Bulletin of the ISI

Vol. LXII

PROCEEDINGS

Lisboa, Portugal

22-29 August Août 2007

56th Session of the ISI

22-29 AUG Lisboa 2007

INTERNATIONAL STATISTICAL INSTITUTE

Bulletin of the International Statistical Institute

Volume LXII (2007)

GUEST EDITORS

****

M. Ivette **Gomes**, *DEIO* (*FCUL*) and *CEAUL, Universidade de Lisboa*

José Alberto **Pinto Martins**, *Instituto Nacional de Estatística*

José Alberto **Silva**, *Instituto Nacional de Estatística*

****

ISBN: 978-972-673-992-0

November 2008

Assessing fat-tailed sequential forecast distributions for the Dow-Jones index with logarithmic scoring rules

Sanfilippo, Giuseppe
Agrò, Gianna
University of Palermo, Dipartimento di Scienze Statistiche e Matematiche "S.Vianelli"
Viale delle Scienze ed. 13
Palermo, 90128, Italy
E-mail: sanfilippo@unipa.it, agro@unipa.it

Lad, Frank
University of Canterbury, Department of Mathematics and Statistics
Private Bag 4800
Christchurch, 8020, New Zealand
E-mail: F.Lad@math.canterbury.ac.nz

1. Introduction

This article displays the computational results of our attempt to address the recognition that sequences of return rates from statistical indices of stock or bond prices exhibit fatter tails in their histograms than are representable via a Normal distribution. Results of the initial investigations by Timmerman (1995) and Mantegna and Stanley (1995) have been supported in investigations of a variety of specific phenomena such as the work of Lim et al (2006) on exchange rates and currency options, to name only one line of work. Many hundreds of related research reports circulate on the web and in published journals.

We shall use the space available in this contribution to report on the substantive results of our analysis without describing every technical detail of their generation. These are available to the interested reader in an extensive research report by Agrò, Lad and Sanfilippo (2007). In a similarly abbreviated style, we shall merely mention relevant technical background in summary and provide appropriate references.

Technical developments in distribution theory have revitalised our ability to assess uncertain prospects for unknown quantities via fat-tailed distributions that are incorporated in the family of so-called *Exponential Power distributions*, or EP distributions, for short. The Normal distribution is one member of this family, demarked by a parameter value in the family structure that distinguishes fat-tailed members from thin-tailed members. The initial work of Subbotin (1923) in defining this family has been extended with varying terminologies in works such as Box and Tiao (1953), Vianelli (1963) and Agrò (1995, 1999).

Finally, our work lies within the operational subjective tradition of statistical analysis, championed in the foundational works of Bruno de Finetti (1937, 1974, 1975) and explained in the practical computational text of Lad (1996). In a word, de Finetti's outlook proposes that sequences of observations are *not generated by* random processes whose distributional structure is meant to be estimated. Rather, probability distributions are meant to represent the uncertain attitudes of investigators, usually via some form of partially exchangeable mixture distributions. The sequential procedure of this methodology proceeds as follows: a mixture forecast distribution is assessed for the first datum; the datum is observed and the forecast density is scored according to the logarithm of the forecast density value at that observation value; that datum is used to update the mixture distribution to be applied to the second datum in the sequence; when observed, that distribution is scored at the observed value in the same way; the datum is used to update the mixture distribution to be applied to the third datum in the sequence; and so on.

The data we shall assess in this way are the daily closing values of the Dow-Jones Index of stock prices, beginning 25 October, 1984 and running through the early months of 2007. These 23 years include some 5678 data points.

With this brief introduction, we present here the following graphical results. We hope they will entice you to study the details of their production and a detailed discussion of their content, in our research report referenced in the second paragraph of this Introduction. The graphs presented here are (perhaps) produced in black-and-white, but are still understandable. The research report on the net displays them in colours, with better effect.

