American Journal of Theoretical and Applied Statistics
2013; 2(2): 38-41
Published online April 2, 2013 (http://www.sciencepublishinggroup.com/j/ajtas)
doi: 10.11648/j.ajtas.20130202.15

Linear scale dilation of asset returns

E. Reschenhofer

Department of Statistics and Operations Research, University of Vienna, Vienna, Austria

Email address:
erhard.reschenhofer@univie.ac.at

To cite this article:
E. Reschenhofer. Linear Scale Dilation of Asset Returns. *American Journal of Theoretical and Applied Statistics.*
Vol. 2, No. 2, 2013, pp. 38-41. doi: 10.11648/j.ajtas.20130202.15

Abstract: Comparing the order statistics of daily returns of the S&P 500 index from 03.01.1950 to 04.03.2013 with the corresponding rankits, a linear scale dilation is observed. This observation is used to derive a five-parameter density function for the parsimonious description of the unconditional distribution of stock returns. The typical graph of this density function looks like a wizard's hat. Its signature feature is the discontinuity at zero.

Keywords: Discontinuity, Rankits, Stock Returns, Unconditional Distribution

1. Introduction

Since the introduction of conditional heteroskedasticity models by Engle [6] and Bollerslev [5], the focus of interest has shifted from the unconditional distribution of stock returns to the conditional distribution. Of course, this does not mean that the unconditional distribution is irrelevant. Not only is it important by itself, but its properties also have implications for the conditional distribution.

The typical unconditional distribution of stock returns is leptokurtic, i.e., it has more probability mass in its center and its tails than a normal distribution. Some decades ago, non-normal stable distributions have therefore been used to model stock returns [11, 7]. However, because of empirical evidence against their crucial properties, in particular the invariance under addition [13], and their incompatibility with finite second moments, these distributions have gone out of fashion. Shifted and scaled t-distributions (or skewed versions of them; see, e.g., [8]) and mixtures of normal distributions [9] are often used instead. While mixture distributions are quite flexible, they require the selection of the number of components as well as the estimation of a possibly large number of parameters. Moreover, they are not designed for the modeling of discontinuities in the probability density function.

In this paper, a simple unconditional distribution is proposed which captures all important characteristics of stock returns and even allows for a discontinuity of the density at zero but requires only five parameters. Section 2 presents evidence of a linear scale dilation of stock returns, which gives some indication of the form of the data generating density. The indicated density is fitted to returns of the S&P 500 index. The results are given in Section 3. Evidence that there is indeed a discontinuity at zero is also provided. Section 4 concludes.

2. Comparing Rankits with Their Empirical Counterparts

The expected value of the rth largest of n independent standard normally distributed random variables is given by

$$E(r,n) = \frac{n!}{(r-1)!(n-r)!} \int_{-\infty}^{\infty} x\left(1-\Phi(x)\right)^{r-1}\left(\Phi(x)\right)^{n-r}\phi(x)dx\,,$$

where ϕ and Φ denote the probability density function and the cumulative distribution function, respectively, of the standard normal distribution. Expected normal order statistics are called rankits. Figure 1.a shows a plot of $E(n-r+1,n)$ against $r=1,...,n=1500$. B. Wheeler's function "normOrder" (R-package "SuppDists") was used for the calculation of the rankits. This function is a modification of M. Maechler's C version of Royston's [15] algorithm.

In the case of a standard normal random sample $(x_1,...,x_n)$ of size n, the rth order statistic $x_{(r)}$ can be regarded as the sample counterpart of $E(n-r+1,n)$ (see Figure 1.b). Clearly, the match is not perfect. Figure 1.c shows a plot of the ratios $x_{(r)}/E(n-r+1,n)$ against $r=1,...,n=1500$. Extreme discrepancies can occur in the neighborhood of zero because of the usually unequal numbers of positive and negative values. The problem can be alleviated by dividing the m negative order statistics $x_{(r)}$, $r=1,...,m$, by the rankits $E(2m-r+1,2m)$, $r=1,...,m$, and the $(n-m)$ positive order statistics $x_{(r)}$, $r=m+1,...,n$ by the rankits $E(2(n-m)-r+1, 2(n-m))$, $r=n-m+1,...,2(n-m)$. Figure 1.d shows a plot of these modified ratios Q_r, $r=1,...,n$, against r=1,...,n.

