{file wigSP.pdf February 21, 2007}

Comparing the distributions of the WIG20 and S&P500 index

Anna Bartkowiak
Institute of Computer Science, University of Wrocław
Joliot-Curie 15, 50-383 Wrocław, PL, e-mail:aba@ii.uni.wroc.pl

Aim of the paper (Draft)

Our aim is to compare the distribution of two stock indices: the Polish WIG20 and the American S&P500. We consider the CLOSING price of these indices (X) in the years 1994–2006. The observed values x_t (denoting the closing price at day t) were transformed to $z_t = ln(x_{t+1} - ln(x_t)$ appropriately. The derived variable Z will be called *daily return*.

We have 2 goals:

1. to compare the daily dynamics of WIG20 and S&P,
2. to model mathematically the distribution of daily returns of these stocks.

Motivation for *goal no. 1.* The two stocks are existing in two different economy systems. WIG20 is relatively young; S&P has long tradition. Is the behavior of the two series - with elapsing time - similar?

Motivation for *goal no. 2.* The attempt of modelling statistically or mathematically the distribution of the variable Z (stock returns) has a long history. Already in 1900, Bachelier proposed the first model for the stochastic process of returns (quoted after Gopikrishnan and al., 1999). Bachelier proposed the model of an uncorrelated random walk with independent, (i.i.d) random variables. However, it seems that this is only a rough approximation of the true model governing the phenomenon. Some researchers (see, e.g. Kon, 1984, Gopikrishnan and al., 1999, Brabazon and O'Neil, and quite a lot of other researchers) argue, that the distribution of returns, when observed in short time intervals (days, or smaller time intervals) is not normal. This is stated by calculating the *kurtosis* of the distribution. Also, it is believed, that the variable Z has heavy tails and is leptokurtic. Other supposed alternatives are: Student's t distribution, mixture of Gaussians or Students' t.

Usually, to our knowledge, only univariate Student's t was considered. We do it considering bi-variate Student's t distribution and apply in that case Mardia's tests for normality of distributions.

We got a lot of interesting results.

1. The autocorrelation functions of the observed variable X look similar for both indices, although that of the Polish WIG20 is decidedly more ragged (see draft figure 1).

2. The hypothesis on the Gaussianity of the distributions should be rejected. The departure from Gaussianity is caused rather by a higher concentration near the mean than by heavy tails (see figures 2 and 4). Formally, the kurtosis for WIG20 is higher than that for SP500 (Figure 3).

3. We stated also a problem about calculating the kurtosis from samples containing single outliers. Applying robust methods one obtains much heavier tails - which is problematic. Other methods - based on quantiles should rather be considered (see for this topic: Schmidt and Trede, 2003).

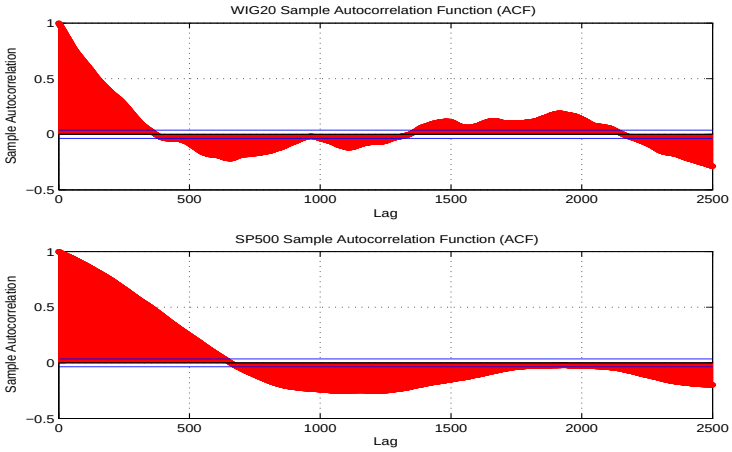

Figure 1: Sample autocorrelation function for WIG20 (top exhibit) and SP500 (bottom exhibit). When comparing both exhibits, one should take into account that the number of stock working days per year is smaller in Poland then in US

References

[1] Archambeau, C., Verleysen, M., Robust Bayesian Clustering. Neural Networks 20 (2007), 129-138.

[2] Brabazon, A., O'Neil, M., Biologically Inspired Algorithms for Financial Modelling. Springer Berlin-Heidelberg, 2006.

[3] Egan W.J., The distribution of S&P Index Returns. (January 6, 2007). Available at SSRN: `http://ssm.com/abstract=955639`

[4] Gopikrishnan, P., et al., Scaling of the distribution of fluctuations of financial market indices. arxiv:cond-mat/9905305 v1 20 May 1999.

[5] Junker, M., Szimayer, A., A probability density applying to different types of heavy tail behaviors. `www.caesar.de/english/research/project_groups/fe_publications.html`

[6] Kon, S. J., Models of Stock returns - A Comparison. The Journal of Finance, 39 (1984), no. 1, 147-163.

[7] Mardia, K. V. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. Sankhya: The Indian Journal of Statistics 36 (1974), B Pt. 2, 115-128.

[8] Schmid, F., Trede, M., Simple tests for peakedness, fat tails and leptokurtosis based on quantiles. Computational Statistics and Data Analysis 43 (2003), 1-12.

[9] Svensen, M., Bishop, C.M., Robust Bayesian mixture modelling. Neurocomputing 64 (2005), 235-252.

[10] Vlassis, N. and Likas, A., A kurtosis-based dynamic approach to Gaussian mixture modelling. IEEE Trans. on Systems, Man and Cybernetics. Part A: Systems and Humans, 29 (1999), No. 4, 393-399.

