Tail distribution of index fluctuations in World markets

Mehmet Eryiğit, Sadik Çukur, Resul Eryiğit

Department of Management, Abant Izzet Baysal University, 14280-Bolu, Turkey

Department of Physics, Abant Izzet Baysal University, 14280-Bolu, Turkey

Article Info

Article history:
Received 2 October 2008
Received in revised form 15 December 2008
Available online 22 January 2009

PACS:
89.65.Gh
89.65.-s
05.45.Tp

Keywords:
Power law
Tail distributions
World stock market indices

Abstract

We have investigated the tail distribution of the daily fluctuations in 202 different indices in the stock markets of 59 countries for the time span of the last 20 years. Power law, log-normal, Weibull, exponential and power law with exponential cutoff distributions are considered as possible candidates for the tail distribution of the normalized returns. It is found that the power exponent depends strongly on the choice of the tail threshold and a sizeable number of indices can be better fitted by a distribution function other than the power law at the region that has power law exponent of 3. Also, we have found that the power exponent is not an indicator of the maturity of the market.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The distribution of fluctuations in the stock and commodity prices and stock market indices has been one of the active areas of research in finance and econophysics for a long time. The character of the tail part of the distribution is especially important in practice because it gives an indication about the maximum risk and reward in a portfolio. Based on the efficient market hypothesis, the distribution of returns is expected to be normal which is found to be not realized in the real data more than 40 years ago [1,2]. Return distributions have fatter tails compared to the normal distribution.

The character of the tails of the distribution of fluctuations in financial time series has been investigated by using many different distribution functions. The extreme returns have been studied by fitting to a generalized Pareto and generalized extreme value distributions [3,4]. On the econophysics area, the power law is the most studied distribution. Stanley’s group has investigated the tail distribution of several stock indices as well as return of individual stocks and found a power law distribution with a universal exponent of 3 [5–7]. Following Ref. [5], many studies on the distribution in different markets have been carried out. Among these Makowiec and Gnacinski [8] have considered a five year period of Warsaw Stock Exchange index and found power law exponents of 3.06 and 3.88 for the negative and the positive tail, respectively. Some reported work find an exponential distribution for the emerging markets, such as Oh, Kim and Um [9] for the 1-minute Korean KOSDAQ index, Couto Miranda and Riera [10] for daily data of Brazilian Bovespa index in the period 1986–2000 and Matia et al. [11] for the Indian Sensex 30 index, (a reanalysis of the Indian index by Pan and Sinha [12] finds a power law tail with an exponent close to that of mature markets). Yan et al. study the Chinese stock exchanges for the 1994–2001 period and report a strong asymmetry for the negative and positive tails with power exponents of 4.29 and 2.44, respectively. Among the many different proposals for the tail distributions normal inverse Gaussian of Barndorff-Nielsen [13], exponential [14, 15], hyperbolic [16] and stretched exponentials. Also, at short time scales, a single $q$-Gaussian is found to fit the return...
distribution of the Polish WIG20-index well [17]. Even for the same market different distributions are proposed to explain the data at different economical conditions [18] (power-law for the inflationary period and exponential for the deflationary period). Recently, Drozdz et al. [19] examined the 1 min returns of S&P 500 index for the May 2004–May 2006 period and concluded that the tail is getting thinner and the exponent is larger than 3 which might be explained in part by the role of better dissemination of information on the efficiency of the markets. One of the outcome of power law research in financial return series is the stylized-fact “powerlaw tail with an exponent which ranges from two to five depending on the data set used in the study” [20]. There has been much effort to explain the microscopic origins of the power law tail [21] and the exponent of 3 [22,23].

The most widely used method for validating the power law and estimating its exponent has been criticized on statistical grounds [24]. The tail index type approaches are also criticized based on bias due to the underlying series not being iid and tail threshold dependent problems in estimating the distribution parameters [3,4].