In the case of zero mean and non-unit variance σ^2, the order statistics must be compared with the rankits multiplied by σ. In the non-normal case, the discrepancies between the order statistics and the rankits can shed light on the nature of the deviation from normality. In the following, the sample will consist of daily log returns of the S&P 500 index. The close prices of the S&P 500 index from 03.01.1950 to 04.03.2013 were downloaded from Yahoo! Finance. To get more robust results, the total sample is divided into nine subsamples of the same size (N=12879, n=N/9=1431). Figure 2 suggests two separate linear relationships between the modified ratios Q_r, $r=1,...,n$, and the order statistics $x_{(r)}$, $r=1,...,n$, one for the negative returns and one for the positive returns. While the intercept and the slope of these linear relationships change over time, the overall picture remains the same. In each subsample, only a negligible fraction of very small returns disturbs the approximate linearity. In the next section, the linear scale dilation observed in Figure 2 will be used to construct a density function for the parsimonious description of stock returns.

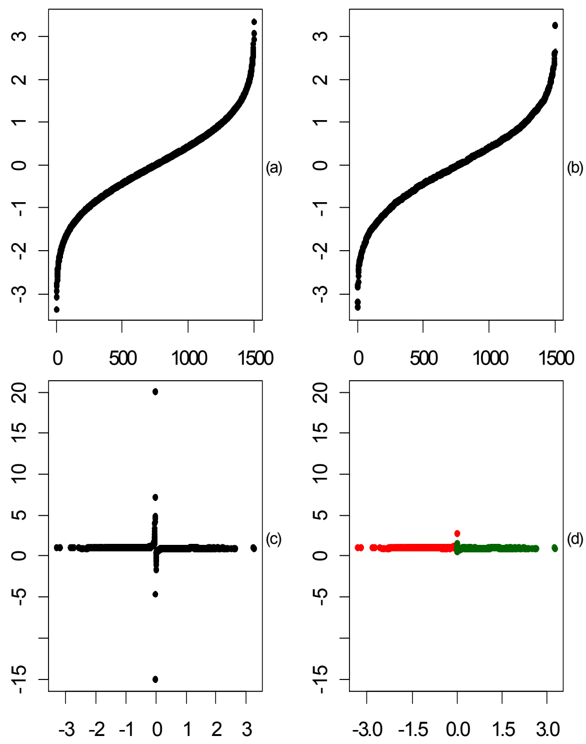

Figure 1. *Comparing rankits with order statistics*

1.a Plot of rankits $E(n-r+1,n)$, $r=1,...,n=1500$
1.b Plot of n standard normal order statistics
1.c Plot of ratios $x_{(r)}/E(n-r+1,n)$, $r=1,...,n$
1.d Plot of modified ratios Q_r, $r=1,...,n$

3. A Discontinuous Density Function for Stock Returns

For large C and not too large n, the expected values of the order statistics of a truncated normal distribution with support on [-C,C] are practically identical to the rankits. This is also true for subsets of order statistics and rankits, e.g., the first half and the second half, respectively. Let $x_{(1)} \leqslant ... \leqslant x_{(m)} < 0$ be the subsample of all negative returns. Assuming that there exists an approximate linear relationship between $Q_r = x_{(r)}/E(2m-r+1,2m)$ and $x_{(r)}$, r=1,...,m, a suitable density function $h(x)$ for the negative returns can be obtained from the truncated halfnormal density function

$$f(z)=\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}}{\frac{1}{2}-\Phi(c)}\ , \ c \leq z < 0 \qquad (1)$$

