Cap 5. Tools in probability theory

5.1 Probability

Building a formal set of tools based on probability theory is fundamental in quantitative finance: although investors may be driven by a personal intuition of probability, we have previously seen in the summary about no-arbitrage that the fair market value of financial assets can be constructed using probability measures constructed "artificially", arriving to risk-neutral valuation.

A particular **state of the world** is denoted by the symbol ω, and **all possible state of the world** are written as Ω. An **event** is a set of elementary ω, and the set of all possible events is represented by I. To each event $A \epsilon I$, we assign the probability $P(A)$. The probabilities must be defined such that:

$P(A) \geq 0 \; \forall \; A \in I$ and $\int_{A \in I} dP(A) = 1$. The triplet $\{\Omega, I, P\}$ is called a **probability space**.

5.2 Random variables and functions

A random variable X is a function defined on the set I: given an event $A \epsilon I$ a random variable will assume a particular numerical value. We have $X: I => B$, where B is the set made of all possible subsets of the real numbers R.

We can now introduce the **distribution function** $G(x) = P(X \leq x)$. $G(\cdot)$ is a function of x, which represents a certain threshold whereas X represents the random variable. If the distribution function is smooth and has a derivative we can define the **density function** of X, denoted by $g(x)$. $g(x) = \frac{dG(x)}{dx}$.

5.3 Moments

The **first moment is the expected value $E[X]$** and it is defined as $E[X] = \int_{-\infty}^{\infty} x f(x) dx$ where $f(x)$ is the corresponding density function. **The second moment is the variance $E[X - E[X]]^2$**, defined in the integral form as $\int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$. The first moment gives information about the center of the distribution, whereas the second one on how much the distribution is spread around this center. The square root of the variance is the standard deviation, which is a measure of the average deviation of an observation from the mean. In the financial markets this measure is called **volatility**. In the case of a normal distributed variable each higher order moments can be linked to the first two: the mean and the variance are sufficient to fully describe a normal variable.

The third moment, the skewness, capture asymmetry in the distribution, whereas the fourth one, kurtosis, captures how "high" the tails are. Higher tails than the normal distribution are usually called "fat tails" and are frequent with financial data. This means that an extreme event is more likely to happen compared to according to a normal distribution.

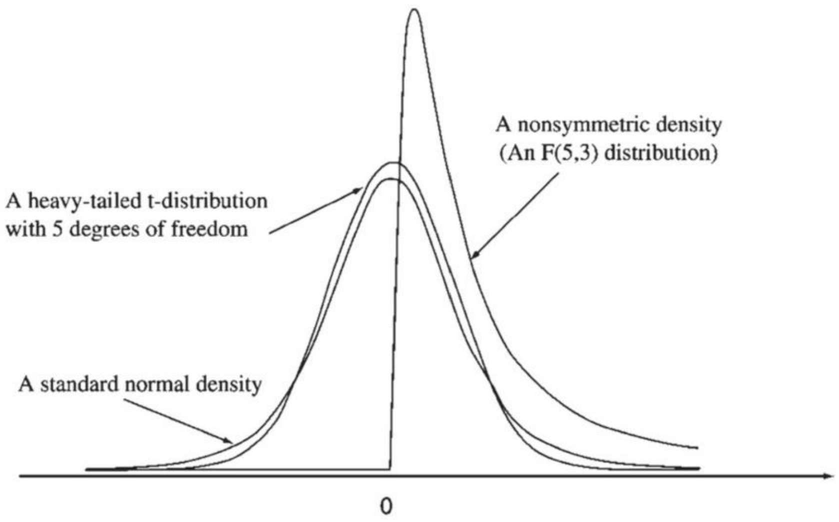

FIGURE 5.2 Examples of symmetric and nonsymmetric distributions.

Heavier tails manifest themselves after a certain threshold in the x-axis.

5.4 Conditional expectations

One forecasting a random variable we some information, denoted by the symbol I_t. Expectations calculated using information are called **conditional expectations** and the conditional mathematical operator is the **conditional expectation operator**. As the information changes according to the time we are in, the conditional expectation operator is indexed by the time index. If we assume that the decision maker never forgets past information, the information sets must become be increasing over time.

We now must define a conditional probability. In particular we define the conditional density of a random variable as $f(x|I_t)$.

The next step is the averaging operator: every forecast is an average of possible values: each values is weighted by its probability. The conditional expectation (forecast) of some random variable S_t can therefore be defined as $\int_{-\infty}^{\infty} S_t f(S_t|I_u) dS_t \quad u < t$.

The conditional expectation operator E_t (a more compact version of $E[x|I_t]$ has the following properties:

1). $E_t[S_t + F_t] = E_t[S_t] + E_t[F_t]$

2). $E_t\left[E_{t+T}[S_{t+T+u}]\right] = E_t[S_{t+T+u}]$

5.5. Markov processes and their relevance

The discussion so far has dealt with the concept of random variables. However what we really need is a model for a *sequence* of random variables. A sequence of random variables $\{X_t\}$ is called stochastic process where t can either be discrete or continuous. A stochastic process should have a well defined joint distribution function $F(x_1, \dots, x_t) = P(X_1 \le x_1, \dots, X_t \le x_t)$.

We will now discuss an important class of stochastic process for finance, the Markov process.

Definition 13. A discrete time process, $\{X_1, \ldots, X_t, \ldots\}$, with joint probability distribution function, $F(x_1, \ldots, x_t)$, is said to be a Markov process if the implied conditional probabilities satisfy

$$P(X_{t+s} \leq x_{t+s} | x_t, \ldots, x_1) = P(X_{t+s} \leq x_{t+s} | x_t) \tag{5.44}$$

where $0 < s$ and $P(\cdot | I_t)$ is the probability conditional on the information set I_t.

In practical terms we are saying that the knowledge of the past is irrelevant for any statement concerning X_{t+s} given x_t. This is extremely relevant for finance application. Suppose X_t represent instantaneous spot rate. Assuming that r_t is Markov we have that the expected future value of r_{t+s} depends only on the latest observation. Splitting changes in interest rates into expected and unexpected components:

$$r_{t+\Delta} - r_t = E[(r_{t+\Delta} - r_t | I_t] + \sigma(I_{t,}t)\Delta W_t$$

If r_t is a Markov process we can write

$$\mathbb{E}[(r_{t+\Delta} - r_t) | I_t] = \mu(r_t, t)\Delta \tag{5.46}$$

and

$$\sigma(I_t, t) = \sigma(r_t, t) \tag{5.47}$$

And as $\Delta ==> 0$

$$dr_t = \mu(r_t, t)\,dt + \sigma(r_t, t)\,dW_t \tag{5.48}$$

But if interest rates were not Markov these steps could not have been followed since the mean and variance could have depended on more observations in the past.

In case of modelling two stochastic processes one has to keep in mind that although two processes can be *jointly* when we model *one* in a univariate setting it will, in general, not be a Markov process. On the other any univariate Markov-process can be converted into a Markov process by increasing the dimensionality of the problems.