Title

On More Robust Estimation of Skewness and Kurtosis: Simulation and Application to the S&P500 Index

Authors
Kim, Tae-Hwan
White, Halbert

Publication Date
2003-09-01

On More Robust Estimation of Skewness and Kurtosis: Simulation and Application to the S&P500 Index

Tae-Hwan Kim
School of Economics, University of Nottingham
University Park, Nottingham NG7 2RD, UK
(Phone) 44-115-951-5466 (Fax) 44-115-951-4159
Tae-Hwan.Kim@nottingham.ac.uk

Halbert White
Department of Economics, University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093-0508
(Phone) 858-534-3502 (Fax) 858-534-7040
hwhite@weber.ucsd.edu

September, 2003

Abstract: For both the academic and the financial communities it is a familiar stylized fact that stock market returns have negative skewness and excess kurtosis. This stylized fact has been supported by a vast collection of empirical studies. Given that the conventional measures of skewness and kurtosis are computed as an average and that averages are not robust, we ask, "How useful are the measures of skewness and kurtosis used in previous empirical studies?" To answer this question we provide a survey of robust measures of skewness and kurtosis from the statistics literature and carry out extensive Monte Carlo simulations that compare the conventional measures with the robust measures of our survey. An application of the robust measures to daily S&P500 index data indicates that the stylized facts might have been accepted too readily. We suggest that looking beyond the standard skewness and kurtosis measures can provide deeper insight into market returns behaviour.

Keywords: Skewness, Kurtosis, Quantiles, Robustness.

1. Introduction

It has long been recognized that the behavior of stock market returns does not agree with the frequently assumed normal distribution. This disagreement is often highlighted by showing the large departures of the skewness and kurtosis of returns from normal distribution counterparts. For both the academic and the financial communities it has become a firm and indisputable stylized fact that stock market returns have negative skewness and excess kurtosis. This stylized fact has been supported by a huge collection of empirical studies. Some recent papers on this issue include Bates (1996), Jorion (1988), Hwang and Satchell (1999), and Harvey and Siddique (1999, 2000).

The role of higher moments has become increasingly important in the literature mainly because the traditional measure of risk, variance (or standard deviation), has failed to capture fully the "true risk" of the distribution of stock market returns. For example, if investors prefer right-skewed portfolios, then more reward should be given to investors willing to invest in left-skewed portfolios even though both portfolios have the same standard deviation. This suggests that the "true risk" may be a multi-dimensional concept and that other measures of distributional shape such as higher moments can be useful in obtaining a better description of multi-dimensional risk. In this context, Harvey and Siddique (2000) proposed an asset pricing model that incorporates skewness and Hwang and Satchell (1999) developed a CAPM for emerging markets taking into account skewness and kurtosis.

Given this emerging interest in skewness and kurtosis in financial markets, one should ask the following question: how useful are the measures of skewness and kurtosis used in previous empirical studies? Practically all of the previous work concerning skewness and kurtosis in financial markets has used the conventional measures of skewness and kurtosis (that is, the standardized third and fourth sample moments or some variants of these). It is well known that the sample mean (also its regression version, the least squares estimator) is very sensitive to outliers. Since the conventional measures of skewness and kurtosis are essentially based on sample averages, they are also sensitive to outliers. Moreover, the impact of outliers is greatly amplified in the conventional measures of skewness and kurtosis due to the fact that they are raised to the third and fourth powers.

In the statistics literature, a great deal of effort has been taken to overcome the non-robustness of the conventional measures of location and dispersion (i.e. mean and variance), and some attention has been paid to the non-robustness of the conventional measures of skewness and kurtosis. In the finance literature, there has been some concern with non-robustness of conventional measures of location and dispersion, but almost no attention to the non-robustness of conventional measures of skewness and kurtosis. In this paper, we consider certain robust alternative measures of skewness and kurtosis based on quantiles that have been previously developed in the statistics literature, and we conduct extensive Monte Carlo simulations to evaluate and compare the conventional measures of skewness and kurtosis and their robust counterparts. Our simulation results demonstrate that the conventional measures are extremely sensitive to single outliers or small groups of outliers, comparable to those observed in U.S. stock returns. An application of robust measures to the daily S&P500 index indicates that the familiar stylized facts (negative skewness and excess kurtosis in financial markets) may have been too readily accepted.