via the transformation

$$\frac{x}{z}=a+bx\ .$$

It follows from

$$x=\frac{az}{1-bz}=g(z)\ ,\quad z=\frac{x}{a+bx}=g^{-1}(z)$$

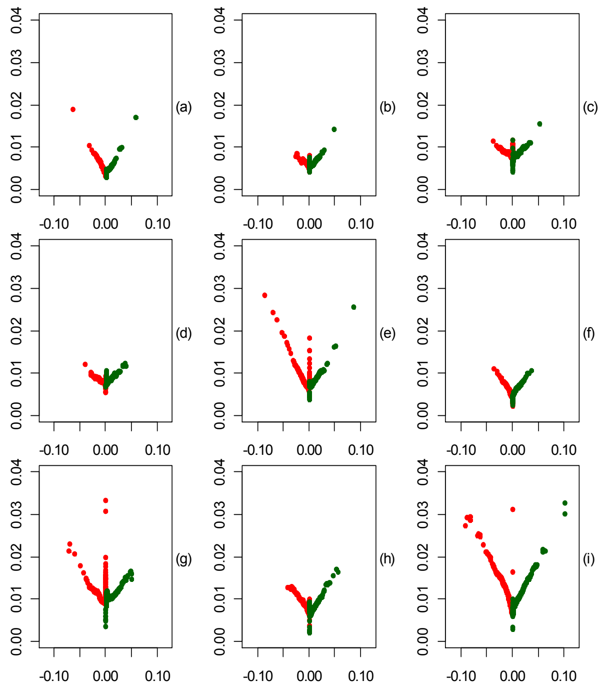

Figure 2. *Linear scale dilation of the daily returns of the S&P index from 03.01.1950 to 04.03.2013. Figures 2.a-2.i show the plots of the modified ratios Q_r against the order statistics $x_{(r)}$, r=1,..., n=12879/9=1431 for nine non-overlapping subperiods (negative returns red, positive green).*

that

$$h(x) = f\left(\tfrac{x}{a+bx}\right) \cdot (g^{-1})'(x)$$

$$= \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{a+bx}\right)^2}}{\frac{1}{2} - \Phi(c)} \cdot \frac{a}{(a+bx)^2} \,. \tag{2}$$

The restriction $g(c) = -1$ implies that

$$c = \frac{1}{b-a} \,.$$

Application of an analogous transformation to the positive returns and combination of the two density functions gives

$$w(x) = I_{[-1,0)}(x) \frac{\lambda \frac{a}{\sqrt{2\pi}}}{\frac{1}{2} - \Phi\left(\frac{1}{b-a}\right)} \cdot \frac{e^{-\frac{1}{2}\left(\frac{x}{a+bx}\right)^2}}{(a+bx)^2}$$

$$+ I_{[0,1]}(x) \frac{(1-\lambda) \frac{A}{\sqrt{2\pi}}}{\Phi\left(\frac{1}{A+B}\right) - \frac{1}{2}} \cdot \frac{e^{-\frac{1}{2}\left(\frac{x}{A+Bx}\right)^2}}{(A+Bx)^2} \,, \tag{3}$$

where λ and 1-λ are the proportions of negative and positive returns, respectively.

Figure 3 compares the density function $w(x)$ fitted to the returns of the last subperiod with a normal density and a histogram. The estimates 0.0094, -0.237, 0.0085, 0.211 of the parameters a, b, A, B were obtained by maximizing the log likelihood. The implied values of $c=1/(b-a)=-4.058$ and $C=1/(A+B)=4.556$ were sufficiently large to confirm that the truncation had no effect. The estimate of λ was 0.461.

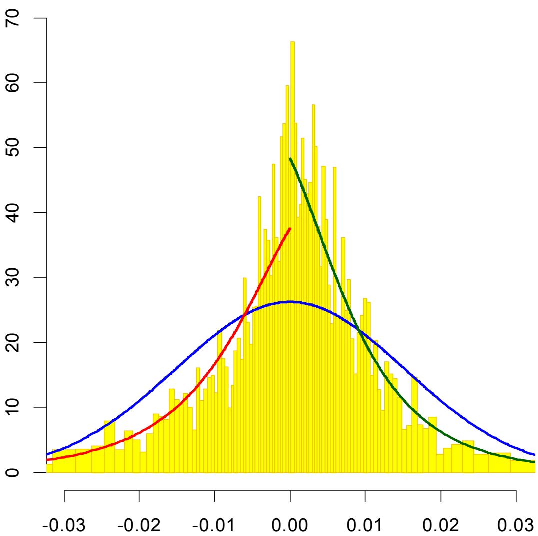

Figure 3. The density function implied by the linear scale dilation of stock returns is fitted to the daily returns of the S&P 500 index from 27.06.2007 to 04.03.2013 and compared with a normal density function (blue) and a histogram.

