\right]^2} = \frac{n\sum_{j=1}^n \left(x_j - \bar{x}\right)^4}{\left[\sum_{j=1}^n \left(x_j - \bar{x}\right)^2\right]^2} = nc(x_1, \ldots, x_n) \leq n.$$

This inequality shows that the function $c(x_1, \ldots, x_n) \leq 1$ and $E[c(x_1, \ldots, x_n)] = E[b_2/n] \leq 1$ would be finite for all distributions, which means that divergence of the sample kurtosis arises because of an increase in the sample size. The behaviour of the sample kurtosis will be like that of a ratio, and hence we should not consider the numerator and denominator separately as is done in the theoretical definition. Using simulation studies it was checked if $c(x_1, \ldots, x_n)$ can be approximated as a function of α. It can also be seen and was confirmed using simulation that the variance of the sample kurtosis is of the form $n^2 var[c(x_1, \ldots, x_n)]$.

This work will focus on symmetrical stable distributed data. Properties and applications of it can for example be found in the work of Čížek et al. (2011) and Chernobai et al. (2007). We use their notation below.

The characteristic function of the family of stable distributions is denoted by $\phi(t)$ where

$$\log(\phi(t)) = \begin{cases} -\sigma^\alpha |t|^\alpha \left[1 - i\beta \text{sign}(t)\tan(\pi\alpha/2)\right] + i\mu t, & \alpha \neq 1, \\ -\sigma|t| \left[1 + i\beta \text{sign}(t)(2/\pi)\log(|t|)\right] + i\mu t, & \alpha = 1. \end{cases}$$

The parameters are the tail index, $\alpha \in (0, 2]$, a scale parameter $\sigma > 0$, coefficient of skewness $\beta \in [-1, 1]$ and location parameter μ.

In Figure 1, 500 random samples were simulated. α's were randomly chosen on the interval $[1, 2]$ with random sample sizes between $n = 200$ and $n = 1500$. The plots show the estimated excess kurtosis. The focus of this study is applications in finance and these sample sizes cover 1 to 6 years when working with daily data.

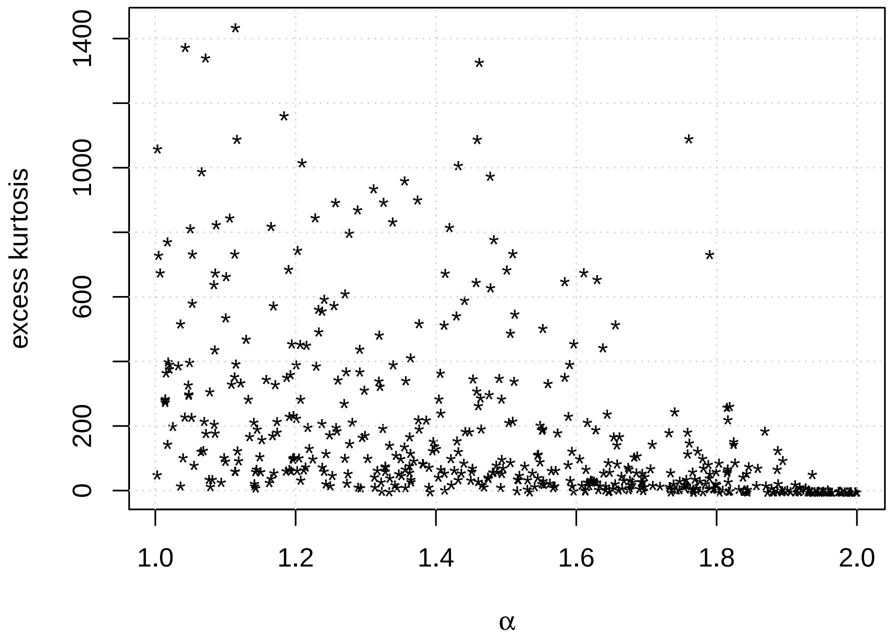

Figure 1. A scatterplot of 500 sample estimates of the excess kurtosis for random $\alpha \in [1, 2]$ and the sample size n between 200 and 1 500. Samples are from a stable distribution.

