Financial Notation Project

Current Release: July, 2006

Introduction

The **Financial Notation Project** is an initiative to standardize the mathematical notation employed in technical financial publications.

Since the early 1970s, finance has become increasingly mathematical, employing essentially every branch of applied math—calculus, linear algebra, differential equations, probability, statistics, stochastic calculus, optimization. Each offers its own notation conventions. Because these conflict, it is impossible to observe them all simultaneously. The solution is to develop a finance-specific system of notation, but financial developments have come so rapidly since the 1970s, that the development of financial notation has been largely ad hoc.

The **Financial Notation Project** offers a system of notation that consistently presents financial concepts. It draws on existing conventions as much as possible. It also introduces some innovations that are both intuitive and simplify the presentation of concepts.

The initiative evolved out of Glyn Holton's own efforts to publish financial content with a consistent system of notation. He put considerable effort into devising notation and decided to make that work available to the financial community. Users who have questions, comments or feedback on the notation are encouraged to write Glyn at **glyn@contingencyanalysis.com**. The notation system has been implemented in Mr. Holton's book *Value-at-Risk: Theory and Practice*, the **www.riskglossary.com** website and other publications.

This document briefly introduces the notation system. For examples of how the notation is used in specific branches of finance, see relevant articles in the **www.riskglossary.com** website. For example, if you want to see how the notation is used in the context of value-at-risk, see the glossary's article on value-at-risk and related articles.

Random Quantities

There is an important difference between a random quantity and a constant quantity. Suppose you read in a paper "Let s be a stock price." You don't immediately know whether the author means a known stock price that has already been observed or a random stock price whose value has not yet been determined. Existing systems of mathematical or financial notation offer no clue.

To remedy this, random quantities are indicated with capital English letters. If they are random variables—that is, univariate—they are italic nonbold: Q, R, S, X, etc. If they are multivariate in some sense—random vectors, random matrices, stochastic processes—they are italic bold: $\boldsymbol{Q}$, $\boldsymbol{R}$, $\boldsymbol{S}$, $\boldsymbol{X}$, etc. Nonrandom quantities are indicated with lowercase italic letters. These are nonbold for scalars: q, r, s, x, etc. They are bold for vectors, matrices, or time series: $\boldsymbol{q}$, $\boldsymbol{r}$, $\boldsymbol{s}$, $\boldsymbol{x}$, etc. The only exception to this is the (nonrandom) identity matrix. Bowing to convention, it is denoted $\boldsymbol{I}$.

With this notation, if a random variable is denoted X, a specific realization of that random variable may be denoted x. Such notational correspondence between random quantities and realizations of those random quantities can dramatically contribute to readers' comprehension.

Subscripts

Components of vectors or matrices are distinguished with subscripts. Consider the random vector

$$\boldsymbol{Q} = \begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \end{pmatrix}, \qquad [1]$$

or the (nonrandom) matrix

$$\boldsymbol{c} = \begin{pmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{pmatrix}. \qquad [2]$$

Time series or stochastic processes arise frequently in finance. To avoid confusion, their time indices are not indicated with subscripts. Instead, they are indicated with superscripts that precede the rest of a symbol. For example, a Libor curve evolves over time. We may represent its value at time t as

$$^{t}\boldsymbol{R} = \begin{pmatrix} ^{t}R_1 \\ ^{t}R_2 \\ \vdots \\ ^{t}R_{12} \end{pmatrix} \sim \begin{pmatrix} \text{1 - month Libor} \\ \text{2 - month Libor} \\ \vdots \\ \text{12 - month Libor} \end{pmatrix}. \qquad [3]$$

The value at time 3 of 2-month Libor is $^{3}R_2$. The entire Libor curve at time 1 is denoted $^{1}\boldsymbol{R}$. The univariate stochastic process representing 2-month Libor over time is represented $\boldsymbol{R}_2$. The 12-dimensional stochastic process representing the entire curve over time is denoted $\boldsymbol{R}$.

Time 0 is generally considered the current time. At time 0, current and past Libor curves are known. As nonrandom quantities, they are represented with lowercase letters: …, $^{-3}\boldsymbol{r}$, $^{-2}\boldsymbol{r}$, $^{-1}\boldsymbol{r}$, $^{0}\boldsymbol{r}$. If time is measured in days, yesterday's value of 9-month Libor is denoted $^{-1}r_9$.

The advantage of using preceding superscripts to denote a stochastic process' time index is clarity. In the above example, suppose time and maturity were both indicated with subscripts and you saw the notation $R_{3,2}$. It would leave you wondering "Is this 3-month Libor at time 2 or 2-month Libor at time 3?" By keeping time and component indices physically separate, the notation ensures one will never be confused for the other. Use of preceding superscripts is unconventional but not without precedent. Actuarial notation makes extensive use of preceding superscripts.

Conditional Distributions

Perhaps no branch of mathematics has worse notation conventions than time series analysis. Not only do authors routinely fail to distinguish between random terms of a stochastic process and realizations of those random terms. They also fail to distinguish between unconditional and conditional distributions. None of this bothers knowledgeable researchers, but it is extremely confusing for beginners. I have already described how capital or lowercase letters are used to distinguish random from nonrandom quantities. Now I will describe how unconditional distributions are distinguished from conditional distributions.

Let $\boldsymbol{X}$ be a univariate stochastic process, so ${}^{t}X$ is the random value of that stochastic process at time t. Suppose the unconditional standard deviation of ${}^{t}X$ is constant (i.e. has the same value for all times t). Then we would denote that unconditional standard deviation σ. If the unconditional standard deviation was not constant (i.e. changes with time) we would denote it ${}^{t}\sigma$. The standard deviation of ${}^{t}X$ conditional on information available at time $t - 1$ would be denoted ${}^{t\,|\,t-1}\sigma$. This convention generalizes in the obvious ways to other parameters, such as the mean μ, or to times other than t or $t - 1$.

In probability, there is standard notation for indicating common probability distributions, including

- $U(a,b)$ for a uniform distribution with nonzero probability on the interval $[a, b]$.
- $N(\mu, \sigma^2)$ for a normal distribution with mean μ and variance σ^2.
- $\Lambda(\mu,\sigma^2)$ for a lognormal distribution with mean μ and variance σ^2.
- $\chi^2(\nu,\delta^2)$ for a chi-squared distribution with ν degrees of freedom and noncentrality parameter δ^2.

These are used unmodified to indicate unconditional distributions. For example, the notation ${}^{t}X \sim N(0, 4)$ indicates that ${}^{t}X$ has an unconditional distribution that is normal with mean 0 and variance 4. Conditional distributions are indicated by adding a preceding superscript. For example, the notation ${}^{t}X \sim {}^{t-5}N(2, 9)$ indicates that, conditional on information available at time $t - 5$, ${}^{t}X$ has a normal distribution with mean 2 and variance 9.