The breakpoints between the histogram cells were determined as the minimum of the returns, the 0.5%, 1.5%, ..., 49.5% quantiles of a normal distribution with mean zero and standard deviation 1.4826 times the median absolute deviation of the negative returns from zero, the 50.5%, 51.5%, ..., 99.5% quantiles of a normal distribution with mean zero and standard deviation 1.4826 times the median absolute deviation of the positive returns from zero, and the maximum of the returns. Figure 4 shows the fitted densities for all subsamples. In some of the older subsamples, there was a non-negligible number of zero returns. This problem was taken care of by adding one half to the negative returns and the other half to the positive returns. In general, the fit of the density funcion $w(x)$ is excellent. The typical graph looks like a wizard's hat. Its signature feature is the discontinuity at zero.

Figure 5 provides evidence that this discontinuity is a genuine feature of the unconditional distribution of stock returns rather than an artifact produced by the estimation method. For each subsample, the cumulative numbers of negative returns and positive returns times -1, respectively, in the interval (-0.0005,0) are plotted. In each case, the positive returns get closer to zero than the negative returns. Moreover, the final gradient is typically steeper in the case of the negative returns.

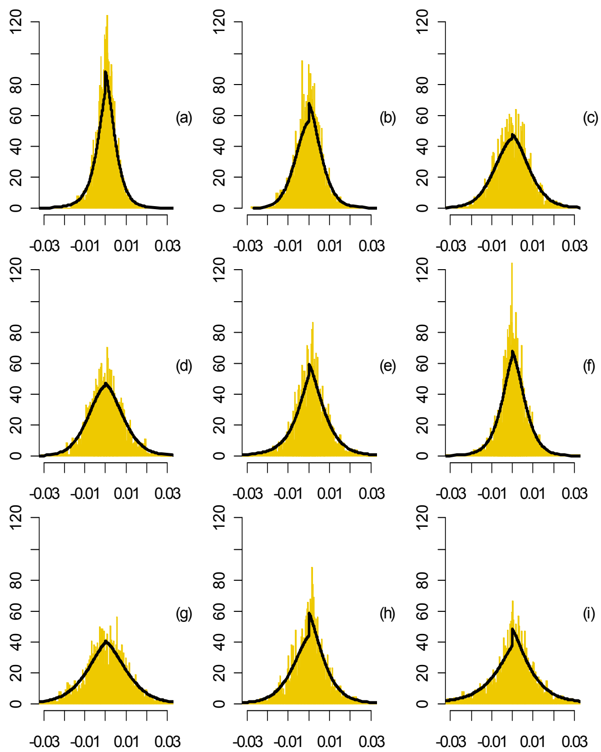

***Figure 4.** Catching returns with the wizard's hat. The daily returns of the S&P 500 index from 03.01.1950 to 04.03.2013 are divided into nine subsamples of the same size and the density w(x) is fitted to each subsample.*