The aim of the present study is to answer following questions: Are power law tails universal across the markets with different efficiencies? Is the power exponent of 3 universal? Is there any difference between the mature and emerging markets in terms of the distribution of the tails of the fluctuations? Is there any asymmetry of positive and negative tails? To answer these questions, we have analyzed the distributions of historical index fluctuation data of 202 indices from the stock markets of 59 countries by using statistically sound methods.

2. Data

The data in the present study consist of daily value of 202 different indices from stock markets of 59 countries. The period covers from February 8th 1988 until February 8th 2008. Since some of the indices have no data going back to 1988 we choose only the indices that were active at least for the last ten years. The data is taken from DataStream servers and a list of indices is given in the Appendix.

We consider the distribution of the normalized log return of indices. The log return for a time period \( \tau \) is given by

\[
R(\tau) = \ln S(t + \tau) - \ln S(t)
\]

where \( S(t) \) is the value of the index at time \( t \). The normalized log return is defined as

\[
r(\tau) = \frac{R(\tau) - \langle R \rangle}{\sigma},
\]

where \( \langle R \rangle \) is the time average of log return and \( \sigma \) is the standard deviation of the returns.

The summary statistics of the data is presented graphically in Fig. 1. The mean of the returns (Fig. 1(a)) is found to be dominantly positive except for some Japanese indices which have a small negative average. The distribution of standard deviations (\( \sigma \)) is shown in Fig. 1(b). We have found that among all the characteristics investigated in the present study, \( \sigma \) is the sole quantity that shows different behavior for mature and emerging market indices. In Fig. 1(b), the highest \( \sigma \)s are for Ukraine, several Chinese indices, Turkey, Taiwan, Indonesia, Poland, Romania and semiconductor and computer indices.
Fig. 2. Cumulative density function of the positive and negative tail of the returns of 202 different world indices.

from USA. Fig. 1(c) compares the distribution of skewness and kurtosis distributions with those of normal distribution. While kurtosis for all the data sets is found to be higher than that of normal distribution, both negative and positive skewness is found for the considered set. No systematic trend in value of skewness or the kurtosis was found. The fatness of the tails can be described by comparing the quantiles of the distribution of the fluctuations with the limit values of the standard normal distributions. The distribution of 1, 5, 95, and 99% quantiles of the normalized returns are displayed in Fig. 1(d). As it can be observed from the figure, both the negative and the positive tails are fat.

The possible form of the distribution of returns has been the subject of many studies, both in finance and econophysics communities. Among them we choose the ones which show fat-tail behavior and have a small number of parameters.

\[ p(x) \approx \begin{cases} x^{-\alpha} & \text{Power law} \\ x^{-\alpha} e^{-\lambda x} & \text{Power law with cutoff} \\ e^{-\lambda x} & \text{Exponential} \\ \left(\frac{x}{\lambda}\right)^{k-1} e^{-x/(\alpha\lambda)^k} & \text{Weibull} \\ \frac{1}{\lambda} \exp \left[ -\frac{(\ln x - \mu)^2}{2\sigma^2} \right] & \text{Log-normal} \end{cases} \]

Estimation of these parameters are done by maximum likelihood estimation method.

3. Results and discussion

The empirical cumulative distribution of the positive and the negative tails of normalized index returns are displayed in the Fig. 2(a) and (b), respectively. Since some of the indices are relatively young, the number of returns is not uniform for all the return series investigated in the present study.

