P1: JYS
c01 JWBK307-Huynh September 15, 2008 7:42 Printer: Yet to come

1 Introduction to Probability

Since financial markets are very volatile, in order to model financial variables we need to characterize randomness. Therefore, to study financial phenomena, we have to use probabilities.

Once defined, we will see how to use probabilities to describe the evolution of random parameters that we later call random processes. The key step here is the quantitative construction of the events' probabilities. First, one must define the events and then the probabilities associated to these events. This is the objective of this first chapter.

1.1 INTUITIVE EXPLANATION

1.1.1 Frequencies

Here is an example to illustrate the notion of relative frequency. We toss a dice N times and observe the outcomes. We suppose that the 6 faces are identified by letters A, B, C, D, E and F. We are interested in the probability of obtaining face A. For that purpose, we count the number of times that face A appears and denote it by $n(A)$. This number represents the frequency of appearance of face A.

Intuitively, we see that the division of the number of times that face A appears, $n(A)$, by the total number N of throws, $\frac{n(A)}{N}$, is a fraction that represents the probability of obtaining face A each time that we toss the dice. In the first series of experiments when we toss the dice N times we get $n_1(A)$ and if we repeat this series of experiments another time by tossing it again N times, we obtain $n_2(A)$ of outcomes A.

It is likely that $n_1(A)$ and $n_2(A)$ are different. The fractions $\frac{n_1(A)}{N}$ and $\frac{n_2(A)}{N}$ are then different. Therefore, how can we say that this fraction quantifies the probability of obtaining face A? To find an answer, we need to continue the experiment. Even if the fractions are different, when the number N of throws becomes very large, we observe that these two fractions converge to the same value of $\frac{1}{6}$.

Intuitively, this fraction measures the probability of obtaining face A, and when N is large, this fraction goes to $\frac{1}{6}$. Thus, each time we toss the dice, it is natural to take $\frac{1}{6}$ as the probability of obtaining face A.

Later, we will see that from the law of large numbers these fractions converge to this limit. This limit, $\frac{1}{6}$, corresponds to the concept of the ratio of the number of favorable cases over the total number of cases.

1.1.2 Number of Favorable Cases Over The Total Number of Cases

When we toss a dice, there is a total of 6 possible outcomes, $\{1, 2, 3, 4, 5, 6\}$, corresponding to the letters on faces $\{A, B, C, D, E, F\}$. If we wish to obtain face A and we have only one such case, then the probability of getting face A is quantified by the fraction $\frac{1}{6}$. However, we may be interested in the event {"the observed face is even"}. What does this mean? The even face can be 2, 4 or 6. Each time that one of these three faces appears, we have a realization of the event {"the observed face is even"}. This means that when we toss a dice, the total number

P1: JYS
c01 JWBK307-Huynh September 15, 2008 7:42 Printer: Yet to come

of possible cases is always 6 and the number of favorable cases associated to even events is 3. Therefore, the probability of obtaining an even face is simply $\frac{3}{6}$ and intuitively this appears to be correct.

From this consideration, in the following section we construct in an axiomatic way the mechanics of what is happening. However, we must first establish what is an event, and then we must define the probabilities associated with an event.

1.2 AXIOMATIC DEFINITION

Let's define an universe in which we can embed all these intuitive considerations in an axiomatic way.

1.2.1 Random Experiment

A random experiment is an experiment in which we cannot precisely predict the outcome. Each result obtained from this experiment is random *a priori* (before the realization of the experiment). Each of these results is called a simple event. This means that each time that we realize this experiment we can obtain only one simple event. Further we say that all simple events are exclusive.

Example 2.1 Tossing a dice is a random experiment because before the toss, we cannot exactly predict the future result. The face that is shown can be 1, 2, 3, 4, 5 or 6. Each of these results is thus a simple event. All these 6 simple events are mutually exclusive.

We denote by Ω the set of all simple events. The number of elements in Ω can be finite, countably infinite, uncountably infinite, etc. The example with the dice corresponds to the first case (the case of a "finite number of results", $\Omega = \{1, 2, 3, 4, 5, 6\}$).