2. Four predictive densities at the beginning and end of 23 years of forecasting

The structure of the analysis was meant to characterise four different statistical assessments of the daily rates of return of the Dow-Jones index in 1984, by different people who largely agreed in their assessment of the first day's prospects, but who disagreed in their claims regarding how the data would be used to change their subsequent assessments of the variability in the index. Figure 1 displays on the left the four densities for the third day's rate of return in the sequence, and, on the right, the four associated forecasting densities after some 23 years of observations. The abbreviated names for the densities should be understood throughout these and the remaining graphs as: MixLC3EP = mixture of a linear combination of three Exponential Power distributions; MixEP = mixture of Exponential Power distributions; MixLC3N = mixture of a linear combination of three Normal distributions; and MixN = mixture of Normal distributions. Details of the mixing structures will need to be read in the referenced technical report. It should be noticed how much fatter the MixN density becomes *in the center* of the density relative to the others. The others there are fatter in their tails.

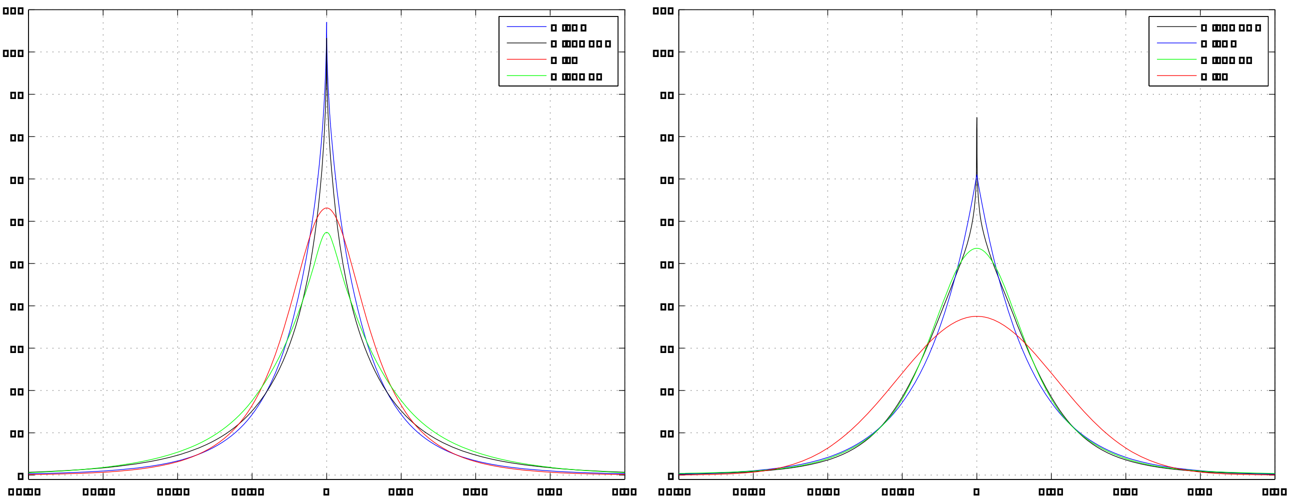
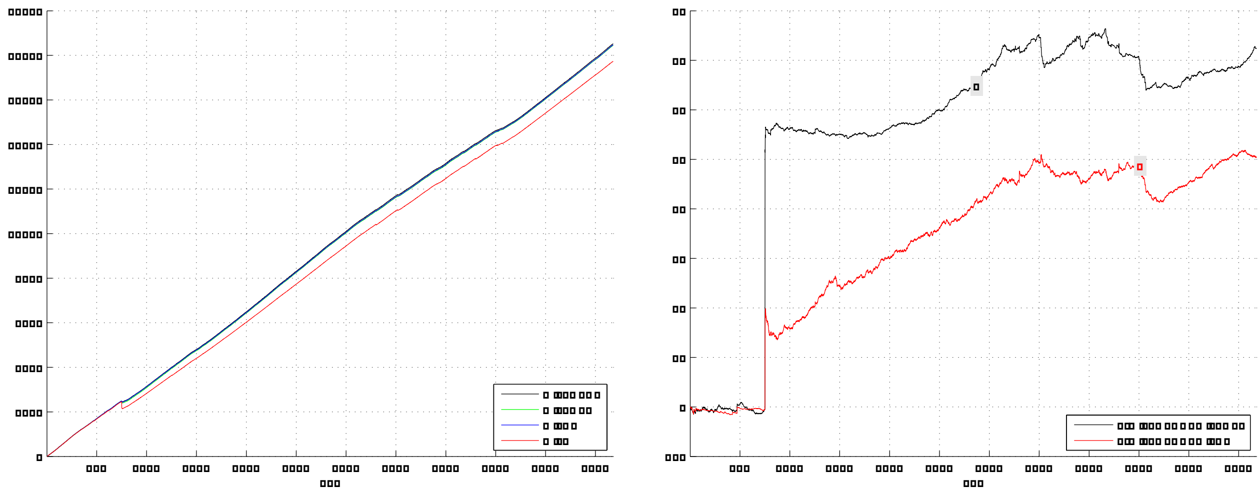