It is well known that determination of \( \alpha \) by least-squares fitting to a power-law gives a biased exponent, while maximum likelihood estimation (MLE) based methods are robust and are more suitable for statistical inference based on the goodness-of-fit test, such as Kolmogorov–Smirnov test [25–27,24]. Since only the tail of the returns is expected to be power law distributed, it is important to find a robust way of choosing the exact starting position of the tail \( x_{\text{min}} \) [4,25,28,27,24]. The particular choice of \( x_{\text{min}} \) has two pitfalls: if \( x_{\text{min}} \) is chosen too small, the part of the distribution that is not power-law in character will be included in the determination of the exponent, so one obtains highly biased \( \alpha \). On the other hand, if \( x_{\text{min}} \) is chosen too high the number of data points included in the sample will be low and the estimated \( \alpha \) would have high uncertainty because of large statistical fluctuations. Graphical determination of \( x_{\text{min}} \) is one of the most common of those methods and can be described as the starting point of the roughly straight-line behavior of CDF or PDF on a log–log graph. The visual selection is subjective and statistically sound selection methods, such as Kolmogorov–Smirnov or Anderson–Darling type distance based selection criteria are expected to lead to better \( x_{\text{min}} \) estimates. Anderson–Darling statistics based methods are known to give a higher threshold parameter which leads to a higher power exponent [27,24]. We have used three different approaches to choosing \( x_{\text{min}} = 1 \) for all the indices, \( x_{\text{min}} \) chosen as the minimum
of Kolmogorov–Smirnov (Anderson–Darling) distance between the experimental and the theoretical distributions. The distribution of $x_{\text{min}}$ values as selected by Kolmogorov–Smirnov and Anderson–Darling statistics are displayed in Fig. 3(a) and (b), respectively. As expected [27,24], the mean value of $x_{\text{min}}$ chosen according to Kolmogorov–Smirnov statistics (left tail = 1.75, right tail = 1.52) is lower than those chosen based on Anderson–Darling statistics (left tail = 2.42, right tail = 2.39).

Initially, we have estimated the parameters of power law, exponential, Weibull, log-normal and power law with exponential cutoff distributions for all the indices by using MLE [24]. The estimation process is carried out for 1212 data sets (202 indices, positive and negative tails and 3 different choices of $x_{\text{min}}$) and the best fit distribution is chosen according to the Vuong statistics. Power law with exponential cutoff and power law distributions were found to better fit the tails for most of the cases, but log-normal and Weibull distributions were found to be comparable with a power law with cutoff for the choice of $x_{\text{min}} = 1$ for some of the indices. Based on the log-likelihood ratios (LR) and Vuong statistics, none of the tails had an exponential distribution as claimed for some of the indices before (Refs. [9,10]). One should note that the time periods covered in Refs. [9,10] and the present study are not same which might be one of the reasons for the discrepancy. Among the investigated distributions, power law-with exponential cutoff was found to be the best in explaining the tail distribution of the most of the indices especially when the threshold of the tail is chosen as $x_{\text{min}} \approx 1$. We have found no systematic trends in the estimated parameters of the distributions concerning the maturity of the market. On this part we present the distribution of the parameters of the power and power law with cutoff distribution parameters across the different indices. The distribution of MLE estimated exponent power law $\alpha$ and the exponent $\alpha$ and damping $\lambda$ for the power law with cutoff distributions for the negative and positive tails of the normalized returns of the 202 indices from the world markets are displayed Fig. 4(a)–(i) for three different choices of tail threshold. In the figure, the rows correspond to $x_{\text{min}} = 1$, $x_{\text{min}}$ determined from the minimum of KS distance and $x_{\text{min}}$ found from the minimum of AD statistics from top to the bottom. The distribution of $\alpha$ for the three different choice of $x_{\text{min}}$ indicates that the stylized-fact $\alpha \approx 3$ is a good approximation if one considers the tail starting at $x_{\text{min}} = 1$. The mean value of $\alpha$ for both the negative and positive tails is 3.03 and 3.18 in Fig. 4(a) with a standard deviation ($\sigma_{\alpha}$) of 0.18 and 0.26 for the negative and positive tails, respectively. From Fig. 4(d) and (g) it is obvious that KS- and AD-distance based choice of the threshold $x_{\text{min}}$ value increases both the mean and the $\sigma_{\alpha}$. The mean (standard deviation) of the exponents in Fig. 4(d) are 3.93 (0.76) and 3.97 (0.87) for the positive and negative tails, respectively. The same quantities for the exponents determined based on AD-tail (Fig. 4(g)) are 4.64 (0.97) and 4.69 (1.06), respectively. In all three of the cases, the power exponent for the negative and the positive tails are equal within the statistical uncertainties. Also, the distribution of left and right tail exponents seems to be almost identical as displayed in Fig. 4(a), (d) and (g). Note that this observation does not automatically mean that the left–right tail symmetry exists for all the indices considered in the present study. As is pointed out in Ref. [24] Anderson–Darling statistics prefer high values of $x_{\text{min}}$ which leads to fewer data points in the tail region and relatively large fluctuations in the estimated power exponent.

