Two experiments in the stability of stock statistics

Burton Rosenberg
Department of Mathematics
and Computer Science
University of Miami

Abstract

This paper announces support in the form of the Spearman rank correlation test for the hypothesis: stock variance is a stable commodity, but the covariance of stocks varies randomly. Among the consequences of this hypothesis are:

1. *Arbitrage equations involving covariances do not constrain the marketplace.*
2. *Variance is a stable commodity whose price is set by the arbitrage opportunities it presents.*
3. *Portfolio theories depending on estimates of future stock covariances are not at present useful theories.*

The result is not unexpected, however the conclusions challenge some of the existing literature.

1 Introduction

The theory of the efficient portfolio aids the investor to stabilize available capital, as well as provides a justification for risk-return payoff. However, to calculate with the theory, estimates of market variances and covariances are required. The question arises as to how well past variances and covariances predict future variances and covariances. Besides this motivation, whether variances and covariances can be valuable assets depends on whether they are stable. A predictable market behavior might be combined with an arbitrage opportunity thereby pricing the market behavior. If, say, covariances had no predictability, they would also be of no value as an asset.

The theory by which an optimal portfolio is calculated is due to Markowitz [6]. For that theory, variance and covariance information is required. One possibility would be to calculate the historical variances and covariances of a universe of stocks and bring these values forward to the next time step. We decided to question this supposition. Weakening the requirements, we tested only how the rankings of stocks from least to most variant and the rankings of stock pairs from least to most covariant change from time step to time step. In other words, is it true that a high variance stock this year will be a high variance stock next year? Will a high covariance stock pair continue to covary strongly during the next year?

According to the methods of this paper, it is true that the variance of a stock moves with the stock into the next time step. Variance is a property of the stock and in this sense we say it is stable. However, they cannot confirm that covariance is stable, in the following sense: the distribution of the Spearman rank correlation coefficient for the two orderings of stock pairs by covariance during consecutive time periods is essentially the distribution achieved by taking two independent, random orderings.

The problem of prediction of stock price movements has been previously studied from the time-series standpoint. The work of Granger and Morgenstern [4] uses the classical techniques of Fourier Analysis to study the spectral qualities of stock price movements. A very detailed study of stock price variance has been undertaken by Shiller [7] in order to bring into accord observed variance and the efficient market hypothesis. In addition, there has been much work done using the ARCH model introduced by Engle [2] and the extension GARCH model introduced by Bollerslev [1]. These heteroscedastic models assume that stock price is a gaussian normal random variable with time varying variance, the variance predicted according to the parameters of the model.

This paper attacks the problem from a different angle. We consider it a problem in hypothesis testing, rather than one of model fitting. Furthermore, we use nonparametric methods and thus have no hypotheses on distributions.

2 The experiments

Two experiments are described. The first experiment, summarized in Figures 1 and 2, uses a data set of 212 stocks containing records of at least 280 closing prices since October 30, 1993. The data was taken from MIT's Stock Market Project [8]. We use this data to test quarter and semi-annual data streams running from the third quarter of 1993 until the fourth quarter of 1994. This data is of limited depth in time, but does give us a large population of stocks to work with.

The second experiment, summarized in Figures 3 and 4, uses the CRSP data set on thirteen stocks running from July, 1962 through December, 1992. We corrected this data for splits but not dividend disbursements. This data was used to test a stream of annualized variance and covariances, normalized for means, of prices from 1963 until 1993.

The experiment on the stability of a stock's variance compares two time periods, j_1 and j_2 for a sample R of N stocks, picked from our universe of stock data. In our experiment, j_1 and j_2 are consecutive quarter, semi-annual or annual periods. Sorting by variances during each of the time periods gives us two orderings α and β of the stocks, from least to most variable:

$$\mathrm{Var}\,{\alpha_1}^{j_1} \le \mathrm{Var}\,{\alpha_2}^{j_1} \le \ldots \le \mathrm{Var}\,{\alpha_N}^{j_1}$$

and,

$$\mathrm{Var}\,{\beta_1}^{j_2} \le \mathrm{Var}\,{\beta_2}^{j_2} \le \ldots \le \mathrm{Var}\,{\beta_N}^{j_2},$$

where α_i, β_i are the various stocks in the sample R.