To get an idea of the relationship involved, regression showed that

$$g_2 \approx n(1 - \alpha/2).$$

There is little variation in the regression coefficients when repeating the simulation and this relationship will be investigated further using simulation studies. Assuming that $g_2(n, \alpha) = n(1 - \alpha/2)$, and by noting that

$$\frac{\partial g_2(n, \alpha)}{\partial n} = 1 - \alpha/2 \quad \text{and} \quad \frac{\partial g_2(n, \alpha)}{\partial \alpha} = -n/2,$$

it can be seen that the sample kurtosis is an increasing function of the sample size, and sensitive to changes in α.

The behaviour of sample skewness was checked using the simulated samples and it was found that the expected value of the sample skewness is zero for symmetric data but the variance is an increasing linear function of the sample size and increases for smaller α. This is not the focus of the work, but a skewness estimate in a large sample might not be a significant indication of skewness if the large variance is taken into account.

A measure to order different symmetric distributions according to the term 'heavy-tailness' was derived by van Zwet (1964) and Groeneveld and Meeden (1984). It was proven that if a distribution is more heavy-tailed than another according to this measure, then kurtosis will also be larger for the heavier-tailed distribution.

2. Simulation study

Say a sample of size n is available and $n = k_1 + \cdots + k_r$. The sample kurtosis will be calculated at increasing sample sizes, say $k_1, k_1 + k_2, k_1 + k_2 + k_3, \ldots, n$, and for different values of α.

In Figure 2 we see the increase in the expected value of the estimated excess kurtosis. The average at each sample size was calculated using $m = 25\,000$ samples. The slope for $\alpha = 1$ is $b = 0.5013$, for $\alpha = 1.5$ it is $b = 0.2474$, and approximately zero for $\alpha = 2$. The proposed relationship (Figure 1) can be considered as an approximation.

The excess kurtosis for simulated data from a t-distribution with $\nu = 3, 4, 5$ degrees of freedom is shown in Figure 3. It can be seen that for small degrees of freedom the relationship between kurtosis and sample size is not linear. The linear trend increase in kurtosis with respect to sample size for samples from a stable distribution can thus be a useful property for model selection. It may not be unique to the stable distribution, but if observed in a practical problem it suggests that a possible candidate to fit might be a stable distribution.

For a given value of α, consider regression between g_2/n and α. If it is assumed that the excess kurtosis is zero for $\alpha = 2$, regression through the origin can be performed. For $n = 250$ and for each $\alpha = 1, 1.1, \ldots, 2$, perform 5 000 regressions. For a specific α calculate the average of the slopes. In Figure 4 the slopes are plotted against $2 - \alpha$, and the slope of this line with respect to $2 - \alpha$ is approximately 0.5, leading to $g_2 \approx 0.5n(2 - \alpha)$. The figure also includes 90% intervals for the slope of a single sample.

The expected value of the sample excess kurtosis increases as a linear function of the number of observations used to calculate the sample kurtosis and the approximate expected excess kurtosis in large samples is thus

$$E(g_2) \approx n(1 - \alpha/2).$$

In the above simulations a fixed sample size was used. To confirm the results it will be checked by using random sample sizes. The consistency of the ordering of kurtosis with tail index will also be considered. As an example 5 000 samples with random sample sizes between $n = 200$ and $n = 1\,500$ were generated from a stable distribution with index $\alpha = 1.25$, estimated excess kurtosis divided by the sample size and the sample mean was 0.1236 compared to $1 - \alpha/2 = 0.1250$. Similarly 5 000 samples with random sample sizes and $\alpha = 1.75$ were generated. The sample mean for $\alpha = 1.75$ is 0.3668 compared to $1 - \alpha/2 = 0.3750$. This confirms the approximate relationship $E(g_2) \approx n(1 - \alpha/2)$.

Comparisons were made in pairs between the sample kurtosis of the two samples. This resulted in approximately 82% correct larger values of the excess kurtosis when the data is more heavy-tailed. If sample kurtosis was divided by the sample size the percentage increases by about 2%. A few such examples were simulated and when comparing between a normal samples $\alpha = 2$ and samples where $\alpha < 2$, the percentage correct ordering of the tail index with respect to kurtosis is very high and often as high as 100%. The conclusion can be made that for symmetric stable distributions, kurtosis is an effective measure to compare the tail-heaviness of samples from two distributions with different parameters.

Table 1 shows the ability of the sample kurtosis (standardised by sample size) to correctly choose the heavier tail in the case of stable distributions, for various combinations of tail index.