Figure 1. Four predictive densities for the rate of return on the third day (left) and for the day following 23 years (right) of trading in the sequence.

3. Cumulative logarithmic scores of the four predictive densities for daily returns

We present the proper logarithmic scoring results in two graphs. Figure 2(a) displays the cumulative sequential scores of all four distributions. It is apparent that the quality of the mixture Normal forecast falls off drastically after the sharp drop in prices that occurred on 19 October, 1987. Cumulating scores for the other three densities are visually so close to each other at this resolution over the long series that Figure 2(b) is required to appreciate the differences among them. As the

MixLC3EP distribution achieves the best score, Figure 2(b) displays the cumulating differences in the scores between MixLC3EP and MixEP (which achieves the second best score) and between MixLC3EP and MixLC3N (which achieves the third best score, thus making this difference greater.)



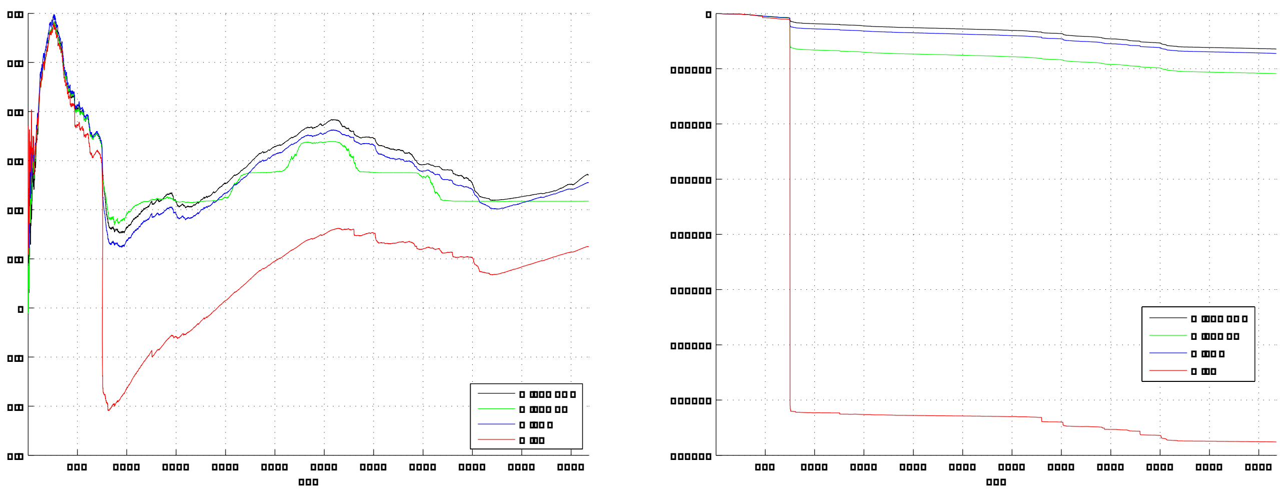
(a) Gross cumulative logarithmic scores of the four forecasting distributions.

(b) Differences between the cumulative logarithmic scores of MixLC3EP and the scores of MixEP and MixLC3N.

Figure 2. Proper logarithmic scoring rules results.

4. Comparing the expected scores: Entropies and their quadratic scores

Just as a forecaster is unsure of the rate of return to be observed each day, assessing it with a distribution, he/she is also uncertain of what logarithmic score will be achieved by the pronounced distribution. The expected logarithmic score by each assessor equals the (negative) *entropy* of the distribution. In turn, this expected score (that is, the assessor's negative entropy) can be scored relative to the achieved logarithmic score via a quadratic score. Figure 3 displays the sequence of the (negative) entropies and their quadratic scores for the four distributions we have studied. It shows that not only does the MixLC3EP distribution develop to entail the greatest negative entropy expectations



(a) Sequential negative entropy information measures of the four forecasting distributions.