The rank of a stock $r \in R$ under the α order is the i such that $\alpha_i = r$,

$$\mathrm{Rank}_\alpha(r) = \{\, i \mid \alpha_i = r \,\}, \quad \text{any } r \in R.$$

Likewise,

$$\mathrm{Rank}_\beta(r) = \{\, i \mid \beta_i = r \,\}, \quad \text{any } r \in R.$$

We wish to compare these two rankings in order to reject the possibility that there is no significant influence of the past on the future. Spearman's rank correlation coefficient [3], [5] is the correlation of ranks under the two orders:

$$r_R = 1 - \frac{6 \sum_{r \in R} \bigl(\mathrm{Rank}_\alpha(r) - \mathrm{Rank}_\beta(r)\bigr)^2}{(N+1)N(N-1)}.$$

Similarly, consider $S = R^{(2)}$ the collection of all distinct stock pairs, and select a size N subset $R \subset S$. Two orders α and β can be defined for consecutive time periods j_1 and j_2,

$$\mathrm{Cov}\,{\alpha_1}^{j_1} \le \mathrm{Cov}\,{\alpha_2}^{j_1} \le \ldots \le \mathrm{Cov}\,{\alpha_N}^{j_1}$$

and,

$$\mathrm{Cov}\,{\beta_1}^{j_2} \le \mathrm{Cov}\,{\beta_2}^{j_2} \le \ldots \le \mathrm{Cov}\,{\beta_N}^{j_2},$$

and Spearman's coefficient is calculated to compare the two rankings.

For Experiment 1, where N is large, if the two rankings were chosen independently at random, r_R would be approximated as a zero-mean, $1/(N-1)$ variance normally distributed random variable. In Experiment 1, the underlying space of events is the choice of subset R. We calculate,

$$z = r_R \sqrt{N-1}.$$

The event,

$$|\, z \,| \ge 2.575,$$

will occur only 1% of the time if α and β were independently chosen orders.

For small N, such as Experiment 2 where $N = 6$, the Spearman coefficients are compared in a table of theoretically calculated values [5]. The approach here is to consider the set of stocks fixed and the randomized event to be the choice of a pair of years.

In fact, we exhaustively use all consecutive years within the range of our data set.

The first experiment uses the MIT data set and is summarized in Figures 1 and 2. Three subexperiments are cited, each subexperiment had eighteen trials. In Figure 1, the third quarter of 1993 is denoted 93.3, and so on, and the first half of 1994 is denoted 94.1–2, and so on.

For the variance subexperiment, two time periods were selected and forty distinct stocks were picked uniformly at random from the population of 212 stocks. The variance of these stocks were calculated and ranked for the two time periods, and the Spearman correlation coefficient derived. Since N is large, the z value is calculated and shown in Figure 1. This was done for each of eighteen trials, that is, eighteen selections of forty stocks.

The covariance subexperiment was similar, however forty distinct stock pairs were selected rather than forty stocks. The random selection was done by selecting uniformly at random twice from the population of 212 stocks and throwing out the choice if the pair has already been chosen or if the two choices happen to be the same stock.

The subexperiment "Random" consisted of selecting forty pairs of values uniformly at random. That is, if variance or covariance were truly random, it could yield z values as in this subexperiment.

The data shows that the hypothesis of independence is rejected for variance. However, with each time period, the ranking of covariance appears to shuffle almost as unpredictably as Random. This is illustrated in Figure 2, where cumulative probabilities have been totaled and graphed, along side a normal distribution.

The second experiment uses the CRSP data set and is summarized in Figures 3 and 4. Fourteen stocks were selected at random from the CRSP data base, provided that their histories ran from 1962 through 1992. The prices were adjusted for splits, at which time one of the fourteen was rejected because of a long period of missing price information. In one covariance subexperiment, the thirteen data sets were arranged into six pairs, leaving one stock out, and the covariances were rank correlated for years y and $y+1$. The covariances were corrected for stock price by dividing by the mean of each stock for the year, thus yielding a dimensionless quantity. Each y in the range $1962, \ldots, 1991$ was considered a trial, and the cumulative distribution function of the thirty resulting Spearman coefficients is shown in Figure 3, curve *cov2–2*. Additional choices of six pairs were performed and yielded similar results, which are not shown.

Figure 3 also shows the results of two variance subexperiments. The six pairs cited in *cov2–2* were broken into two disjoint sets of six stocks, and Spearman coefficients calculated for rankings of variance divided by mean price squared at time y versus $y+1$, for $y = 1962, \ldots, 1991$. These thirty trials were cumulated to form curves *var–2* and *var–2bis*. Finally, the theoretical null hypothesis curve for $N = 6$ is given as curve *n–6*.