A more detailed study of the distributions for a small subset of the considered indices is presented in the Table 1. Here, we present the MLE estimated parameters and $p$-values of power law, log-normal, Weibull and power-law with cutoff distributions as well as Vuong log-likelihood and nested-hypothesis testing $p$-values. We also provide the plots of all the fitted distributions along with the empirical data for the DJI30 index in Fig. 5. The most important observation from the Table 1 and Fig. 5 is the fact that if one chooses threshold $x_{\text{min}} = 1$, all three other distributions fit the data much better than the power law. Note that the Vuong statistics take into account the fact that the power law is a one-parameter
distribution while the other three have two parameters. Based on these results, although disdained by Mandelbrot [29] on the nonexistence of financial order parameter and scale-invariance breaking properties, power law with cutoff seems to be the best distribution to explain the tail if one considers the cutoff value of $x_{\text{min}} = 1$. A similar exponentially dampened power law model was used by Wu [30] to account for the decline of return non-normality with time aggregation.
We have also considered the possibility of the return fluctuations having a power distribution with different exponents in shorter time intervals. In such a situation, the distribution over all the period would be a superposition of power laws with different exponents which might be another distribution. The power exponent of DJI index in the period 10.01.1928–12.09.2008 is estimated and displayed in Fig. 6. A sliding time window of length 3000 days with sliding time of 300 days is used to estimate the time-dependent parameters of the distributions and the Vuong statistics. In Fig. 6 we display the power exponent $\alpha$ for the positive and negative tails for the choice of $x_{\text{min}} = 1$, but the general behavior of $\alpha$ for...
the other choices of $x_{\min}$ is similar. The time variation of $\alpha$ is not monotonous since in the mid 1970s the fluctuations in $\alpha$ are small. The Vuong statistics in Fig. 6(b) indicate that for the particular choice of $x_{\min} = 1$, power law with exponential cutoff distribution is a better choice for explaining the tail of the normalized returns for DJI index.

4. Conclusions

We have investigated the tail distribution of the returns of a large number of world indices. Log-normal, Weibull, power law and power law with cutoff are considered as candidate distributions. It is found that the answer to the question “which distribution fits the tail the best” depends on the choice of the threshold of the tail. If a common $x_{\min} = 1$ is chosen as the threshold, the widely accepted $\alpha \approx 3$ exponent is found, but Vuong and nested-hypothesis testing statistics suggest that all three other distributions are much better at explaining the tail distribution. On the other hand, if $x_{\min}$ is chosen based on some statistically sound criteria (such as Kolmogorov–Smirnov or Anderson–Darling distance) the mean value of power exponent is larger than 3 and its dispersion is larger compared to the $x_{\min} = 1$ case. We have found no indication about the tail index or tail distribution–maturity relation. Concerning negative–positive tail asymmetry, we have found that although there might be some differences in the tail exponents for the left and right tails for a particular index, the distribution of tail exponents across different markets is symmetrical.