(b) Cumulative quadratic scores of the sequential entropies of the four forecasting distributions.

Figure 3. Sequence of the negative entropies and their quadratic scores.

among the four, but it achieves the best score of its entropy too. Intuitively, this means that this distribution both contains the most information among the four, and also has the best understanding of the amount of information it asserts regarding the price series.

REFERENCES

Agrò, G. (1995) Maximum likelihood estimation for the exponential power function parameters, *Communications in Statistics: simulation and computation*, **24**, 523-536.

Agrò, G. (1999) Parameter orthogonality and conditional profile likelihood: the exponential power function case, *Communications in Statistics: theory and methods*, **28**, 1759-1768.

Agrò, G., Lad, F. and Sanfilippo, G. (2007) Sequential forecasting of economic indices using mixtures of linear combinations of EP distributions. Available on www.unipa.it/∼sanfilippo/isireport/

Box, G.E.P. and Tiao, G.C. (1973) *Bayesian Inference in Statistical Analysis*, New York: John Wiley.

de Finetti, B. (1937) La prévision, ses lois logiques, ses sources subjectives, *Annales de L'Institute Henri Poincaré*, **7**, pp. 1-68. H. Kyburg (tr.) Foresight, its logical laws, its subjective sources, in H. Kyburg and H. Smokler (eds.) *Studies in Subective Probability*, second edition, 1980, New York: Krieger.

de Finetti, B. *Theory of Probability: a critical introductory treatment* (1974,1975) 2 volumes, A.F.M. Smith and A. Machi (trs.), New York: Wiley. Translation of *Teoria della probabilità: sintesi introduttiva con appendice critica* (1970) Torino: Einaudi.

Lad, F. (1996) *Operational Subjective Statistical Methods: a mathematical, philosophical, and historical introduction*, New York: John Wiley.

Lim, G.C., Martin, G.M. and Martin, V.L. (2006) Pricing currency options in the presence of time-varying volatility and non-normalities, *Journal of Multinational Financial Management*, **16**(3), 291-314.

Mantegna, R.N. and Stanley, H.E. (1995) Scaling behaviour in the dynamics of an economic index, *Nature*, **376**, 46-49.

Subbotin, M.T. (1923) On the law of frequency of error, *Mathematicheskii Sbornik*, 296-330.

Timmerman, A. (1995) Scales and stock markets, *Nature*, **376**, 18-19.

Vianelli, S. (1963) La misura della variabilità condizionata in uno schema generale della Curve Normali di frequenza, *Statistica*, **23**, 447-473.

ABSTRACT

We use the logarithmic scoring rule for distributions to assess a variety of fat-tailed sequential forecasting distributions for the Dow-Jones industrial stock index from 1980 to the present. The methodology applies Bruno de Finetti's contributions to understanding how to compare the quality of different coherent forecasting distributions for the same sequence of observations, using proper scoring rules. Four different forms of forecasting distributions are compared: a mixture Normal, a mixture of convex combinations of three Normal distributions, a mixture exponential power distribution, and a mixture of a convex combination of three exponential power distributions. The mixture linear combination of three EP distributions achieves the best score on a fairly regular basis, followed by the mixture EP and the mixture linear combination of three Normals. All three make marked regular improvements in assessing volatility phenomena (tail behaviour of the distributions) compared to the Normal distribution, with an especially noticeable improvement achieved after the information gained from the drastic fall in stock prices that occurred in October, 1987. The mixture EP distributions are designed to incorporate fat-tailed properties into the forecast probabilities. It is surprising that the mixture linear combination of three Normal distributions fairs as well as it does in the comparison of scores. Overall, the methodology provides a practical improvement in comparing the quality of statistical forecasts over the so-called "model-fitting" procedures that have long been used in the statistical assessment of contentious economic issues. The results of the scored forecasting exercise are supported by an analysis of the expected scores via the entropies of the forecast distributions, and a quadratic score of these expectations, as well.