Financial Theory – Lecture 3

Fredrik Armerin, Uppsala university, 2024

Agenda

- Brief recap of probability theory.
- Measuring risk.
- Vectors, matrices and multivariate random variables.

The lecture is based on

- Sections 3.1-3.5 and 4.2 in the course book.

Risk in financial economics

From the course book:

> *Two key themes of this book are exactly how to measure the risk of an investment and by how much such risks are compensated in financial markets.*

(Munk, p. 16.)

Random variables

There are discrete and continuous random variables.

For a discrete random variable we define the probability function

$$p_s = P(X = x_s)$$

of getting the outcome x_s, and for a continuous random variable we define the probability density function (pdf) $f_X(x)$ which has the property that for $a < b$

$$P(a < X \leq b) = \int_a^b f_X(x)dx.$$

Remark. There are random variables that has both a discrete and a continuous part. They are called mixed random variables.

Random variables

For any random variable we define its cumulative probability function (cdf) as

$$F_X(x) = P(X \leq x).$$

It holds that

$$P(a < X \leq b) = F_X(b) - F_X(b).$$

The $k\%$ percentile for a contionuous random variable is the value x that satisfies

$$P(X \leq x) = k\% \quad \Leftrightarrow \quad F_X(x) = k\% \quad \Leftrightarrow \quad x = F_X^{-1}(k\%).$$

Expected values

The expected value of a discrete random variable is defined as

$$E[X] = \sum_s p_s x_s,$$

and for a continuous random variable as

$$E[X] = \int_{-\infty}^{\infty} x\, f_X(x) dx.$$

The expected value of a function of a random variable are given by

$$E[g(X)] = \sum_s p_s g(x_s)$$

and

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f_X(x) dx$$

for a discrete and continuous random variable respectively.

Variance and standard deviation

To measure the variability in the outcome of a random variable we use the variance

$$\mathrm{Var}[X] = E\left[\left(X - E\left[X\right]\right)^2\right]$$

or the standard deviation

$$\mathrm{Std}[X] = \sqrt{\mathrm{Var}[X]}.$$

The variance satisfies

$$\mathrm{Var}[X] = E\left[X^2\right] - \left(E\left[X\right]\right)^2.$$

Calculation rules

Let X be a random variable, and let a and b be two real numbers. Then

$$\begin{aligned} E[aX+b] &= aE[X]+b \\ \mathrm{Var}[aX+b] &= a^2\mathrm{Var}[X] \\ \mathrm{Std}[aX+b] &= |a|\mathrm{Std}[X]. \end{aligned}$$

Higher order moments

Skew or skewness:

$$\mathrm{Skew}[X] = \frac{E\left[\left(X - E\left[X\right]\right)^3\right]}{\mathrm{Std}[X]^3}.$$

Kurtosis:

$$\mathrm{Kurt}[X] = \frac{E\left[\left(X - E\left[X\right]\right)^4\right]}{\mathrm{Std}[X]^4} - 3.$$

If $\mathrm{Kurt}[X] > 0$, then we say that the distibution of X is lepokurtic, or that it has heavy tails or fat tails.

Higher order moments

If X is normally distributed with mean μ and variance σ^2, which we write as

$$X \sim N(\mu, \sigma^2),$$

then

$$\text{Skew}[X] = 0 \quad \text{and} \quad \text{Kurt}[X] = 0.$$

The risk-return tradeoff

Let r be a random rate of return over some time interval.

The expected value $E[r]$ is a measure of the reward we get from the investment.

Since we also can put our money in the bank and get the risk-free rate of return r_f, it is common to look at the excess return

$$E[r] - r_f.$$

This difference is called the risk premium.

To measure the risk in an investment, the standard deviation of the rate of return $\mathrm{Std}[r]$ is often used.

The risk-return tradeoff

In order to measure the tradeoff between the risk and the reward in terms of the expected rate of return, the Sharpe ratio is often used:

$$\mathrm{SR} = \frac{E[r] - r_f}{\mathrm{Std}[r]}.$$

We will later in the course see that the Sharpe ratio arises in a natural way in finance, but at this point it is just one suggestion of how to measure the risk-return tradeoff.