Example 2.2 We count the number of phone calls to one center during one hour. The number of calls can be 0, 1, 2, 3, etc. up to infinity. An infinite number of calls is evidently an event that will never occur. However, to consider it in the theoretical development allows us to build useful models in a relatively simple fashion. This phone calls example corresponds to the countably infinite case ($\Omega = \{ 0, 1, 2, 3, \ldots, \infty\}$).

Example 2.3 When we throw a marble on the floor of a room, the position on which the marble will stop is a simple event of the experiment. However, the number of simple events is infinite and uncountable. It corresponds to the set of all points on the floor.

Building a probability theory for the case of finite experiments is relatively easy, the generalization to the countably infinite case is straightforward. However, the uncountably infinite case is different. We will point out these differences and technicalities but we will not dwell on the complex mathematical aspects.

1.2.2 Event

We consider the experiment of a dice toss. We want to study the "even face" event. This event happens when the face shown is even, that is, one of 2, 4, or 6.

Thus, we can say that this event "even face" contains three simple events $\{2, 4, 6\}$. This brings us to the definition:

Definition 2.4 *Let Ω be the set of simple events of a given random experiment. Ω is called the sample space or the universe. An event is simply a sub-set of Ω.*

Is any subset of Ω an event? This question will be answered below. We must not forget that an event occurs if the realized simple event belongs to this event.

1.2.3 Algebra of Events

We saw that an event is a subset of Ω. We would like to construct events from Ω. Let Ω be the universe and let ξ be the set of events we are interested in. We consider the set of all events. ξ is an algebra of events if the following axioms are satisfied:

A1: $\Omega \in \xi$,
A2: $\forall A \in \xi, A^c = \Omega \backslash A \in \xi$ (where $\Omega \backslash A$, called the complementary of A, is the set of all elements of Ω which do not belong to A),
A3: $\forall A_1, A_2, \ldots, A_n \in \xi, A_1 \cup A_2 \cup \ldots \cup A_n \in \xi$.

Axiom A1 says that the universe is an event. This event is certain since it happens each time that we undertake the experiment. Axiom A1 and axiom A2 imply that the empty set, denoted by $\emptyset$, is also an event but it is impossible since it never happens. Axiom A3 says that the union of a finite number of events is also an event. To be able to build an algebra of events associated with a random experiment encompassing a countable infinity of simple events, axiom A3 will be replaced by:

A3': $\cup_{n=1}^{\infty} A_n = A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots \in \xi$.

This algebra of events plays a very important role in the construction of the probability of events. The probabilities that we derive should follow the intuition developed previously.

1.2.4 Probability Axioms

Let Ω be the universe associated with a given random experiment on which we build the algebra of events ξ. We associate to each event $A \in \xi$ a probability noted Prob(A), representing the probability of event A occurring when we realize the experiment. From our intuitive setup, this probability must satisfy the following axioms:

P1: $\text{Prob}(\Omega) = 1$,
P2: $\forall A \in \xi, 0 \leq \text{Prob}(A) \leq 1$,
P3: if $A_1, A_2, \ldots, A_n, \ldots$ is a series of mutually exclusive events, that is: $\forall i \neq j, A_i \cap A_j = \emptyset$, then

$$\text{Prob}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \text{Prob}(A_n). \tag{1.1}$$

P1: JYS
c01 JWBK307-Huynh September 15, 2008 7:42 Printer: Yet to come

Axiom P3 is called σ−additivity of probabilities. This axiom allows us to consider random experiments with an infinity of possible outcomes. From these axioms, we can see that

$$\text{Prob}(\emptyset) = 0 \quad \text{and} \quad \text{Prob}(A^c) = 1 - \text{Prob}(A) \tag{1.2}$$

which are intuitively true.

A very important property easy to derive is presented below.

Property 2.5 *Consider two events A and B, then*

$$\text{Prob}(A \cup B) = \text{Prob}(A) + \text{Prob}(B) - \text{Prob}(A \cap B). \tag{1.3}$$

The mathematical proof is immediate.