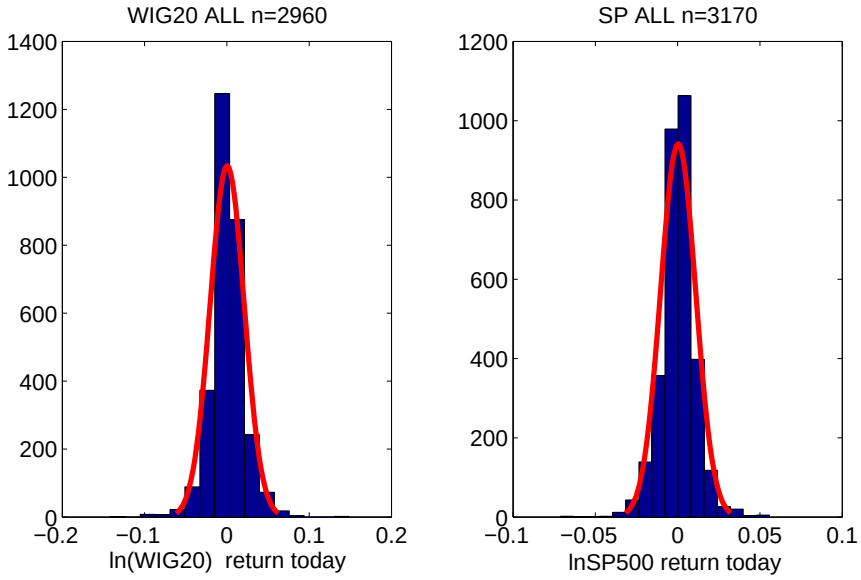

Figure 2: Histogram - with normal curve overlaid - for all notations of WIG20 (left) and SP500 (right) for all observations over the the period of 13 years. Notice the peakedness at the centers and lack of pronounced tails

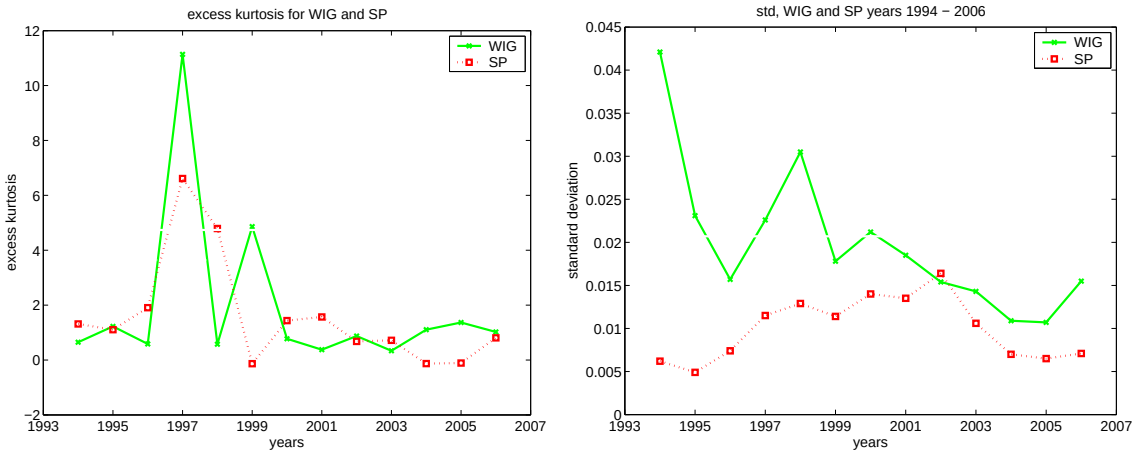

Figure 3: Excess Kurtosis (left) i standard deviation (right) in the years 1994-2006. The Reader should remind that excess kurtosis for normal distributions amounts zero. One may notice that WIG20 exhibits for all investigated years values markedly greater then zero. The kurtosis for SP500 is milder. Both indices attain in 2002 and 2003 $\approx$ equal values. The total kurtosis - not shown in the graph - amounts for Wig20: 4.74, and for SP500: 3.64 respectively

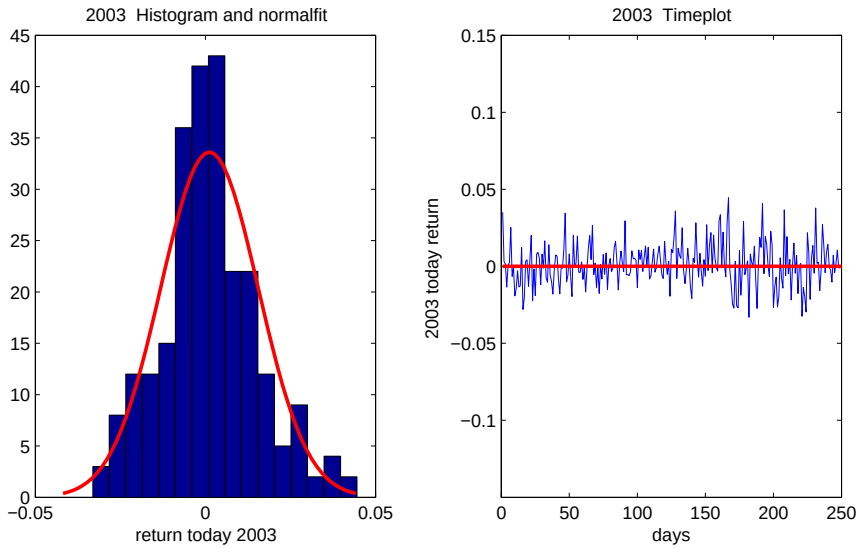

Figure 4: Distributions of WIG20 returns in 2003. Again one may notice the concentration about the mean and short tails