The standard deviation of the sample estimate as a function of α is seen in Figure 5 and the variance

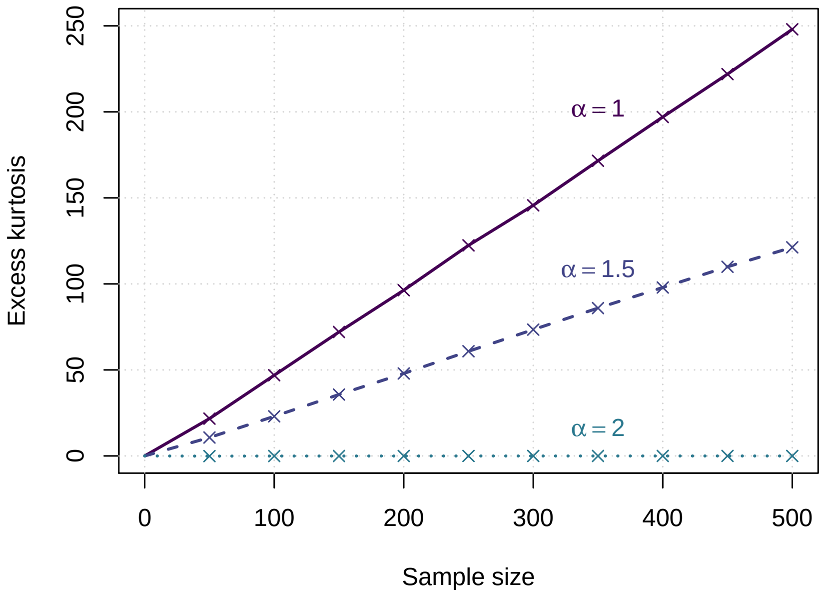

Figure 2. Plot of average of excess sample kurtosis using simulated samples from a symmetric stable distribution. The averages were calculated using 25 000 samples for each sample size ranging from 50 to 500 in steps of 50.

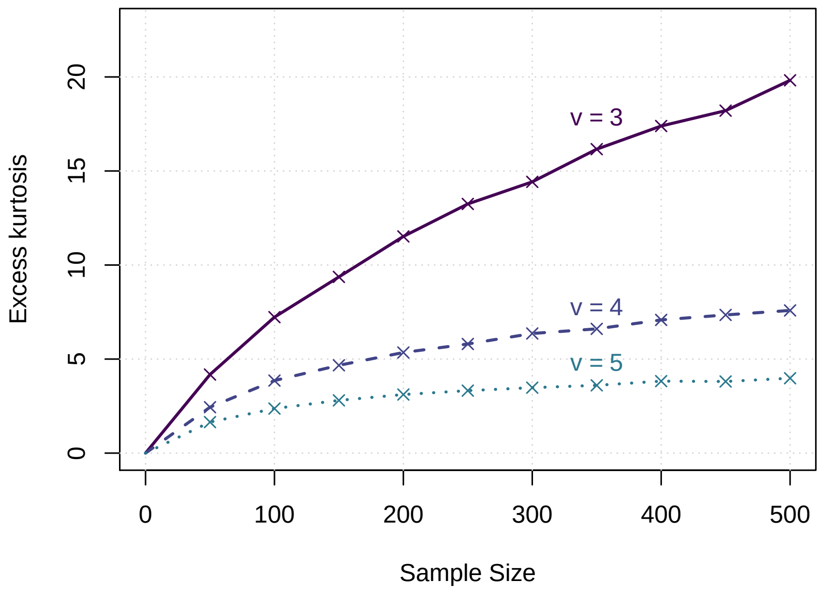

Figure 3. Plot of average of excess sample kurtosis using simulated samples from a *t*-distribution. The averages were calculated using 25 000 samples for each sample size ranging from 50 to 500 in steps of 50.