2. Review of Robust Measures of Skewness and Kurtosis

We consider a process $\{y_t\}_{t=1,2,...,N}$ and assume that the y_t's are independent and identically distributed with a cumulative distribution function F. The conventional coefficients of skewness and kurtosis for y_t are given by:

$$SK_1 = E\left(\frac{y_t - m}{s}\right)^3, \qquad KR_1 = E\left(\frac{y_t - m}{s}\right)^4 - 3,$$

where $m = E(y_t)$ and $s^2 = E(y_t - m)^2$, and expectation E is taken with respect to F. Given the data $\{y_t\}_{t=1,2,...,N}$, SK_1 and KR_1 are usually estimated by the sample averages

$$\hat{SK}_1 = T^{-1}\sum_{t=1}^{N}\left(\frac{y_t - \hat{m}}{\hat{s}}\right)^3, \qquad \hat{KR}_1 = T^{-1}\sum_{t=1}^{N}\left(\frac{y_t - \hat{m}}{\hat{s}}\right)^4 - 3,$$

where $\hat{m} = T^{-1}\sum_{t=1}^{N} y_t$, $\hat{s}^2 = T^{-1}\sum_{t=1}^{N}(y_t - \hat{m})^2$.

Due to the third and fourth power terms in SK_1 and KR_1, the values of these measures can be arbitrarily large, especially when there are one or more large outliers in the data. For this reason, it can sometimes be difficult to give a sensible interpretation to large values of these measures simply because we do not know whether the true values are indeed large or there exist some outliers. One seemingly simple solution is to eliminate the outliers from the data. Two problems arise in this approach. One is that the decision to eliminate outliers is taken usually after visually inspecting the data; this can invalidate subsequent statistical inference. The other is that deciding which observations are outliers can be somewhat arbitrary.

Hence, eliminating outliers manually is not as simple as it may appear, and it is desirable to have non-subjective robust measures of skewness and kurtosis that are not too sensitive to outliers. We turn now to a description of a number of more robust measures of skewness and kurtosis that have been proposed in the statistics literature. It is interesting to note that among the robust measures we discuss below, only one requires the second moment and all other measures do not require any moments to exist.

2.1 Robust Measures of Skewness

Robust measures of location and dispersion are well known in the literature. For example, the median can be used for location and the interquartile range for dispersion. Both the median and the interquartile range are based on quantiles. Following this tradition, Bowley (1920) proposed a coefficient of skewness based on quantiles:

$$SK_2 = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1},$$

where Q_i is the i^{th} quartile of y_t: that is, $Q_1 = F^{-1}(0.25)$, $Q_2 = F^{-1}(0.5)$ and $Q_3 = F^{-1}(0.75)$. It is easily seen that for any symmetric distribution, the Bowley coefficient of skewness is zero. The denominator $Q_3 - Q_1$ re-scales the coefficient so that the maximum value for SK_2 is 1, representing extreme right skewness and the minimum value for SK_2 is –1, representing extreme left skewness.

The Bowley coefficient of skewness has been generalized by Hinkley (1975):

$$SK_3(\alpha) = \frac{F^{-1}(1-\alpha) + F^{-1}(\alpha) - 2Q_2}{F^{-1}(1-\alpha) - F^{-1}(\alpha)},$$

for any α between 0 and 0.5. Note that the Bowley coefficient of skewness is a special case of Hinkley's coefficient when α = 0.25. This measure, however, depends on the value of α and it is not clear what value should be used for α. One way of removing this dependence is to integrate out α, as done in Groeneveld and Meeden (1984):

$$SK_3 = \frac{\int_0^{0.5}\left\{F^{-1}(1-a)+F^{-1}(a)-2Q_2\right\}da}{\int_0^{0.5}\left\{F^{-1}(1-a)-F^{-1}(a)\right\}da}$$

$$= \frac{m-Q_2}{E\,|\,y_t-Q_2\,|}.$$

This measure is also zero for any symmetric distributions and is bounded by –1 and 1.