In practical asset management other measures are also used, such as the Sortino ratio or the Maximum drawdown.

Normally distributed log-returns

Assume that we want to model the rate of return r as a normally distributed random variable.

Since r can take both positive and negative values, this seems as an OK model.

One drawback, however, is that the rate of return cannot be lower than -100%, i.e. $r \geq -1$.

The solution to this potential problem is to instead assume that the log-return is normally distributed.

Normally distributed log-returns

A random variable X is said to be lognormally distributed with parameters m and s^2 if

$$\ln X \sim N(m, s^2).$$

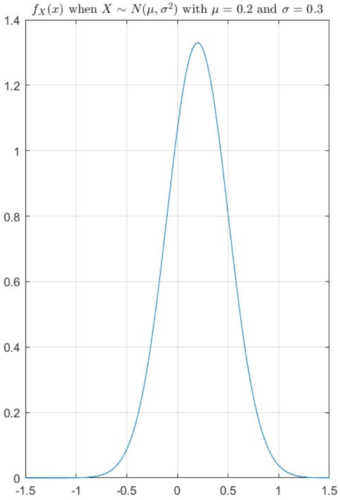
The first two moments of a lognormally distributed random variable X are

$$\begin{aligned} E[X] &= e^{m+\frac{s^2}{2}} \\ \mathrm{Var}[X] &= e^{2m+s^2}\left(e^{s^2}-1\right). \end{aligned}$$

Since $r^{\log} = \ln R$ we see that

$$r^{\log} \text{ is normally distributed } \Leftrightarrow R \text{ is lognormally distributed.}$$

Normally distributed log-returns



Covariance

The covariance between two random variables X_1 and X_2 is defined as

$$\text{Cov}[X_1, X_2] = E\left[\left(X_1 - E\left[X_1\right]\right) \cdot \left(X_2 - E\left[X_2\right]\right)\right].$$

We see that

$$\text{Cov}[X_1, X_1] = \text{Var}[X_1].$$

It can further be shown that

$$\text{Cov}[X_1, X_2] = E\left[X_1 X_2\right] - E\left[X_1\right] E\left[X_2\right],$$

which can be written

$$E\left[X_1 X_2\right] = E\left[X_1\right] E\left[X_2\right] + \text{Cov}[X_1, X_2].$$

For random variables X, Y and real numbers a, b, c, d

$$\text{Cov}[aX + b, cY + d] = ac\text{Cov}[X, Y].$$

Correlation

The correlation between two random variables X_1 and X_2 is defined as

$$\operatorname{Corr}[X_1, X_2] = \frac{\operatorname{Cov}[X_1, X_2]}{\operatorname{Std}[X_1]\operatorname{Std}[X_2]}.$$

The correlation satisfies

$$-1 \leq \operatorname{Corr}[X_1, X_2] \leq 1 \quad \Leftrightarrow \quad \big|\operatorname{Corr}[X_1, X_2]\big| \leq 1.$$

Note that

$$\operatorname{Cov}[X_1, X_2] = \operatorname{Corr}[X_1, X_2]\operatorname{Std}[X_1]\operatorname{Std}[X_2].$$

Alternative ways of measuring risk

Using the variance (standard deviation) to measure the risk goes back to Markowitz 1952.

There have been other suggestions.

- Mean absolute deviation (MAD): $E\left[\left|r - E[r]\right|\right]$.
- Semivariance.
- Value-at-Risk (VaR).
- Expected shortfall (ES) (sometimes called Conditional Value-at-Risk (CVaR) or Tail Value-at-Risk (TVaR)).
- Coherent risk measures.
- Convex risk measures.