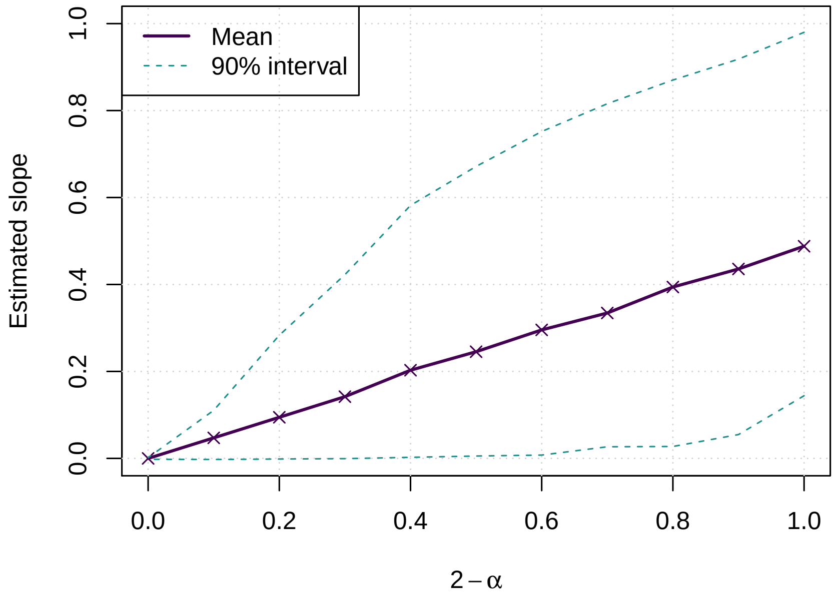

Figure 4. The relationship between the estimated slope of increase of excess kurtosis with respect to sample size and $2 - \alpha$. Each point calculated as the average of 5 000 estimated values.

for at sample size n is proportional to n^2 for all values of α, except when $\alpha = 2$.

It is worth noting that the sample kurtosis of a sample from a stable distribution does not appear to be affected by the skewness parameter value. Repeating the primary simulation study with varying skewness values produced regression results simultaneously showing a very strong relationship between sample kurtosis and α on the one hand and an insignificant relationship with skewness (as well as the interaction of α and skewness) on the other hand.

3. An application to log-returns

The change in kurtosis with respect to the number of observations used to calculate the sample kurtosis was investigated when applied to log-returns of the New York Stock Exchange (NYSE). The daily closing values of 5 years from May 2013 to May 2018 were used. These log-returns are approximately symmetrically distributed with sample mean zero and the stable distribution is considered as a possible distribution. Log-returns are assumed independent for the purpose of this analysis.

The index and log returns are shown in Figure 6. There is an initial period, a major correction and the period after the correction.

A histogram of the log returns are shown in Figure 7 to illustrate the values that are modelled from here on.

The sample excess kurtosis of the log-returns using increasing sample sizes is plotted in Figure 8. It can be seen that the distribution of the log-returns seems to change and then stay the same for a period if one considers a change in slope as an indication of a change in the distribution.

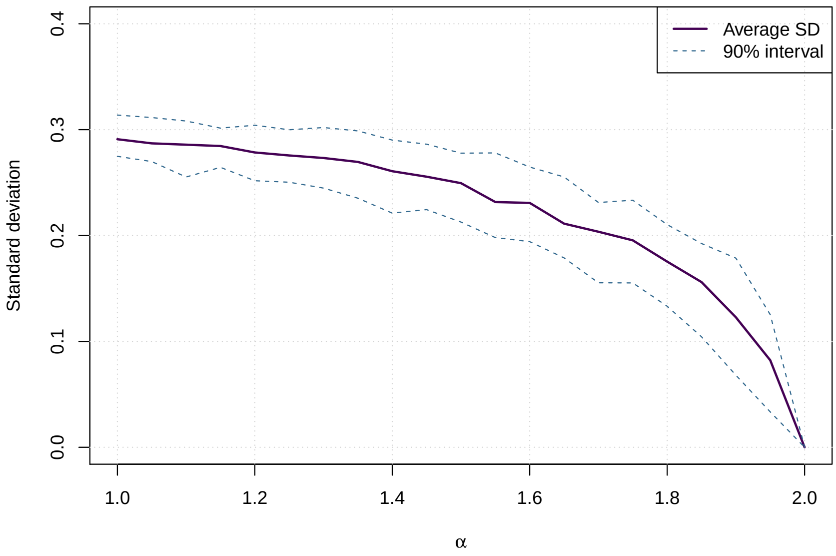

Figure 5. The standard deviation of the sample size standardised excess kurtosis estimate (g_2), as a function of α, based on 50 sets of 100 simulated samples of size 500 for each α.