Proof: Let $A \setminus C$ be the event built from elements of A that do not belong to C.

$$A = (A \setminus C) \cup C \quad \text{where} \quad C = A \cap B. \tag{1.4}$$

Since $A \setminus C$ and C are disjoint, from axiom P3,

$$\text{Prob}(A) = \text{Prob}(A \setminus C) + \text{Prob}(C). \tag{1.5}$$

Similarly

$$\text{Prob}(B) = \text{Prob}(B \setminus C) + \text{Prob}(C). \tag{1.6}$$

Adding these two equations yields:

$$\text{Prob}(A \setminus C) + \text{Prob}(B \setminus C) + \text{Prob}(C) = \text{Prob}(A) + \text{Prob}(B) - \text{Prob}(C). \tag{1.7}$$

Moreover,

$$A \cup B = (A \setminus C) \cup (B \setminus C) \cup C, \tag{1.8}$$

and since $A \setminus C$, $B \setminus C$ and C are disjoint, we have

$$\text{Prob}(A \cup B) = \text{Prob}(A \setminus C) + \text{Prob}(B \setminus C) + \text{Prob}(C), \tag{1.9}$$

thus,

$$\text{Prob}(A \cup B) = \text{Prob}(A) + \text{Prob}(B) - \text{Prob}(C). \tag{1.10}$$

Example 2.6 Let's go back to the dice toss experiment with

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

and consider the events:

(a) $A = \{\text{“face smaller than 5”}\} = \{1, 2, 3, 4\}$.

P1: JYS
c01 JWBK307-Huynh September 15, 2008 7:42 Printer: Yet to come

Since events $\{1\}$, $\{2\}$, $\{3\}$, and $\{4\}$ are mutually exclusive, we know from axiom P3 that:

$$\text{Prob}(A) = \text{Prob}(\{1\}) + \text{Prob}(\{2\}) + \text{Prob}(\{3\}) + \text{Prob}(\{4\}) = \frac{4}{6}.$$

(b) $B = \{\text{"even faces"}\} = \{2, 4, 6\}$.

Thus, $A \cup B = \{1, 2, 3, 4, 6\}$ and $A \cap B = \{2, 4\}$. We also have, $\text{Prob}(A) = \frac{4}{6}$, $\text{Prob}(B) = \frac{3}{6}$, $\text{Prob}(A \cap B) = \text{Prob}(\{2, 4\}) = \frac{2}{6}$, which implies

$$\text{Prob}(A \cup B) = \text{Prob}(A) + \text{Prob}(B) - \text{Prob}(A \cap B) = \frac{4}{6} + \frac{3}{6} - \frac{2}{6} = \frac{5}{6}.$$

Next, we discuss events that may be considered as independent. To present this, we must first discuss the concept of conditional probability, i.e., the probability of an event occurring given that another event already happened.

1.2.5 Conditional Probabilities

Let A and B be any two events belonging to the same algebra of events. We suppose that B has occurred. We are interested in the probability of getting event A. To define it, we must look back to the construction of the algebras of events.

Within the universe Ω in which A and B are two well-defined events, if B has already happened, the elementary event associated with the result of this random experiment must be an element belonging to event B. This means that given B has already happened, the result of the experiment is an element of event B.

Intuitively, the probability of A occurring is simply the probability that this result is also an event of B. If B has already happened, the probability of getting A knowing B is the probability of $A \cap B$ divided by the probability of B. Therefore, we obtain

$$\text{Prob}(A|B) = \frac{\text{Prob}(A \cap B)}{\text{Prob}(B)}. \tag{1.11}$$

This definition of the conditional probability is called Bayes' rule.

This probability satisfies the set of axioms for probabilities introduced at the beginning of the section:

$$\text{Prob}(\Omega|B) = 1, \tag{1.12}$$

$$0 \leq \text{Prob}(A|B) \leq 1, \tag{1.13}$$

$$\text{Prob}(A^c|B) = 1 - \text{Prob}(A|B), \tag{1.14}$$

and

$$\text{Prob}(\cup_{n=1}^{\infty} A_n|B) = \sum_{n=1}^{\infty} \text{Prob}(A_n|B), \quad \forall i \neq j, \quad A_i \cap A_j = \emptyset. \tag{1.15}$$

This definition is illustrated next by way of examples.

