Market Models

A Guide to
Financial Data Analysis

Carol Alexander

JOHN WILEY & SONS, LTD
Chichester • New York • Weinheim • Brisbane • Singapore • Toronto

Market Models

A Guide to
Financial Data Analysis

Carol Alexander

JOHN WILEY & SONS, LTD

Chichester • New York • Weinheim • Brisbane • Singapore • Toronto

Preface

This book is about the financial market models that are used by risk managers and investment analysts. It aims to provide a rigorous explanation of the theoretical ideas, but in practical and very clear terms. As concepts are introduced, real-world examples are provided in the text and, interactively, on the accompanying CD.

I have heard it said that too much academic research is focused on finding very precise answers to irrelevant questions. This book aims to provide academically acceptable answers to the questions that are really important for practitioners. It is written for a wide audience of practitioners, academics and students interested in the data analysis of financial asset prices.

It aims to help practitioners cut through the vast literature on financial market models, to focus on the most important and useful theoretical concepts. For academics the book highlights interesting research problems that are relevant to the day-to-day work of risk managers and investment analysts. For students, the comprehensive and self-contained nature of the text should appeal.

The book is divided into three parts:

Part I: Volatility and Correlation Analysis covers the estimation and forecasting of volatility and correlation for the pricing and hedging of options portfolios.

Part II: Modelling the Market Risk of Portfolios concerns factor modelling and the measurement of portfolio risk: the main focus is on modelling relationships between assets and/or risk factors using linear models.

Part III: Statistical Models for Financial Markets focuses on the time series analysis of financial markets.

A detailed summary of the content is provided in the introduction to each part. At the end of the book a low-level technical appendix is included; this covers the basic statistical theory that is necessary for the book to be self-contained.

Practitioners and academics share many important problems, and the communication between theory and practice is an essential part of model development. However, it is not always easy to straddle the divide between academic research and the practice of risk management and investment analysis. A common language, a common terminology and, above all, a common approach are necessary. It is hoped that this book will help to enhance the communication between these two schools.

Carol Alexander

July 2001

Contents

An asterisk '*' denotes that illustrative software is on the CD.

Part I
Volatility and Correlation Analysis

Part I provides insights into the pricing and hedging of options through the understanding of volatility and correlation, and the uncertainty which surrounds these key determinants of portfolio risk. The first chapter introduces volatility and correlation as parameters of the stochastic processes that are used to model variations in financial asset prices. They are not observable in the market and can only be measured in the context of a model.

Option pricing, which models asset prices in continuous time, is covered in Chapter 2. This chapter focuses on the consequences of using the Black–Scholes model to price options. Although there can only be one true volatility for the underlying price process, different volatilities are implied by the market prices of options on the same underlying asset. If one is willing to accept these volatilities, rather than invent better option pricing models, then their behaviour can be described by modelling the 'smile' or 'skew' patterns that emerge. The relationship between underlying price changes and changes in the implied volatility of an option is analysed to support the use of different volatility assumptions for pricing and hedging.

Statistical forecasts of volatility and correlation employ discrete time series models on historical return data. Chapter 3 explains how to obtain moving average estimates of volatility and correlation and outlines their advantages and limitations. A weighted average is a method for estimation. The current estimate of volatility or correlation is sometimes used as a forecast, but this requires returns to be independent and identically distributed, an assumption which is not always supported by empirical evidence. Chapter 4 introduces generalized autoregressive conditional heteroscedasticity (GARCH) models, which are based on more realistic assumptions about asset price dynamics. This chapter aims to cut through a vast academic literature on the subject to present the concepts and models that are most relevant to practitioners. A step-by-step guide to the implementation of the GARCH models that are commonly used by risk managers and investment analysts is followed by a description of the application of GARCH models to option pricing and hedging.

Most statistical models for forecasting volatility are actually models for forecasting variance: the volatility forecast is taken as the square root of the variance forecast. However, a forecast is an expectation, taken under some probability measure, and the expectation of a square root is not equal to the square root of an expectation. The last chapter in this part of the book examines this and other key issues surrounding the use of volatility and correlation forecasts. Quite different results can be obtained, depending on the model used and on the market conditions so, since volatility can only be measured in the context of a model, how does one assess the accuracy of a volatility forecast? Rather than employ point forecasts of volatility, this part of the book ends by advocating the use of standard errors, or other measures of uncertainty in volatility forecasts, to improve the valuation of options.