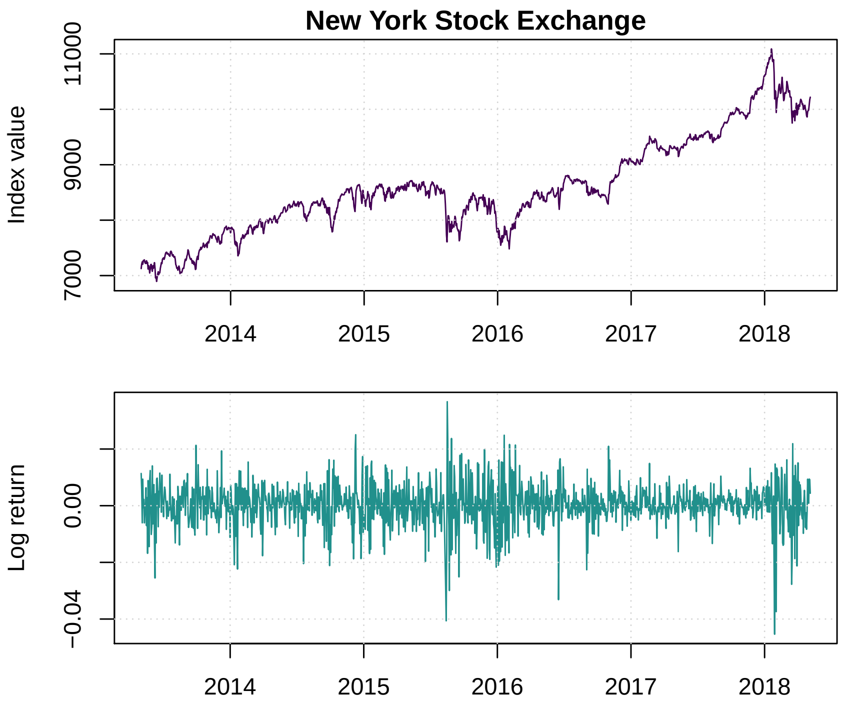

Figure 6. Index and log returns of the NY stock exchange, 5 years daily data.

Table 1. Approximate probability of the sample kurtosis of a sample from the stable distribution with row label tail index exceeding the kurtosis from a sample with column label tail index.

α	1.25	1.5	1.75	2
1	0.635	0.772	0.899	1.000
1.25		0.661	0.839	0.999
1.5			0.730	0.997
1.75				0.991

By using the relationship, $E(g_2) \approx n(1 - \alpha/2)$, one finds that $\hat{\alpha} \approx 2(1 - g_2/n)$. This can be used as an approximate method to calculate an estimate of the tail index. The closer to 2, the less the volatility. It can be seen in Figure 9 that there are periods with a changing index and constant periods in terms of the estimated tail index.

The Kogon-Williams estimation method (Kogon and Williams, 1998) was applied and it suggested a tail index of roughly 1.76 for the data as a whole. When we consider two series of observations, 100-–400 and 800-–1100, we obtain estimates of roughly 1.86 and 1.72, suggesting a systemic change. Thus, we expand on this using a rolling window approach. See Figure 8, where the systemic change is clearly visible.

Kurtosis is very sensitive with respect to changes in the index even though the kurtosis is calculated using log-returns and the difference in behaviour of the sample kurtosis over time before and after the correction is very clear.

The Kogon-Williams estimation method (Kogon and Williams, 1998) was also applied to a rolling window of size 300, and plotted in Figure 9 along with the tail index implied by the estimated kurtosis, to see if there was a change in the tail-index as indicated by the change in sample kurtosis.

The results are consistent with the change in kurtosis, showing that the second period can be more volatile and heavy-tailed. These are periods before and after the 2016 presidential election in the USA.

4. Conclusion

There is a relationship between kurtosis and the tail-index for samples from the stable distributions. For a sample of size n, the sample kurtosis can be considered as n times the ratio of two polynomials which are both of degree 4, and the expected value of the ratio is finite, even if the expected value of the numerator or denominator does not exist. This property makes kurtosis useful in heavy-tailed data if the proportionality to n is taken into account. Thus for $\alpha > 0$,

$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n} \left(x_j - \bar{x}\right)^4}{\left[\sum_{j=1}^{n} \left(x_j - \bar{x}\right)^2\right]^2} \approx 1 - \alpha/2.$$