Noting that the denominator $E\,|\,y_t-Q_2\,|$ is a kind of dispersion measure, we observe that the Pearson coefficient of skewness [Kendall and Stuart (1977)] is obtained by replacing the denominator with the standard deviation as follows:

$$SK_4 = \frac{m-Q_2}{s}.$$

Groeneveld and Meeden (1984) have put forward the following four properties that any reasonable coefficient of skewness $g(y_t)$ should satisfy: (i) for any $a \in (0,\infty)$ and $b \in (-\infty,\infty)$, $g(y_t) = g(ay_t+b)$; (ii) if y_t is symmetrically distributed, then $g(y_t) = 0$; (iii) $-g(y_t) = g(-y_t)$; (iv) if F and G are cumulative distribution functions of y_t and x_t, and $F <_c G$, then $g(y_t) \leq g(x_t)$ where $<_c$ is a skewness-ordering among distributions (see Zwet (1964) for the definition). Groeneveld and Meeden proved that SK_1, SK_2 and SK_3 satisfy (i)-(iv), but SK_4 satisfies (i)-(iii) only.

2.2 Robust Measures of Kurtosis

Moors (1988) showed that the conventional measure of kurtosis (KR_1) can be interpreted as a measure of the dispersion of a distribution around the two values $m \pm s$. Hence, KR_1 can be large when probability mass is concentrated either near the mean m or in the tails of the distributions. Based on this interpretation, Moors (1988) proposed a robust alternative to KR_1: $\frac{(E_7-E_5)+(E_3-E_1)}{E_6-E_2}$, where E_i is the i^{th} octile: that is, $E_i = F^{-1}(i/8)$ for $i = 1, 2, \ldots, 7$. Moors justified this estimator on the ground that the two terms, (E_7-E_5) and (E_3-E_1), are large (small) if relatively little (much) probability mass is concentrated in the neighbourhood of E_6 and E_2, corresponding to large (small) dispersion around $m \pm s$. The denominator is a scaling factor ensuring that the statistic is invariant under linear transformation. As we do for KR_1, we center the Moors coefficient of kurtosis at the value for the standard normal distribution. It is easy to calculate that $E_1 = -E_7 = -1.15$, $E_2 = -E_6 = -0.68$, $E_3 = -E_5 = -0.32$ and $E_4 = 0$ for $N(0,1)$ and therefore the Moors coefficient of kurtosis is 1.23. Hence, the centered coefficient is given by:

$$KR_2 = \frac{(E_7 - E_5) + (E_3 - E_1)}{E_6 - E_2} - 1.23.$$

While investigating how to test light-tailed distributions against heavy-tailed distributions, Hogg (1972, 1974) found that the following measure of kurtosis performs better than the traditional measure KR_1 in detecting heavy-tailed distributions: $\frac{U_\alpha - L_\alpha}{U_\beta - L_\beta}$ where U_α (L_α) is the average of the upper (lower) α quantiles defined as: $U_\alpha = \frac{1}{\alpha}\int_{1-\alpha}^{1} F^{-1}(y)dy$, $L_\alpha = \frac{1}{\alpha}\int_{0}^{\alpha} F^{-1}(y)dy$, for $\alpha \in (0,1)$. According to Hogg's simulation experiments, α = 0.05 and β = 0.5 gave the most satisfactory results. Here, we adopt these values for α and β. For $N(0,1)$, we have $U_{0.05} = -L_{0.05} = 2.06$ and $U_{0.5} = -L_{0.5} = 0.80$, implying that the Hogg coefficient with α = 0.05 and β = 0.5 is 2.59. Hence, the centered Hogg coefficient is given by:

$$KR_3 = \frac{U_{0.05} - L_{0.05}}{U_{0.5} - L_{0.5}} - 2.59.$$

Another interesting measure based on quantiles has been used in Crow and Siddiqui (1967), which is given by $\frac{F^{-1}(1-\alpha) + F^{-1}(\alpha)}{F^{-1}(1-\beta) - F^{-1}(\beta)}$ for $\alpha, \beta \in (0,1)$. Their choices for α and β are 0.025 and 0.25 respectively. For these values, we obtain $F^{-1}(0.975) = -F^{-1}(0.025) = 1.96$ and $F^{-1}(0.75) = -F^{-1}(0.25) = -0.68$ for N(0,1) and the coefficient is 2.91. Hence, the centered coefficient is:

$$KR_4 = \frac{F^{-1}(0.975) + F^{-1}(0.025)}{F^{-1}(0.75) - F^{-1}(0.25)} - 2.91.$$

3. Monte Carlo Simulations

In this section we conduct Monte Carlo simulations designed to investigate how robust the alternative measures of skewness and kurtosis are in finite samples. The simulations were carried out on a 700MHz PC using MATLAB. The random number generator used in the simulations is that from the MATLAB Statistics Toolbox.

