Empirical Properties of Financial Data

MFM Practitioner Module: Quantitative Risk Management

John Dodson

September 14, 2016

Introduction

The analysis of financial timeseries naturally separates into the analysis of marginal univariate (scalar) random variables and dependent multivariate (vector) random variables. We will discuss empirical properties of both aspects.

Samples

Traditionally statisticians seek out i.i.d. samples. We will not be so lucky as to observe these directly with financial data. We will generally be able to retain the independence assumption if we use innovations such as log-returns for risk factors, but we will not be able to assume that observations through time are identically distributed.

- For example, the efficient market hypothesis says that changes in asset prices should be independent from period to period.

Independence

Before we define dependence, it is useful to define

Independence

Random variables X and Y are independent iff

$$F_{(X,Y)}(x,y) = F_X(x)F_Y(y) \qquad (*)$$

For all x, y. In particular,

$$\mathrm{E}\,(XY) = (\mathrm{E}\,X)\,(\mathrm{E}\,Y)$$

We can differentiate $(*)$ to see that

$$f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$$

It is also true of the characteristic functions $\phi_X(t) \triangleq \mathrm{E}\,e^{itX}$

$$\phi_{(X,Y)}(t_X,t_Y) = \phi_X(t_X)\phi_Y(t_Y)$$

Marginal Density

From Fubini's theorem, it is generally possible to derive marginal densities for a joint density, regardless of any dependence.

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dy$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dx$$

Of course, if X and Y are independent, then

$$f_{(X,Y)}(x,y) = f_X(x) f_Y(y)$$

but this does not need to be true in general.

Conditional Density

Conditioning a random variable is a powerful concept!

- The marginal characterization of a dependent variable is adequate if we do not know or care about the value of any potentially related dependent variables
- Conditioning, on the other hand, allows us to incorporate synthesis

Say we know the joint density of (X, Y), and we have learned that an event, say $Y = y$, is true. We can adjust the marginal distribution of X to account for this fact

$$f_{X|Y}(x) = f_X(x) \frac{f_{(X,Y)}(x, y)}{f_X(x) f_Y(y)}$$

- Note the analogy here to the Radon-Nikodým change of measure.

Conditional Expectation

A natural application of conditioning is the conditional expectation of a random variable.

$$\mathrm{E}\, X|Y = \int_{-\infty}^{\infty} x\, f_{X|Y}(x)\, dx$$

Tower Property

Sometimes it is useful to condition on unknown events. In this case, the conditional expectation is the same as the unconditional expectation.

$$\mathrm{E}\,(\mathrm{E}\, X|Y) = \mathrm{E}\, X$$

The lesson here is that conditioning has to exclude some outcomes in order to be consequential.

Dependence

To the extent that the *joint* density is not just a product of the *marginal* densities, there is dependence.

Factorization

This ratio can be expressed as

$$f_U\left(F_{X_1}(x_1), F_{X_2}(x_2), \ldots\right) \triangleq \frac{f_{(X_1, X_{2,\ldots})}(x_1, x_2, \ldots)}{f_{X_1}(x_1) f_{X_2}(x_2) \cdots}$$

Copula

Sklar's theorem says $f_U : [0,1]^N \mapsto \mathbb{R}^+$ is a density function that characterizes a new random variable, U, that encapsulates the dependence structure of X. Independence means $f_U \equiv 1$.

Two random variables that have the same copula are said to be co-monotonic.

Stylized Facts

Univariate

Daily timeseries of asset returns have certain general patterns that have been persistent enough to have become stylized facts:

- Returns are not i.i.d. but show little serial correlation
- Absolute returns show profound serial correlation
- Conditional expected returns are close to zero
- Conditional variance appears to vary over time
- Extreme return appear in clusters
- Returns appear to be fat-tailed or leptokurtotic

Modern econometric models are able to reflect all of these phenomena, and we will discuss this extensively in this module.

Stylized Facts
Multivariate

In the spirit of Sklar's theorem, ideally we would like to isolate common observations about multivariate financial timeseries into marginal and dependence phenomena.

- Only contemporaneous panel correlations are materially non-zero
- Absolute returns show profound panel and serial correlation
- Panel correlations vary over time
- Extreme returns tend to affect a number of components together

A focus on linear correlations complicates the analysis, because this measure of dependence is not strictly determined by the copula. Other dependence measures, such as Kendall's concordance, may be more useful.

Stylized Facts
Multivariate

In a multi-normal model, the conditional expectation of the dependent variable (Y) is affine in the independent variable (X).

$$\mathrm{E}\, Y|X = \mathrm{E}\, Y + \beta\,(X - \mathrm{E}\, X)$$

This relationship is particular to normal margins combined with a Gaussian copula.
More generally we might write

$$\mathrm{E}\, Y|X = \mathrm{E}\, Y + \beta(X)\,(X - \mathrm{E}\, X)$$

If $\beta(\cdot)$ is an increasing function, this might be interpreted as correlations increasing in extreme scenarios. In fact, it is possible that the copula parameters (Gaussian or otherwise) might be constant, but marginal leptokurtosis might be responsible.