Part I introduces some challenging concepts that will be returned to later as further models are introduced. For example, the principal component models of implied volatility in §6.3, the orthogonal method for generating covariance matrices in §7.4 and the normal mixture density models of §10.3 will all continue the exposition of ideas that are introduced in Part I.

1

Understanding Volatility and Correlation

This chapter introduces some of the concepts that are fundamental to the analysis of volatility and correlation of financial assets. This is a vast subject that has been approached from two different technical perspectives. On the one hand, the option pricing school models the variation in asset prices in continuous time; this perspective will be taken in Chapter 2. On the other hand, the statistical forecasting school models volatility and correlation from the perspective of a discrete time series analyst; this is the approach used in Chapters 3 and 4.

The basic concepts are introduced within a unified framework that, I hope, will be accessible to both schools. Some of these concepts are quite complex and their exposition has necessitated many footnotes and numerous pointers to other parts of the book. First, volatility and correlation are described as parameters of stochastic processes that are used to model variations in financial asset prices. Then the differing needs of various market participants to assess volatility and correlation are examined. The needs of the analyst will

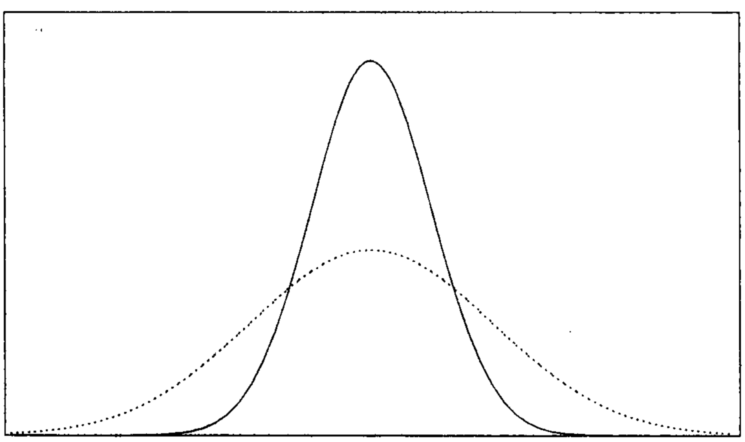

Figure 1.1 Volatility and scale.

determine whether an option pricing (implied volatility) approach or a statistical modelling (covariance matrix) approach is required (or both). Implied volatility and statistical volatility normally refer to the same process volatility, but volatility estimates often turn out to be quite different and because volatility can only be measured in the context of a model it is very difficult to assess the accuracy of estimates and forecasts. The chapter concludes with remarks on the decisions about the data and the models that will need to be made when volatility and correlation forecasts are implemented.

1.1 The Statistical Nature of Volatility and Correlation

Financial asset prices are observed in the present, and will have been observed in the past, but it is not possible to determine exactly what they will be in the future. Financial asset prices are random variables, not deterministic variables.[1] Variations of financial asset prices over a short holding period are often assumed to be lognormal random variables. Therefore returns to financial assets, the relative price changes, are usually measured by the difference in log prices, which will be normally distributed.[2]

Volatility is a measure of the dispersion in a probability density. The two density functions shown in Figure 1.1 have the same mean but the density function indicated by the dotted line has greater dispersion than the density indicated by the continuous line.[3] The most common measure of dispersion is the *standard deviation* σ of a random variable, that is, the square root of its variance.

---

[1] A *random variable*, also called a 'stochastic variable' or 'variate', is a real-valued function that is defined over a sample space with probability measure. A value x of a random variable X may be thought of as a number that is associated with a chance outcome. Each outcome is determined by a chance event, and so has a probability measure. This probability measure is represented by the *probability density function* of the random variable. For any probability density function $g(x)$, the corresponding *distribution function* is defined as $G(x) = \text{Prob}(X < x) = \int_{-\infty}^{x} g(x)\,dx$. It is not necessary to specify both density and distribution: given the density one can calculate the distribution, and conversely since $g(x) = G'(x)$.