We choose three symmetric distributions and one non-symmetric distribution for our simulation study. Our symmetric distributions are the standard normal distribution N(0,1), and the Student $t-$distributions with 10 and 5 degrees of freedom [T-10, T-5]. These represent moderate, heavy and very heavy tailed distributions. For the non-symmetric distribution we use the log-normal distribution with $m = 1$, $s = 0.4$ (shifted by $-e^{1+0.5(0.4)^2}$ so that the mean is zero) denoted Log-N(1,0.4). For a range of values for N (50, 250, 500, 1000, 2500 and 5000), we generate $\{y_t\}_{t=1,2,...,N}$ using the four distributions and calculate the various measures of skewness and kurtosis discussed in the previous section. The number of replications for each experiment is 1000. The true values of the skewness and kurtosis measures for the various distributions are provided in Table 1. If our statistics are consistent, then their Monte Carlo distributions should collapse around these values as $N \to \infty$.

The simulation results are reported in Figures 1-8. Each figure is divided into two sections: A and B. Figure A shows smoothed histograms of estimated coefficients of skewness or kurtosis for the 5 different sample sizes and 4 data generating distributions. We use a kernel density method to smooth the histograms. In order to provide additional information, Figure B displays box-plots for the corresponding smoothed histograms. As usual, each box represents the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the box and their length is chosen to be the same as the length of the corresponding box (i.e. the inter-quartile range). The number at the end of each whisker is the number of observations beyond the end of the whiskers.

As expected, the performance of SK_1 (Figure 1) deteriorates as the distribution moves from N(0,1) to T-10 and T-5. The sampling distributions tend to have large dispersion and the number of observations outside the whiskers is noticeably increasing. For the lognormal case, SK_1 is not a good measure when N is small, which is indicated by the fact that the center of the box-plot is quite different from the limiting value (1.32), although it does move towards the limiting value as

$N \to \infty$. These problems, however, are not present in the other robust measures, SK_2 (Figure 2), SK_3 (Figure 3), SK_4 (Figure 4). That is, the sampling distributions are fairly stable and similar across various distributions and also, for the lognormal case, the center of the box-plot stays near the limiting value even for $N = 50$.

The performance of KR_1 (Figure 5) is even worse than SK_1 as the distribution moves from N(0,1) to T-10 and T-5. For the T-5 case, the center of the box-plot is still far away from the true value 6 even for $N = 5000$. The small sample bias of KR_1 for Log-N(1,0.4) is also evident in the box-plot. Other robust estimates (KR_2 and KR_4) do not exhibit these problems at all. The sampling distribution of KR_3 indicates that there is a small finite sample bias when the number of observations is less than 1000.

Next we add a single outlier to the same set of generated random numbers $\{y_t\}_{t=1,2,...,N}$ and calculate the same set of measures in order to see the impact of the single outlier on the various measures. The outlier is constructed to occur at time $t \in (0,1)$. In order to inject some realism into the simulation study, we use the daily S&P500 index to calculate the size and location of the outlier. The sample period is January 1, 1982 through June 29, 2001 with 5085. The largest outlier (-20.41%) is caused by the 1987 stock market crash. The timing of the crash in this sample period is $t = 0.3$ (the location of the outlier in the sample divided by the total number of observations). We calculate the 25th percentile of the sampling distribution of the S&P index. It turns out that the 25th percentile is -0.42. The size of the outlier relative to the 25th percentile is then given by $m = y_{[tT]} / F^{-1}(0.25) = -20.41/-0.42 = 48.62$. Therefore, we first generate random numbers $\{y_t\}_{t=1,2,...,N}$ and calculate the 25th percentile $F^{-1}(0.25)$. The outlier is then $mF^{-1}(0.25)$ and we replace the observation $y_{[0.3T]}$ with the outlier.

The results are reported in Figures 9-16. The impact of one outlier on SK_1 is clearly visible in Figure 9. Note that the center of the sampling distributions of SK_1 is moving toward to zero (the true value for all symmetric distributions) once N is greater than 500, but even for $N = 5000$, the center is far from zero. No whiskers from the box-plots contain the true value 0 in their upper tails. The maximum impact occurs approximately when $N = 500$. In that case the center of the box-plots is around between -13 and -10. In contrast, the outlier has no impact on SK_2 at all. This is expected since SK_2 is based on quantiles whose values are not changed by a single observation.

