Empirical Properties of Financial Data

MFM Practitioner Module: Quantitative Risk Management

John Dodson

October 30, 2019

Introduction

The analysis of financial timeseries naturally separates into the analysis of marginal univariate 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.

Random Variables

Definitions

A random variable is a (real) quantity whose value is not known with current information (but may be known in the future). These can be represented mathematically as measurable functions of a sample space, $X : \Omega \mapsto \mathbb{R}$. We try to denote them by upper-case Roman letters. We try to denote a placeholder for a particular value obtained by a random variable in its state space by the corresponding lower-case letter $x \in X(\Omega) \subset \mathbb{R}$. There is a probability associated with every measurable subset $\mathcal{A} \subset \mathbb{R}$, which might consist of a combination of intervals and points. The pre-image $\omega = X^{-1}(\mathcal{A}) \in \mathcal{F}$ of such sets are called events. The corresponding probability is the measure of the event $\mu(\omega) = \mathrm{P}\left[X \in \mathcal{A} | \mathcal{F}\right]$. $\mathcal{F}$ is the conditioning sigma algebra.

Random Variables

This structure lends itself to a measure theory interpretation, where the probability associated with a set is simply the integral of the probability density over that set.

$$\mathrm{P}\left[X \in (a, b)\right] = \int_a^b f_X(x)\, dx$$

If X ranges over the real numbers, $\mathbb{R}$, then $f_X(\cdot)$ must have certain properties. In particular, it must be a non-negative (generalized) function and

$$\lim_{x \to -\infty} f_X(x) = 0 \qquad \lim_{x \to \infty} f_X(x) = 0$$

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

Characterizations

In addition to the density function, there are at least three other characterizations of a real random variable

- density function $f_X(x)$
- distribution function $F_X(x) = \int_{-\infty}^{x} f_X(x')\, dx'$
- quantile function $Q_X(p) = F_X^{-1}(p)$
- characteristic function $\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)\, dx$

This last is based on the Fourier transform, where $i^2 = -1$. While the density function is the most common, these four representations are all equivalent.

Hint: they can be distinguished by the nature of their arguments; resp. a state, an upper bound on a state, a probability, and the reciprocal of a state.

Transformations

If we have a characterization of a random variable X, it is natural to ask if we can derive the characterization of a function of that variable $Y = h(X)$, or of a sample, $\{X_1, X_2, \ldots, X_n\}$, $Y = h(X_1, X_2, \ldots, X_n)$.
This is in general difficult, but there are some notable easy cases.

- f, F, and Q for an increasing, invertible function

$$f_Y(y) = \frac{f_X\left(h^{-1}(y)\right)}{h'\left(h^{-1}(y)\right)}$$

$$F_Y(y) = F_X\left(h^{-1}(y)\right) \qquad Q_Y(p) = h\left(Q_X(p)\right)$$

- ϕ for the mean of a sample, $Y_n = \frac{1}{n}\sum_{j=1}^{n} X_j$

$$\phi_{Y_n}(t) = \left[\phi_X\left(\frac{t}{n}\right)\right]^n$$

Expectation

We will talk in detail about the topic of estimation later, but if you were asked to give a "best guess" about the value of random variable, the expected value would be a natural answer.

The expected value is a density-weighted average of all possible values

$$
\begin{aligned}
\mathrm{E}\,X &= \int_{-\infty}^{\infty} x\, f_X(x)\, dx \\
&= \int_{-\infty}^{\infty} \left(\tfrac{1}{2} - F_X(x)\right)\, dx \qquad (!) \\
&= \int_0^1 Q_X(p)\, dp \\
&= -i\phi_X'(0)
\end{aligned}
$$

Ex. evaluate each for a Dirac spike, $f_X(x) = \delta(x - x_0)$

Expectation

While it may not be possible to evaluate the characterization of a function of a random variable, it is generally possible to evaluate the expected value of a function of a random variable.

$$\mathrm{E}\, h(X) = \int_{-\infty}^{\infty} h(x) f_X(x)\, dx$$

The probability measure of a measurable subet $\mathcal{A}$ is the expectation of the indicator function for the subset

$$\mathrm{P}\left[X \in \mathcal{A}\right] = \mathrm{E}\, 1_{\mathcal{A}}(X)$$

In general, the expected value of a function of a random variable is not just the value of the function at the expected value of the random variable.

$$\mathrm{E}\, h(X) \neq h(\mathrm{E}\, X)$$

Do not make this mistake!

Law of Large Numbers

A foundational result connects the expectation of a random variable with the sample mean.

Law of Large Numbers

$$\lim_{n\to\infty} \tfrac{1}{n} \sum_{j=1}^{n} X_j = \mathrm{E}\, X \quad \text{almost surely}$$

We will demonstrate the weak version of this result using the characteristic function and then look at some numerical results.