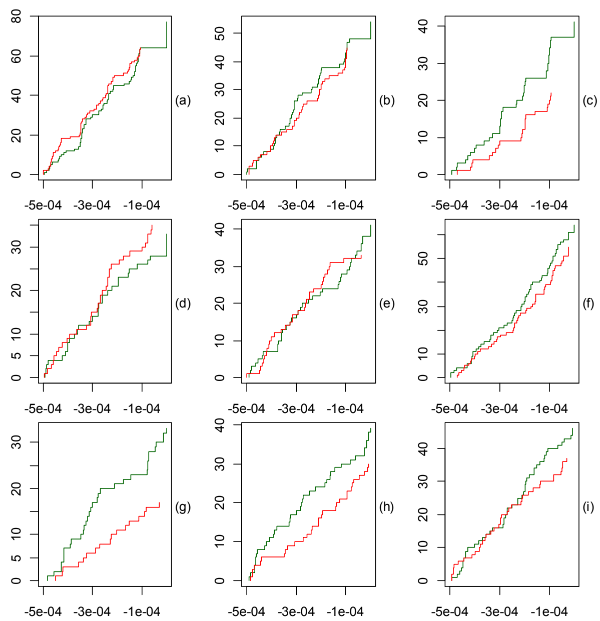

Figure 5. *Evidence in favor of a discontinuity at zero. For each of nine subsamples of the daily returns of the S&P 500 index from 03.01.1950 to 04.03.2013, the cumulative numbers of negative returns (red) and positive returns (green) times -1, respectively, in the interval (-0.0005,0) are plotted.*

4. Conclusion

In view of the many generalizations of the normal distribution [1-3, 12, 14], it is somehow surprising that none of them prevailed in financial applications. Perhaps it is just too difficult to capture all characteristics of stock returns with a single smooth density function that depends only on few parameters. The problem is that little can be gained by modeling negative and positive returns simultaneously. The apparent discontinuity at zero is an aggravating factor.

The solution proposed in this paper is to fit a two-parameter density function separately to the negative and positive returns. The form of the density was derived from a linear scale dilation observed in daily index returns. Unless additional restrictions are imposed, this approach implies a discontinuity at zero. But this is not necessarily a disadvantage because the empirical evidence presented in this paper indeed points in that direction. Ultimately, future studies will have to determine whether the wizard's hat distribution (3) is suitable for the majority of assets and sample periods.

It is hard to tell whether the two parameters for each half-axis are sufficient to deal also with extreme observations. Perhaps this must be decided on a case-by-case basis because this largely depends on the respective application. If necessary, there is, of course, always the possibility to introduce an additional parameter. A possible generalization of (2) is given by

$$h_p(x) \propto \frac{e^{-\frac{1}{2}\left(\frac{x}{a+bx}\right)^2}}{|a+bx|^p}. \tag{4}$$

Alternatively, instead of using a parametric distribution for the extremes, the tail behavior could be investigated separately with extreme values theory [10].

References

[1] A. Azzalini, A class of distributions which includes the normal ones, Scandinavian Journal of Statistics 12, pp. 171-178, 1985.

[2] A. Azzalini, Further results on a class of distributions which includes the normal ones, Statistica 46, pp. 199-208, 1986.

[3] A. Azzalini, The skew-normal distribution and related multivariate families, Scandinavian Journal of Statistics 32, pp. 159-188, 2005.

[4] F. M. Aparicio and J. Estrada, Empirical distributions of stock returns: European securities markets, 1990-95, The European Journal of Finance 7, pp. 1-21, 2001.

[5] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, pp. 307-327, 1986.

[6] R. F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, pp. 987-1007, 1982.

[7] E. F. Fama, The behavior of stock market prices, Journal of Business 38, pp. 34-105, 1965.

[8] C. Fernandez and M. F. Steel, On Bayesian modeling of fat tails and skewness, Journal of the American Statistical Association 93, pp. 359–371, 1998.

[9] S. J. Kon, Models of stock returns - a comparison, The Journal of Finance 39, pp. 147-165, 1984.

[10] F. M. Longin, The choice of the distribution of asset returns: how extreme value theory can help?, Journal of Banking and Finance 29, pp. 1017–1035, 2005.

[11] B. Mandelbrot, The variation of certain speculative prices, Journal of Business 34, pp. 392-417, 1963.

[12] S. Nadarajah, A generalized normal distribution, Journal of Applied Statistics 32, pp. 685–694, 2005.

[13] R. R. Officer, The distribution of stock returns, Journal of the American Statistical Association 67, pp. 807-812, 1972.

[14] H. T. Robertson and D. B. Allison, A novel generalized normal distribution for human longevity and other negatively skewed data, PLoS ONE 7, e37025, 2012.

[15] J. P. Royston, Expected normal order statistics (exact and approximate), Journal of the Royal Statistical Society Series C (Applied Statistics) 31, pp. 161-165, 1982.