[2] The *normal density function* $\phi(x)$ is defined by two parameters, the mean μ and the variance σ^2: $\phi(x) = ((2\pi\sigma^2)^{-1/2})\exp(\frac{1}{2}(x-\mu)^2/\sigma^2)$ for $-\infty < x < \infty$. This gives the familiar symmetric bell-shaped curve, which is centred on the mean μ and has a dispersion that is determined by the variance σ^2.

A random variable is said to be *lognormally distributed* when its logarithm is logonormally distributed. A lognormal density function is not symmetrical; it is bounded by zero on the low side but can, in theory, reach infinitely high values. For this reason it is commonly assumed that financial assets (bonds and shares) and possibly commodity prices are better represented by lognormal than by normal variates. Conversely, investors compare financial assets on the basis of their returns; it is therefore returns that are comparable whatever the price of the underlying asset, and it is simplest to assume that returns are normally distributed. It follows that the price is lognormally distributed: indeed if $r_t = (P_t - P_{t-1})/P_{t-1}$ is normally distributed then $P_t/P_{t-1} = 1 + r_t$ and $\ln(P_t/P_{t-1}) \approx r_t$ (note that when x is small, $\ln(1+x) \approx x$). Therefore $\ln(P_t/P_0)$ is normally distributed and P_t/P_0, is lognormally distributed. Note that this argument is based on investment assets and would not apply to interest rates. The argument has also shown that the return over small time intervals is approximated by the first difference in the log prices.

[3] If a random variable X has density function $f(x)$ then its *mean* is $\mu = E(X) = \int xf(x)dx$. The mean is like the centre of gravity of a density. It is a fundamental parameter of any density, the parameter that describes the *location* of the density. It is also called the *first moment* of the density function. The *variance* is $\sigma^2 = V(X) = \int(x-\mu)^2 f(x)dx = E(X^2) - [E(X)]^2$. This parameter measures the *dispersion* of the density function about the mean. It is also called the *second moment* about the mean of the density function.

It is hard to predict price variations of financial assets so it is usual to assume that successive returns are relatively independent of each other. This means that uncertainty will increase as the holding period increases, the distribution will become more dispersed and its variance will increase. Put another way, the variance of n-day returns will increase with n. Therefore it is not possible to compare n-day variance with m-day variance on the same scale. It is standard to assume statistically independent returns[4] and to express a standard deviation in annual terms. Thus in financial markets we define

Uncertainty will increase as the holding period increases, the distribution will become more dispersed and its variance will increase

$$\text{Annual volatility} = (100\sigma\sqrt{A})\%, \tag{1.2}$$

where A is an annualizing factor, the number of returns per year.[5] In this way volatilities of returns of different frequencies may be compared on the same scale in a volatility term structure (§2.2.2, §3.3 and §4.4.1).

To understand what correlation is, consider a *joint density* of two random variables.[6] A joint density may be visualized as a mountain: the more symmetric this mountain is about both the axes representing the two variables, the less information can be gained about the value of one variable by knowing the value of the other; that is, the lower the *correlation* between the two variables. For highly correlated variables the joint density will have more of a ridge in a direction between the axes of the two variables.

Figure 1.2 shows three 'scatter plots', where synchronous observations on each of the returns are plotted as horizontal and vertical coordinates. A scatter plot is a sample from the joint density of the two returns series, and so if the returns have no correlation their scatter plot will be symmetrically dispersed, like the one in Figure 1.2a; a high value on one axis will be no indication that the corresponding value on the other axis will be high or low. But if they have a high positive correlation the joint density will have a ridge sloping upwards, as in Figure 1.2b; when one variable has a high value the other will also tend to have a high value. If they have negative correlation the joint density will have a downwards sloping ridge as in Figure 1.2c; when one variable has a high value the other will tend to have a low value, and vice versa.