Value-at-Risk and Expected shortfall

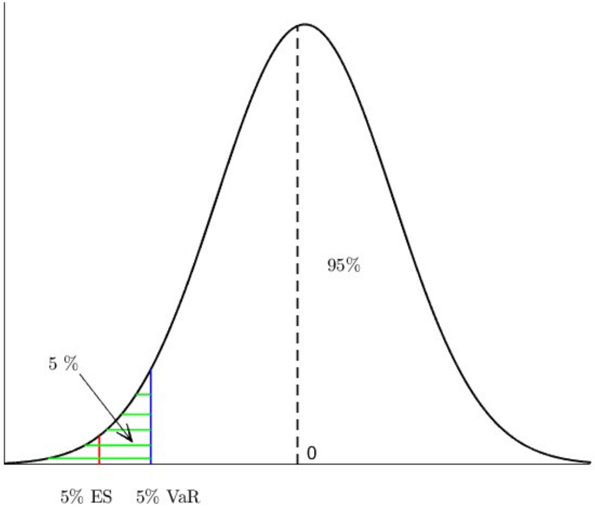
Say that we want to measure the risk of what can happen on the 5% worst days of trading for an investor.

Value-at-Risk
On 95% of the days, the maximum loss is equal to VaR.

Expected shortfall
On the 5% worst days, the mean loss is equal to ES.

VaR is about what can happen on a "good" day, while ES is about what can happen on a "bad" day.

Value-at-Risk and Expected shortfall



Value-at-Risk and Expected shortfall

From a computational point of view, the Value-at-Risk is a percentile of the profit distribution.

The profit can be defined either in units of currency or as the rate of return from the investment. See the book for details.

In some cases VaR and ES are defined in terms of the loss distribution. The only difference is that the sign of VaR and ES are changed.

Risk measures

The systematic study of different types of risk measures starts with "Thinking coherently" from 1997 and "Coherent measures of risk" from 2001, both by Artzner, Delbaen, Embrechts and Heath.

- Value-at-Risk is not a coherent risk measure.
- Expected shortfall is a coherent risk measure.

Some argue that the definition of a coherent risk measure is too restrictive. This has lead to the more general concept of convex risk measures (and other alternatives as well).

The Basel accords

The Basel accords regulates the supervision of banks from a risk perspective.

From Basel Committee on Banking Supervision, "MAR33 Internal models approach: capital requirements calculation":

33.2 ES must be computed on a daily basis for the bank-wide internal models to determine market risk capital requirements. ES must also be computed on a daily basis for each trading desk that uses the internal models approach (IMA).

33.3 In calculating ES, a bank must use a 97.5th percentile, one-tailed confidence level.

(This is equivalent to what the book calls 2.5% ES.)

Previously 1% VaR was used.

Vectors and matrices

A vector is an orded collection of N elements:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = (x_1, x_2, \ldots, x_N)^\top.$$

Here $\top$ denotes the transpose of a vector.

Remark. The notation $'$ is also used to denote transposition.

Vectors and matrices

Recall that

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_N + y_N \end{pmatrix}$$

and for a scalar a

$$a\mathbf{x} = a \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_N \end{pmatrix}.$$

Vectors and matrices

The inner product (or vector product or dot product) between two vectors is given by

$$\mathbf{x} \cdot \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_N y_N = \sum_{i=1}^{N} x_i y_i.$$

Note that

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}.$$

Vectors and matrices

We let

$$\mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

denote the vector of only 1's. For any vector $\mathbf{x}$

$$\mathbf{x} \cdot \mathbf{1} = x_1 + x_2 + \cdots + x_N = \sum_{i=1}^{N} x_i.$$

Vectors and matrices

A matrix A is a collection of elements in a rectangular array:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{pmatrix}.$$

This vector has M rows, N columns and a total of MN number of elements.

We say that A is an $M \times N$ matrix.

In the book matrices are denoted using a double underline: $\underline{\underline{A}}$.

Vectors and matrices

Let A be an $N \times N$ matrix, and let $\mathbf{x}$ and $\mathbf{y}$ be two column vectors of length N. Then

$$A\mathbf{x} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} \square \\ \square \\ \vdots \\ \square \end{pmatrix}$$

is a column vector of length N.

We can now multiply this vector with $\mathbf{y}$:

$$\mathbf{y} \cdot A\mathbf{x} = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i y_j A_{ij},$$

and the result is a scalar.