The calculations for this project were done in Perl on a DEC–5000/125 workstation under Ultrix 4.3. Further details of the programs and data sets are included in an extended Technical Report.

3 Conclusions

It appears that stock volatility is stable in time: a high variance stock yesterday will be a high variance stock tomorrow. However, the same is not true for the covariance of two stocks. A strong correlation of two stocks yesterday does not lead to a strong correlation of those stocks tomorrow. To test this idea, we applied nonparametric tests to the ranking of stocks from least to most variant and to the ranking of stock pairs from least to most covariant. For variance, there is this stability. However, for covariance, we cannot distinguish between actual stock data and a purely random shuffle at each period of covariance ranking.

This means that a portfolio adjusted correctly for the previous period, according to the methods of classical portfolio theory, should have no advantage over a neutral portfolio for the next period, since the facts upon which the adjustment is predicated are no more likely to stay put than is a pack of cards to remain unmodified after a thorough shuffling.

Also, this work underlines a subtlety in the theory of portfolio diversification. The stabilizing effect of diversification is not due to deterministic occurrences of negatively correlated industry cycles. Rather, negatively correlated stocks arise haphazardly, provided that the portfolio is large enough.

Furthermore, since covariance cannot be relied upon to retain its ranking, it cannot be bought and sold. This would lead one to believe, but one cannot conclude, that arbitrage equations involving covariance are unlikely to constrain the marketplace. On the other hand, the stability of variance that is confirmed in this paper concords with current use of variance, for example in the Black-Scholes option pricing formula, as a salable commodity and a source of arbitrage opportunities.

Acknowledgement: The author acknowledges the help of Prof. Thomas Gosnell of the Finance Department. Without his generous offer of assistance, the analysis of the CRSP data sets would not have been accomplished in time for these proceedings.

References

[1] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[2] Robert F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[3] John E. Freund. *Mathematical Statistics*. Prentice-Hall, Englewood Cliffs, New Jersey, 1992.

[4] Clive W. J. Granger and Oskar Morgenstern. *Predictability of Stock Market Prices*. Heath Lexington Books, Lexington, Massachusetts, 1970.

[5] Maurice G. Kendall. *Rank Correlation Methods*. Hafner Publishing Co., 1955.

[6] Harry M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. Cowles Foundation Monographs, Yale University Press, 1959.

[7] Robert J. Shiller. *Market Volatility*. The MIT Press, Cambridge, Massachusetts, 1989.

[8] Mark Torrance. Experimental stock market data. http://www.ai.mit.edu/stocks/.

Trial	Time Period	Experiment		
		Variance	Covariance	Random
1	93.3 vs. 93.4	3.42	0.21	1.44
2	93.3 vs. 93.4	4.32	−1.58	−0.55
3	93.3 vs. 93.4	3.63	−0.23	−0.90
4	93.4 vs. 94.1	4.44	0.28	−0.24
5	93.4 vs. 94.1	2.98	1.05	1.48
6	93.4 vs. 94.1	2.53	0.0058	−1.31
7	94.1 vs. 94.2	2.30	−0.76	0.59
8	94.1 vs. 94.2	1.58	1.41	−0.24
9	94.1 vs. 94.2	3.13	0.92	−1.56
10	94.2 vs. 94.3	3.83	2.73	0.58
11	94.2 vs. 94.3	3.74	1.75	0.20
12	94.2 vs. 94.3	2.88	−0.51	1.16
13	94.3 vs. 94.4	4.16	1.29	0.55
14	94.3 vs. 94.4	3.68	0.13	0.90
15	94.3 vs. 94.4	4.32	−0.86	0.58
16	94.1-2 vs. 94.3-4	3.94	0.096	0.060
17	94.1-2 vs. 94.3-4	1.12	−2.50	−0.26
18	94.1-2 vs. 94.3-4	2.03	−0.32	−1.17

Figure 1: Table of experiment one results.