On the other hand, we see the moving-box type of convergence for the other robust skewness estimators (SK_3 and SK_4). This is because these measures involve m and s, and the outlier must have some impact on the sample mean and sample standard deviation. Even for N = 5000 the medians of the sampling distributions of SK_3 and SK_4 for the symmetric generating distributions are slightly less than zero. If one does not want to have any impact of a single outlier when measuring skewness, then SK_2 is preferable. However, if one wishes to take into account the outlier in the calculation, but does not want to have as severe a distortion as appears in SK_1, then SK_3 or SK_4 might be a better measure.

The impact of the single outlier on KR_1 is truly spectacular (Figure 13). The effect is maximized around at N = 1000 in which the center of box-plots are between 150 and 300 for various distributions when the true value should be between 0 and 6. Once N becomes larger than 1000, KR_1 converges towards its true value, but even with N = 5000 the medians are still between 80 and 200. The results indicate that it may not be possible to attach any meaningful interpretation to a large value of KR_1 even when there are sufficiently many observations. The Moors coefficient KR_2, on the other hand, is not influenced by the outlier at all (Figure 14). For the same reason as for SK_2, this is because KR_2 is based on octiles, which are not affected by a single observation. The centered Hogg coefficient KR_3 is mildly influenced by the outlier, especially when the number of observations is very small, but the influence disappears quickly as N increases (Figure 15). The source of the influence is the terms $L_{0.05}$ and $L_{0.5}$. As explained before, L_a is the average of the lower a percentile tails. Hence, when N is small (e.g. 50), the outlier becomes one of the percentiles and enters the calculation of L_a. On the other hand, the impact on KR_4 is present only when N = 50 (Figure 16). This is because the term $F^{-1}(0.025)$ in the formula is equal to the outlier when the number of observations is very small. The same comment can be applied to the choice of kurtosis measures: KR_2 completely ignores a single outlier while KR_3 and KR_4 reflects to some degree the presence of an outlier without much distortion.

If an outlier occurs only once and its size does not depend on the sample size (N), then its impact on any statistic will eventually disappear as the sample size goes to infinity. This must be true for even SK_1 or KR_1, despite their being severely degraded in finite samples. In reality, however, we tend to find that large outliers recur through time: for example, the 1987 stock market crash and the 1998 Asian crisis. One way of modelling this phenomenon is to use a mixture distribution. Suppose that a process $\{y_t\}$ is generated from $D(m_1, s_1)$ with probability p and from $D(m_2, s_2 = g s_1)$ with probability $1 - p$. If the probability p is very close to one and s_2 is fairly large compared to s_1, then the process has recurring outliers with probability $1 - p$ through time. In our simulations we determine the values of p and g, again using the daily S&P index. We treat a change larger than 7% per day as an outlier. Over the sample period, there are six observations whose absolute values are greater than 7%. Our estimate for p is then given by $\hat{p} = 5079/5085 = 0.9988$ and $1 - \hat{p}$ = 0.0012. The sample standard deviation of the sample without the six outliers is 0.94, and the sample standard deviation of the six outliers is 9.99. Hence, the ratio is $9.99/0.94 = 10.63$. The sample mean of the six outliers is -7.322. Taking into account these estimates, we choose the following set of parameter values for our simulations: $m_1 = 0$, $s_1 = 1$, $m_2 = -7$, $g = 10$ and $p = 0.9988$. Hence, the random numbers are generated by $0.9988D(0,1)$ + $0.0012D(-7,10)$ using the same four distributions for D. We present the true values of the skewness and kurtosis measures for these mixture distributions in Table 2. We provide a short discussion of how to obtain these values in the Appendix.

The simulation results are displayed in Figures 17-24. The behaviour of SK_1 (Figure 17) is quite different from that of SK_1 with a single outlier in that the dispersion of sampling distributions *increases* as the sample size becomes larger. This may seem surprising because we usually expect the dispersion of a sampling distribution to shrink as N increases. In the single outlier case we guaranteed the occurrence of the outlier in each sample regardless of the value of N. In the mixture distribution case, on the other hand, when N is small, the sample may not have any outliers due to the high value of $p = 0.9988$. This may explain why the dispersion of the sampling distribution of SK_1 is small when N is small, and becomes larger as N increases. In most distributions, the center of box-plots converges to the true value rather slowly: with N = 5000 it is still far from the true value; -2.27 for N(0,1), -2.00 for T-10, -1.71 for T-5 and 0.52 for Log-N(1,0.4). The other robust skewness measures (SK_2, SK_3, SK_4) are in general not influenced by this type of outlier: the dispersion is quite small even for small N and there is no finite sample bias. As for SK_1, the sampling distributions of KR_1 (Figure 21) tend to have larger dispersion as N approaches 5000. Two other measures (KR_2, KR_4) are robust to these recurring outliers and their medians converge to the true limiting values reasonably quickly (Figures 22 and 24). The Hogg coefficient KR_3 (Figure 23) has a finite sample bias for small N, but this disappears once the number of observations becomes larger than 500.