The key to the derivation is to recognize that

$$\phi_{Y_n}(t) = \left(1 + \frac{it\,\mathrm{E}\,X}{n} + o\left(\frac{t}{n}\right)\right)^n \to e^{it\,\mathrm{E}\,X}$$

Monte Carlo

The LLN, combined with information technology, brought about a revolution in applied mathematics last century, introducing a completely novel way to evaluate integrals.

- integrals are expectations of functions of random variables

$$\int h(x)\,dx = \int \frac{h(x)}{f_X(x)} f_X(x)\,dx = \mathrm{E}\,\frac{h(X)}{f_X(X)}$$

- expectations are means of random samples

$$\mathrm{E}\,\frac{h(X)}{f_X(X)} \overset{a.s.}{=} \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \frac{h(x_j)}{f_X(x_j)}$$

- you can evaluate an arbitrarily complicated integral if you can
    1. identify an appropriate random variable
    2. generate a very large sample of random variates
    3. cheaply evaluate the integrand and density function

Central Moments

Denote the expected value and standard deviation of a random variable X by μ and σ.
The standardized transformation of X is

$$Z = \frac{X-\mu}{\sigma}$$

Its characteristic function is

$$\phi_Z(t) = e^{-i\frac{\mu}{\sigma}t}\phi_X\left(\frac{t}{\sigma}\right)$$

The moments of Z measure the skewness, kurtosis, etc. of X. The easiest way to evaluate these is to note that

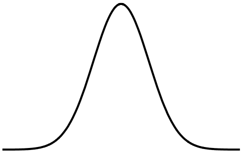
$$\operatorname{E} Z^n = (-i)^n \phi_Z^{(n)}(0)$$

N.B.: Moments do not always exist, and they are generally not adequate to characterize a random variable.

Normal Distribution

The most important distribution for $X \in \mathbb{R}$ is the normal or gaussian distribution. It takes two parameters, and is denoted $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$. μ is often called the mean.

$$
\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \frac{1}{\sigma} \\
F_X(x) &= \tfrac{1}{2} \operatorname{erfc}\left(\frac{\mu - x}{\sqrt{2}\sigma}\right) \\
Q_X(p) &= \mu - \sqrt{2}\sigma \operatorname{erfc}^{-1}(2p) \\
\phi_X(t) &= e^{i\mu t - \frac{1}{2}\sigma^2 t^2}
\end{aligned}
$$



Normal density

Central Limit Theorem

A foundational result connects the normal distribution with the Law of Large Numbers.

Central Limit Theorem

For any random variable X with finite expected value μ and standard deviation σ,

$$\lim_{n\to\infty} \sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{n} X_j - \mu \right) \sim \mathcal{N}\left(0, \sigma^2\right) \quad \text{in distribution}$$

That is, the deviation between the sample mean and the expected value is approximately normal, with standard deviation equal to the standard deviation of the random variable divided by the square root of the sample size.

This result can be demonstrated by considering the characteristic function of the RHS above.

Change of Measure

Recall, a random variable is a measurable function of the sample space with respect to a probability measure.
It can be useful to work with the same random variable under an alternate probability measure.

Radon-Nikodým theorem

If P and P′ are equivalent measures, then there is a random variable $Z \geq 0$ such that for any random variable X

$$\mathrm{E}'X = \mathrm{E}\left(ZX\right)$$

For example, say that $X \sim \mathrm{N}(\mu, \sigma^2)$ under P, but $X \sim \mathrm{N}(\mu', \sigma^2)$ under P′. Then the Radon-Nikodým derivative is

$$Z = e^{-\frac{X}{\sigma}\left(\frac{\mu}{\sigma}-\frac{\mu'}{\sigma}\right)+\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2-\frac{1}{2}\left(\frac{\mu'}{\sigma}\right)^2}$$

and this same Z would apply to any function of X.

Multivariate Random Variables

The support of a random variable is the union of all measurable sets in the state space. So far, we have discussed r.v.s whose support is $\mathbb{R}$. This naturally specializes to r.v.s whose support is an interval in $\mathbb{R}$ such as $(0, \infty)$ or $[0, 1]$. An important special case is when the support is countable, such as a just the naturals $\mathbb{N} = \{0, 1, 2, \ldots\}$, e.g. a Poisson r.v.

Random Vectors

When the support is $\mathbb{R} \otimes \cdots \otimes \mathbb{R} = \mathbb{R}^n$, we have random vectors. In this setting, rather then working in an abstract version of the preceding, it can be useful to consider this as n univariate random variables connected by a copula. We will introduce this below.

Another class of random variables that we will encounter are random scalars. The state space for these are positive-definite tensors of some particular dimension and rank, e.g. a Gamma or Wishart r.v.

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 for Financial Risk Factors
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 for Financial Risk Factors
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.

- 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 for Financial Risk Factors

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.