A square matrix A is symmetric if

$$A = A^\top,$$

which is the same as requiring that

$$A_{ij} = A_{ji}, \text{ for every } i, j = 1, 2, \ldots, N \text{ when } i \neq j.$$

An import observation is that if A is symmetric, then

$$\mathbf{y} \cdot A\mathbf{x} = \mathbf{x} \cdot A\mathbf{y}.$$

The main example of a symmetric matrix in this course is the variance-covariance matrix.

Vectors and matrices

We let I denote the identity matrix:

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

i.e. it is a square matrix with 1's on the diagonal an 0's off the diagonal. This matrix has the property that

$$A\,I = I\,A = A$$

for any matrix A such that the multiplication makes sense (i.e. if A is $N \times N$, then I must be $N \times N$).

In the book the notation $\underline{\underline{1}}$ is used for the identity matrix.

The inverse of a square matrix A (if it exist!) is a matrix denoted A^{-1} which satisfies

$$AA^{-1} = A^{-1}A = I.$$

Example If the matrix A is 1×1 then it is equal to a scalar a. If $a \neq 0$, then

$$a \cdot \frac{1}{a} = \frac{1}{a} \cdot a = 1 \leftarrow \text{The } 1 \times 1 \text{ identity matrix.}$$

As for numbers, not every matrix is invertible.

Vectors and matrices

Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be a 2×2 matrix.

Then it is invertible if and only if $ad - bc \neq 0$ and in this case

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Exercise: Check this!

Multivariate random variables

Let $X_1, X_2, \ldots, X_N$ be random variables and let $a_1, a_2, \ldots, a_N$ be real numbers. Then

$$E\left[\sum_{i=1}^{N} a_i X_i\right] = \sum_{i=1}^{N} a_i E\left[X_i\right]$$

$$\mathrm{Var}\left[\sum_{i=1}^{N} a_i X_i\right] = \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \mathrm{Cov}[X_i, X_j]$$

The last formula is sometimes written

$$\mathrm{Var}\left[\sum_{i=1}^{N} a_i X_i\right] = \sum_{i=1}^{N} a_i^2 \mathrm{Var}[X_i] + \sum_{i \neq j}^{N} a_i a_j \mathrm{Cov}[X_i, X_j]$$

(and see Theorem 3.2 in Munk for a third way).

Multivariate random variables

An N-dimensional random variable $\boldsymbol{X}$ is a random vector:

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix} = (X_1, X_2, \ldots, X_N)^\top.$$

Multivariate random variables

The mean vector is given by

$$\boldsymbol{\mu} = E[\boldsymbol{X}] = \begin{pmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_N] \end{pmatrix}.$$

Using this we can write

$$E\left[\sum_{i=1}^{N} a_i X_i\right] = \sum_{i=1}^{N} a_i E[X_i] = \mathbf{a} \cdot \boldsymbol{\mu}.$$

Multivariate random variables

The variance-covariance (or covariance) matrix is given by

$$\Sigma = \mathrm{Var}(\boldsymbol{X}) = \begin{pmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots & \mathrm{Cov}[X_1, X_N] \\ \mathrm{Cov}[X_2, X_1] & \mathrm{Var}[X_2] & \cdots & \mathrm{Cov}[X_2, X_N] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}[X_N, X_1] & \mathrm{Cov}[X_N, X_2] & \cdots & \mathrm{Var}[X_N] \end{pmatrix},$$

that is

$$\Sigma_{ij} = \mathrm{Cov}[X_i, X_j], \ \ i, j = 1, \ldots, N.$$

Since $\mathrm{Cov}[X_i, X_j] = \mathrm{Cov}[X_j, X_i]$, the matrix Σ is symmetric:

$$\Sigma = \Sigma^\top.$$

Futhermore

$$\mathrm{Var}\left[\sum_{i=1}^N a_i X_i\right] = \sum_{i=1}^N \sum_{j=1}^N a_i a_j \mathrm{Cov}[X_i, X_j] = \mathbf{a} \cdot \Sigma \mathbf{a}.$$