4. Application: Skewness and Kurtosis of the S&P500 Index

In this section we apply conventional and robust measures of skewness and kurtosis to the same S&P500 index data described in the previous section. The sample period of January 1, 1982 through Jun 29, 2001 yields 5085 observations, and the unit is percent return.

First we compute the conventional measures of skewness and kurtosis using all observations. For this sample period, we have SK_1 = -2.39 and KR_1 = 53.62, which is consistent with the previous findings in the literature; i.e. negative skewness and excess kurtosis. Next we use robust measures to estimate skewness and kurtosis. The results are displayed in the first column of Table 3. The outcomes are interesting in that all but one robust estimate support the opposite characterization of skewness and kurtosis. All the robust skewness measures are pretty close to zero and hence indicate that there is little skewness in the distribution of the S&P500 index. Given that all robust kurtosis measure KR_i (i = 2,3,4) are centered by the values for N(0,1), positive values (KR_2 = 0.28, KR_3 = 0.77, KR_4 = 1.16) indicate that there exists excess kurtosis, but this is considerably more mild than usually thought.

Next we re-compute all the statistics after removing the single observation corresponding the 1987 stock market crash. The results are reported in the second column of Table 3. The values for SK_1 and KR_1 are dramatically reduced to SK_1 = -0.26 and KR_1 = 6.80, but the other robust measures hardly change. This clearly shows that the single observation must be very influential in the calculation of SK_1 and KR_1, which is consistent with what we find in our simulations. On that basis, we may argue that it is difficult to attach a meaningful interpretation to the value of KR_1 (53.62) calculated using all observations. Finally, as we have done in the simulations, we remove the six observations whose absolute values are larger than 7%. The results are in the third column of Table 3. Not only the conventional measures become substantially smaller (SK_1 = -0.04 and KR_1 = 3.43) as in the case where the single crash observation is removed, but also they indicate that there exist no negative skewness and the magnitude of kurtosis is not as large as previously believed. The implication of all other robust measures are qualitatively the same; that is, no negative skewness and quite mild kurtosis.

5. Conclusion

The use of robust measures of skewness and kurtosis reveals interesting evidence, which is in sharp contrast to what heretofore has been firmly regarded as true in the finance literature. Rather than arguing that we have obtained definite evidence to refute the long-believed stylized facts about skewness and kurtosis, we hope that the current paper will serve as a starting point for further constructive research on these important issues. We do propose however, that the standard measures of skewness and kurtosis be viewed with skepticism, that the robust measures described here be routinely computed, and, finally, that it may be more productive to think of the S&P500 index returns studied here as being better described as a mixture containing a predominant component that is nearly symmetric with mild kurtosis and a relatively rare component that generates highly extreme behaviour. Viewing financial markets in this way further suggests that useful extensions of asset pricing models now embodying only skewness and kurtosis may be obtained by accommodating mixtures similar to those discussed here.

Appendix

Everitt and Hand (1981) provided the first five central moments of a two-component univariate normal mixture. Extending their results, we consider the following mixture distribution.

$$f = pg_1 + (1-p)g_2$$

where g_i has mean m_i and variance s_i^2, and is assumed to have up to 4th moment. Let $m = \int xf(x)dx$ and $v_r = \int (x-m)^r f(x)dx$. Then it can be proved that

$$m = pm_1 + (1-p)m_2$$

$$v_2 = p(s_1^2 + d_1^2) + (1-p)(s_2^2 + d_2^2)$$

$$v_3 = p(S_1 s_1^3 + 3d_1 s_1^2 + d_1^3) + (1-p)(S_2 s_2^3 + 3d_2 s_2^2 + d_2^3)$$