P1: JYS
c01 JWBK307-Huynh September 15, 2008 7:42 Printer: Yet to come

Example 2.7 Consider the dice toss experiment with event

$$A = \{\text{“face smaller than 5"}\} = \{1, 2, 3, 4\}$$

and event

$$B = \{\text{“even face"}\} = \{2, 4, 6\}.$$

We know that

$$\text{Prob}(B) = \text{Prob}(\{2, 4, 6\}) = \frac{3}{6}$$

and

$$\text{Prob}(A) = \text{Prob}(\{1, 2, 3, 4\}) = \frac{4}{6}.$$

However, we want to know what is the probability of obtaining an even face knowing that the face is smaller than 5 (in other words, A has already happened). From Bayes' rule:

$$\begin{aligned}\text{Prob}(B|A) &= \frac{\text{Prob}(A \cap B)}{\text{Prob}(A)} \\ &= \frac{\text{Prob}(\{2, 4\})}{\text{Prob}(\{1, 2, 3, 4\})} \\ &= \frac{2/6}{4/6} \\ &= \frac{1}{2}.\end{aligned}$$

Example 2.8 From a population of N persons, we observe n_s smokers and n_c people with cancer. From these n_s smokers we observe $n_{s,c}$ individuals suffering from cancer. For this population, we can say that the probability that a person is a smoker is $\frac{n_s}{N}$ and the probability that a person has cancer is $\frac{n_c}{N}$. The probability that a person has cancer given that he is already a smoker is:

$$\text{Prob}(\text{cancer}|\text{smoker}) = \frac{\text{Prob}(\text{smoker and cancer})}{\text{Prob}(\text{smoker})} = \frac{n_{s,c}}{n_s}.$$

From this experiment, we note that the conditional probability can be smaller or greater than the probability considered *a priori*. Following this definition of the conditional probability, we examine next the independence of two events.

1.2.6 Independent Events

Two events are said to be statistically independent when the occurrence of one of them doesn't affect the probability of getting the other. A and B are said to be statistically independent if

$$\text{Prob}(A|B) = \text{Prob}(A). \tag{1.16}$$

From Bayes' rule, if A and B are two independent events then

$$\text{Prob}(A \cap B) = \text{Prob}(A)\text{Prob}(B). \tag{1.17}$$

Example 2.9 Consider the experiment of tossing a dice twice. Intuitively, we hope that the result of the first toss would be independent of the second one. From our preceding exposition, we can establish this independence as follows. Indeed, the universe of this experiment contains 36 simple events denoted by $(R1, R2)$ where $R1$ and $R2$ are respectively the results of the first and second tosses, with $(R1, R2)$ taking values (n, m) in

$$\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}.$$

The probability the first element $R1$ equals n is

$$\text{Prob}(R1 = n) = \frac{1}{6}, \quad \forall n \in \{1, 2, 3, 4, 5, 6\}$$

and the probability the second element $R2$ equals m is

$$\text{Prob}(R2 = m) = \frac{1}{6}, \quad \forall m \in \{1, 2, 3, 4, 5, 6\}.$$

Since $\text{Prob}(R1 = n, R2 = m) = \frac{1}{36}$, then the conditional probability

$$\begin{aligned}\text{Prob}(R2 = m|R1 = n) &= \frac{\text{Prob}(R1 = n, R2 = m)}{\text{Prob}(R1 = n)} \\ &= \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6},\end{aligned}$$

which gives us $\text{Prob}(R2 = m|R1 = n) = \text{Prob}(R2 = m) = \frac{1}{6}$. Hence, we conclude that $R2$ and $R1$ are independent.

Notes and Complementary Readings

The concepts presented in this chapter are fundamentals of the theory of probabilities. The reader could refer to the books written by Ross (2002 a and b) for example.