Appendix

List of the DataStream index codes investigated in the present study:
Argentina(ARGSIBO, ARGMERV), Australia(ASXAORD, ASX100I, ASX20LD, TTOSP60), Austria(AXIN50, AXIN5Z), Belgium(TOTMKBG, BGMIDCP), Canada(TTOSP60, TTOCOMP, SPDCNX), Check Republic(CZPX50I), China(CHSH180, CHSASHR, CHSSBHR, CHSCOMP, CHZASHR, CHZASUB, CHZBSHR, CHZBSUB, CHZCOMP), Chile(IGPAGEN, IPSASEL), Croatia(CTCROBE), Denmark(TOTMKDK), Egypt(EGHRMGL, EGHRMBP), Estonia (ESARIPA, ESTALSE), Finland(TOTMFN), France(FRC40AC, FSB250), Germany(CDAXTEC, DAX200A, DAXINDEX, DAX40AVAL, MDAIXDAX, SDAIXDAX), Greece(TOTMKGR, FTASE4P, FTASE80), Hong Kong(HNGKNGI, HKHCHA, HKCHIE, HKHCOMP, HKHSFIN, HKHHKCI, HKHHKLC, HKHMLCI), Hungary(BUXINDX, HUNESCI, TOTMKHN, ATX-HUN), Iceland(ICEX15I, ICEXALL), India(IBOMBSBE, IBOM200, IBOM500, IBOMSEEN), Ireland(TOTMKIR), Israel(ISTA100, ISTA75I, ISTGNR, ISTGNS, ISTMAOF), Indonesia(JAKDBXI, JAKMIXI, JAKCOMP, JAKLQ45), Italy(ITALMBI, MITBELNI, ITMIDEX), Japan(JAPDOWA, NIK300, JAPOTCI, TOKYOSE, TSET100, TSET500, TSECR30, TSEG70, TSE400, TSESCOS), Jordan(AMMANFM), Korea(KORCOMP, KOR100I, KOR200I, KOR500I, KORLCAP), Malaysia(KLPCOMP, KLSEM), Kuwait(KSEICOS), Luxemburg(LXLUXXI, TOTMKLX), Mexico(MXSEANAM, MXIMC30, MXINMEX, MXIP35), Morocco(MDCFG25), Netherland(AMSOKAP, AMSSMKAZ, NLALSHR), Norway(TOTMKNW), New Zealand(NZ30CAP, NZSEALL, NZ10CAP), Pakistan(PKSE100), Peru(PERUTXI), Philippines(PSEALL, PSEXCOMP), Poland(TOTMKPO), Portugal(POP210, POPSIGN), Romania(RMBETRL, RMBETU), Russia(RSRTSB, RSRTSFX), Singapore(TOTMKSG), Slovakia(SXSA12), Slovenia(SLOEBOIO, SLOESI, SLOPIX), South Africa(JSEOVER, JSEIN25, JSEAL40), Spain(TOTMARKB, IBEX35I, SWISS(SWSEALI), Sweden(SWEDOMX, SWEBCEN, SWSEALU), Thailand(BGT, TOTMKT), Tunisia(TUBVMIN, TUTUNIN), Turkey(TRKISTB, TRKNT30, TKNATA, TOTMKTK), Ukraine(UKRRDRAG, UKRPFTS), USA(AMEX/AMXCOMP, AMXMINT, AMXMAJ, AMXOILS), Dow Jones(DJCOMP65, DJINDUS, DJTRSP, DJTILS, WILEQTY), NYSE(NYSEALL, NYSEFIN, NYSENEI, NYSEMT, NYSE100, NYSEWDL), Philadelphia(PHLXAU, PHLXSOX, PHLXY), Russia(ROSSEL, ROSSL3, ROSSL2, ROSSL1), NASDAQ/NAS100, NASBANK, NASBIOT, NASCOMP, NASMCT, NASINDS, NASINR, NAS- FINS, NASTELC, NASTRAIN), S&P100, S&P400, S&P500, KCVALR), UK(FTSE100, FTSE100E, FT250VA, FTSEAI, FTALLS, FTALLX, FTSESIO, FTSESC, FTTM10E, FTTM100), Venezuela(TOTMV).

References