Correlation is a measure of co-movements between two returns series. Strong positive correlation indicates that upward movements in one returns series tend to be accompanied by upward movements in the other, and similarly

[4] Two random variables X and Y are *independent* if and only if their joint density function $h(x, y)$ is simply the product of the two marginal densities. That is, if X has density $f(x)$ and Y has density $g(y)$ then X and Y are independent if and only if $h(x, y) = f(x)g(y)$.

[5] The annualizing factor is a normalizing constant: the variance increases with the holding period but the annualizing factor decreases. The number of *trading days* (or 'risk days') per year is usually taken for the conversion of a daily standard deviation into an annualized percentage; that is, often A=250 or 252 in (1.2). Note the continuation of this footnote in §2.1.1 (footnote 5).

[6] The joint density $f(x, y)$ of two random variables X and Y is a real-valued function of the two variables where the total area underneath the surface is one: $\int\int f(x, y)dxdy = 1$. The joint probability that X takes values in one range and Y takes values in another range is the area under the function defined by these two ranges.

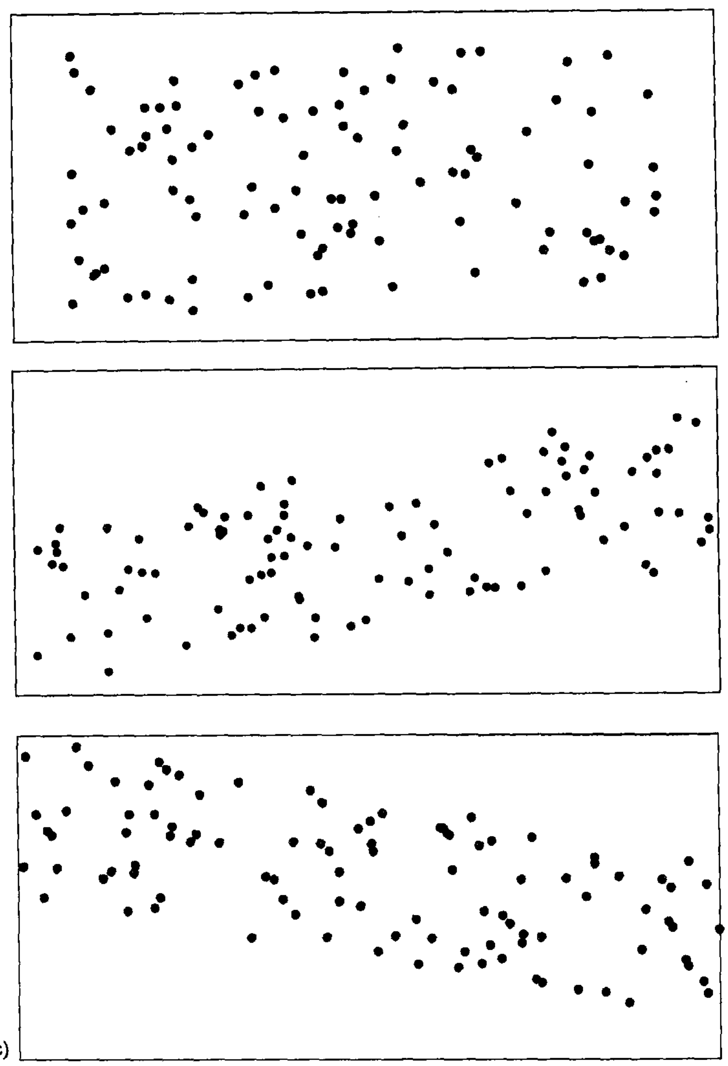
(a)
(b)
(c)

Figure 1.2 (a) Zero correlation; (b) positive correlation; (c) negative correlation.

downward movements of the two series tend to go together. If there is a strong negative correlation then upward movements in one series are associated with downward movements in the other.

A simple statistical measure of co-movements between two random variables is *covariance*, the first product moment about the mean of the joint density function. That is, $\text{cov}(X, Y) = E[(X-\mu_X)(Y-\mu_Y)]$, where $\mu_X = E(X)$ and $\mu_Y = E(Y)$. Covariance is determined not only by the degree of co-movement but also by the size of the returns. For example, monthly returns are of a much