$$v_3 = p((K_1+3)s_1^4 + 4d_1 S_1 s_1^3 + 6d_1^2 s_1^2 + d_1^4) + (1-p)((K_2+3)s_2^4 + 4d_2 S_2 s_2^3 + 6d_2^2 s_2^2 + d_2^4)$$

where $d_i = m_i - m$, and S_i and K_i are the skewness and kurtosis of g_i respectively. Moreover, let X_1 and X_2 be the random variables governed by g_1 and g_2. In our simulation set-up, we have $X_2 = gX_2 - m_2$. Then it is easily seen that both skewness and kurtosis are invariant under this linear transformation. Hence, we have $S_1 = S_2$ and $K_1 = K_2$, and the values of S_1 and K_1 are in Table 1. All central moments up to v_4 can be now calculated in order to obtain the skewness and kurtosis of the corresponding mixture distribution.

References

Bates, D.S. (1996), "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69-107.

Bowley, A.L. (1920), Elements of Statistics, New York: Charles Scribner's Sons.

Crow, E.L. and Siddiqui, M.M. (1967), "Robust Estimation of Location," *Journal of the American Statistical Association*, 62, 353-389.

Groeneveld, R.A. and Meeden, G. (1984), "Measuring Skewness and Kurtosis," *The Statistician*, 33, 391-399.

Everitt, B.S. and Hand, D.J. (1981), Finite Mixture Distributions, London, New York: Chapman and Hall.

Harvey, C.R. and Siddique, A. (1999), "Autoregressive Conditional Skewness," *Journal of Financial and Quantitative Analysis*, 34, 465-487.

Harvey, C.R. and Siddique, A. (2000), "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, LV, 1263-1295.

Hinkley, D.V. (1975), "On Power Transformations to Symmetry," *Biometrika*, 62, 101-111.

Hogg, R.V. (1972), "More Light on the Kurtosis and Related Statistics," *Journal of the American Statistical Association*, 67, 422-424.

Hogg, R.V. (1974), "Adaptive Robust Procedures: A Partial Review and Some Suggestions for Future Applications and Theory," *Journal of The American Statistical Association*, 69, 909-923.

Hwang, S. and Satchell, S.E. (1999), "Modelling Emerging Market Risk Premia Using Higher Moments," *International Journal of Finance and Economics*, 4, 271-296.

Jorion, P. (1988), "On Jump Processes in the Foreign Exchange and Stock Markets," *Review of Financial Studies*, 1, 27-445.

Kendall, M.G. and Stuart, A. (1977), The Advanced Theory of Statistics, Vol. 1, London: Charles Griffin and Co.

Moors, J. J. A. (1988), "A Quantile Alternative for Kurtosis," *The Statistician*, 37, 25-32.

Zwet, W.R. van (1964), "Convex Transformations of Random Variables," Math. Centrum, Amsterdam.

Table 1. Values of skewness and kurtosis for various distributions

(no outlier and single outlier cases)

	N(0,1)	T-10	T-5	Log-N(1,0.4)
SK_1	0	0	0	1.32
SK_2	0	0	0	0.13
SK_3	0	0	0	0.25
SK_4	0	0	0	0.18
KR_1	3	1	6	3.26
KR_2	1.23	0.04	0.10	0.04
KR_3	2.59	0.20	0.46	0.19
KR_4	2.91	0.28	0.63	0.27

Table 2. Values of skewness and kurtosis for various distributions

(mixture distribution case)

	N(0,1)	T-10	T-5	Log-N(1,0.4)
SK_1	-2.27	-2.00	-1.71	0.52
SK_2	0	0	0	0.14
SK_3	-0.01	-0.01	-0.01	0.24
SK_4	-0.01	-0.01	-0.01	0.17
KR_1	52.76	57.33	101.47	49.18
KR_2	0.01	0.04	0.10	0.40
KR_3	0.09	0.29	0.53	0.27
KR_4	0.01	0.28	0.66	0.27

Table 3. Skewness and kurtosis of the S&P500 Index

	Using all observations	Without the 1987 crash observation	Without the 6 observations ($\geq$ 7%)
SK_1	-2.39	-0.26	-0.04
SK_2	0.08	0.08	0.08
SK_3	0.04	0.04	0.05
SK_4	0.02	0.03	0.03
KR_1	53.62	6.80	3.44
KR_2	0.28	0.28	0.28
KR_3	0.77	0.73	0.67
KR_4	1.16	1.16	1.13