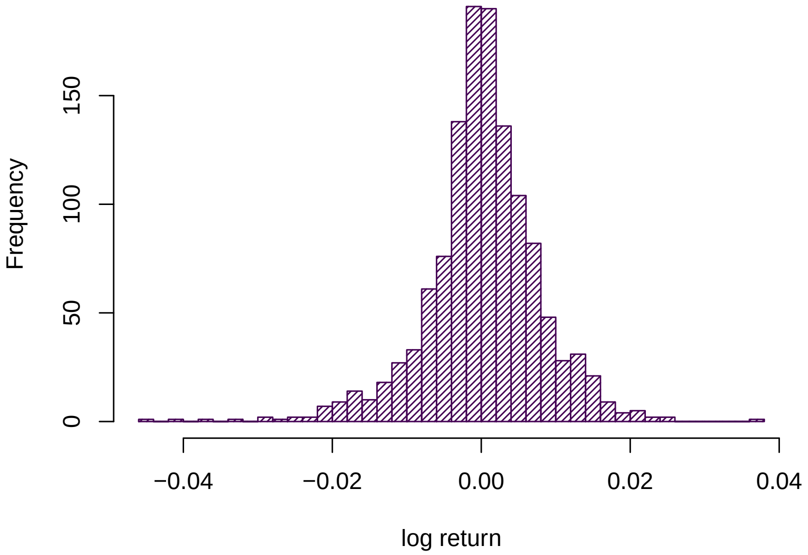

Figure 7. Histogram of log returns of the NY stock exchange, 5 years daily data.

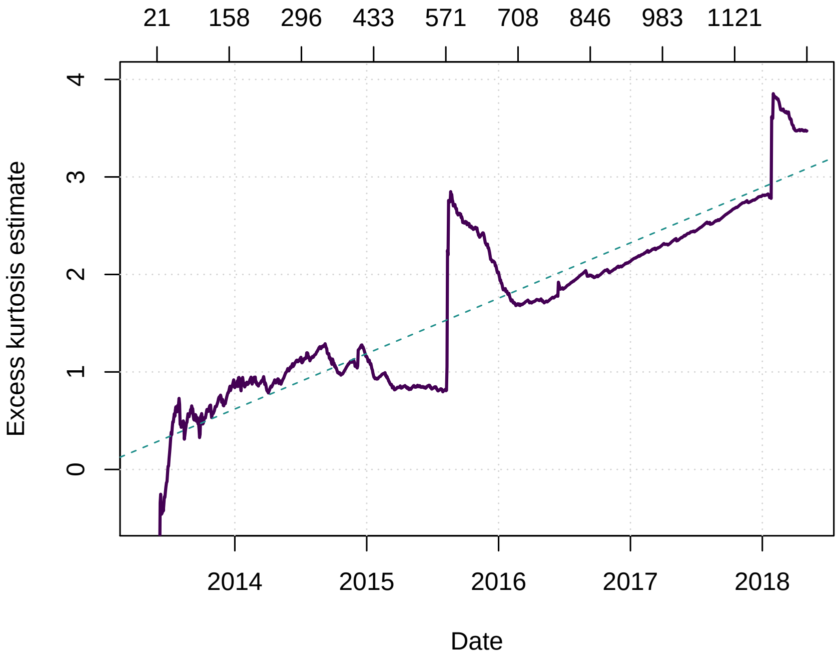

Figure 8. Excess kurtosis of the log-returns calculated as a function of the number of points used.

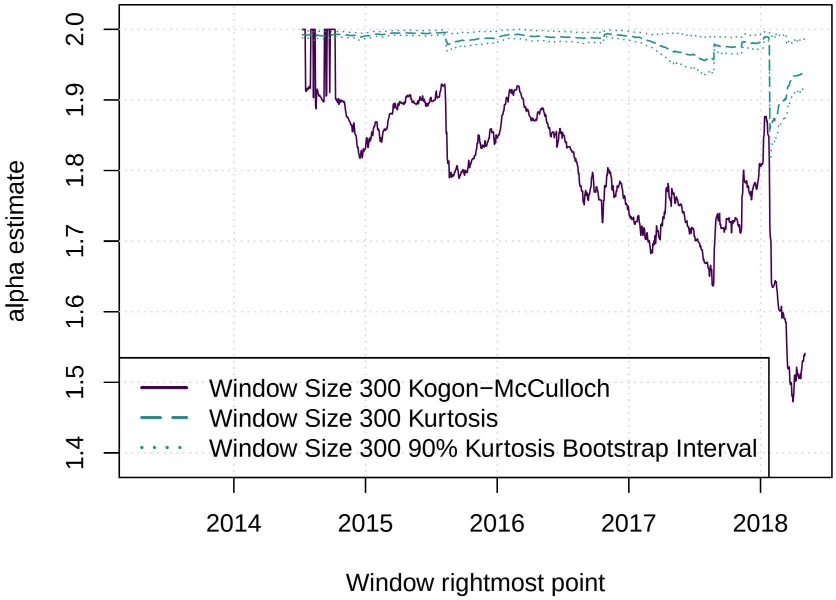

Figure 9. Estimated tail index of the log-returns.