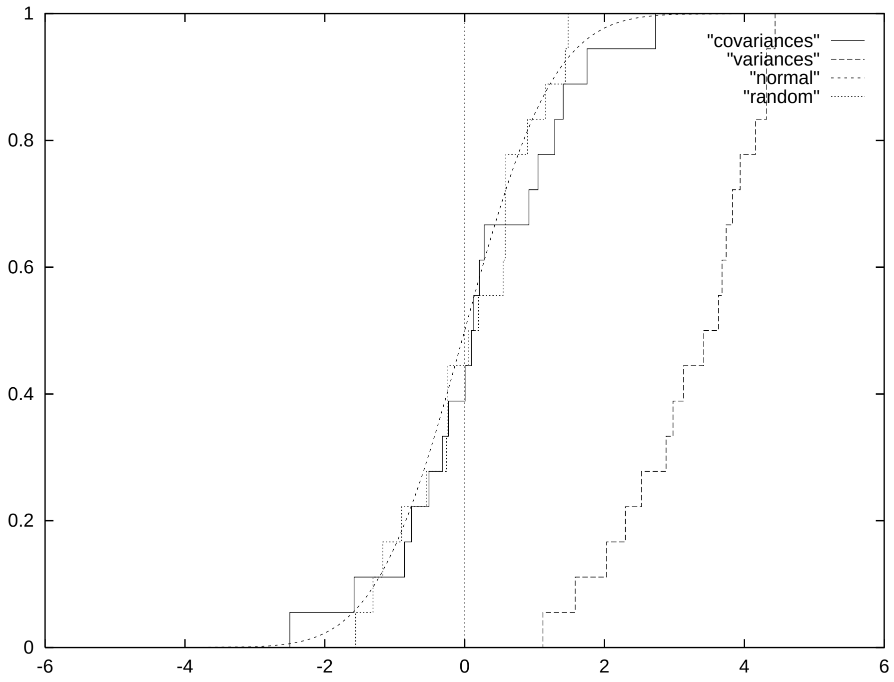

Figure 2: Summary of the first experiment.

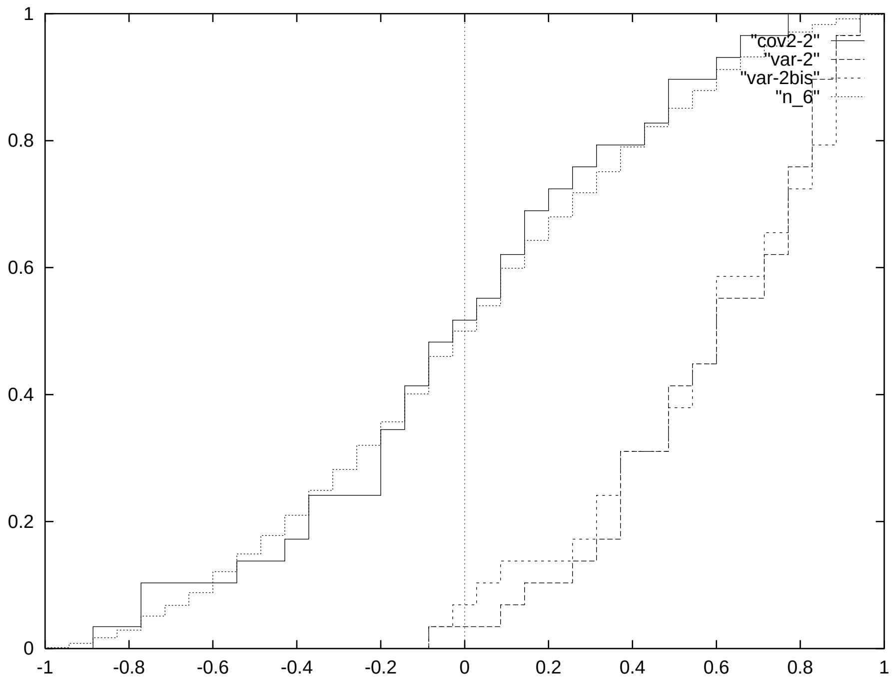

Figure 3: Summary of the second experiment.

Sym	Description	n/g	var-2	var-2bis	cov2-2
ADX	Adams Express Co, NYSE		•		•[1]
AL	Alcan Aluminum Ltd, NYSE		•		•[2]
BHY	Belding Hemingway Inc New , NYSE				
CAN	Continental Can Inc Del, NYSE		•		•[3]
DYA	Dynamics Corp of America, NYSE	*			
FP	Fischer & Porter, AMEX			•	•[1]
GQ	Grumman Corp., NYSE		•		•[4]
IP	International Paper Co, NYSE		•		•[5]
LDR	Landauer Inc, AMEX		•		•[6]
NMK	Niagara Mohawk Pwr Co, NYSE			•	•[4]
PKE	Park Electrochemical Corp, NYSE			•	•[6]
RGS	Rochester Gas & Elec. Corp, NYSE			•	•[3]
SCX	Starrett L. S. 'A', NYSE			•	•[2]
UIS	Unisys Corp, NYSE			•	•[5]

Figure 4: Table of stocks in experiment two.