This property can be used to compare the 'tail-heaviness' by using kurtosis of two samples from a stable distribution.

The linear relationship between the increase as more points are used to calculate kurtosis can be used as a property to exclude or include a stable distribution as a possible distribution which can be fitted to for example log-returns.

For GARCH models the fourth moment should be finite when fitted to log-returns. By plotting the estimated kurtosis as a function of an increasing number of observations an increase in sample kurtosis might be an indication that the fourth moment is not finite.

Using bootstrap methods to estimate a variance, the relationship $g_2/n \approx 1 - \alpha/2$ can be used in large samples to test hypotheses concerning α and especially to test if $\alpha < 2$.

Notes. The simulation study was done in the 'R' statistical programming language (R Core Team, 2019). Additional packages loaded include the 'StableEstim' package (Kharrat and Boshnakov, 2016) and the 'stabledist' package (Wuertz et al., 2016).

The annotated code and data are available from the corresponding author.

Acknowledgements. The authors would like to thank all those involved in the review and editorial process for their contributions to improving the paper.

References

Bielinskyi, A., Soloviev, V. N., Semerikov, S. O., and Solovieva, V. (2019). Detecting stock crashes using Levy distribution. *In Proceedings of the Selected Papers of the 8th International Conference on Monitoring, Modeling & Management of Emergent Economy*. Odessa, Ukraine, 420–433.

Chernobai, A. S., Rachev, S. T., and Fabozzi, F. J. (2007). *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis*. Wiley & Sons, Hoboken, New Jersey.

Cox, N. J. (2010). Speaking stata: The limits of sample skewness and kurtosis. *The Stata Journal*, **10**, 482–495.

Čížek, P., Weron, R., and Härdle, W. K. (2011). *Statistical Tools for Finance and Insurance*. Hugo Steinhaus Center, Wroclaw University of Technology, Berlin.

Engle, R. F. and Patton, A. J. (2001). What good is a volatility model? *Quantitative Finance*, **1**, 237–245.

Fiori, A. M. and Zenga, M. (2009). Karl Pearson and the origin of kurtosis. *International Statistical Review*, **77**, 40–50.

Groeneveld, R. A. and Meeden, G. (1984). Measuring skewness and kurtosis. *Journal of the Royal Statistical Society, Series D*, **33**, 391–399.

Johnson, M. E. and Lowe, V. W. (1979). Bounds on the sample skewness and kurtosis. *Technometrics*, **21**, 377–378.

Kharrat, T. and Boshnakov, G. N. (2016). *StableEstim: Estimate the Four Parameters of Stable Laws using Different Methods*. R package version 2.1.
URL: *https://CRAN.R-project.org/package=StableEstim*

Kogon, S. M. and Williams, D. B. (1998). Characteristic function based estimation of stable distribution parameters. *In A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhäuser, Boston, 311–338.

McCulloch, J. H. (1996). 13 financial applications of stable distributions. *In Statistical Methods in Finance*, volume 14 of *Handbook of Statistics*. Elsevier, 393–425.

Mikosch, T. (1999). Regular variation, subexponentiality and their applications in probability theory. Lecture notes for the workshop: Heavy Tails and Queues, EURANDOM Institute, Eindhoven, Netherlands.

Ord, K. and Stuart, A. (1994). *Kendall's Advanced Theory of Statistics: Distribution Theory*. Wiley, Chichester.

R Core Team (2019). R: A language and environment for statistical computing.
URL: *http://www.R-project.org/*

van Zwet, W. (1964). Convex transformations of random variables. *In Mathematical Centre Tracts*, volume 7. Mathematisch Centrum, Amsterdam.

Wuertz, D., Maechler, M., and core team members., R. (2016). *stabledist: Stable Distribution Functions*. R package version 0.7-1.
URL: *https://CRAN.R-project.org/package=stabledist*

Xu, W., Wu, C., Dong, Y., and Xiao, W. (2011). Modeling chinese stock returns with stable distribution. *Mathematical and Computer Modelling*, **54**, 610–617.

Manuscript received 2020-03-11, revised 2020-05-27, accepted 2020-05-30.