When will I ever use this?

Volume 2: Probabilistic Decision Modeling

Kenneth Chelst and Thomas Edwards

Table of Contents

Introduction Probabilistic Decision Making

The world around us is filled with uncertainty, risk, and variability that complicate day-to-day decisions and the development of long-term plans. This applies to personal decisions as well as decisions made by companies and government agencies. When you work on college applications, you cannot be sure which schools will accept you. In a rush to get a meal between classes, you will face uncertainty in the time it takes to be served at the school cafeteria. The school newspaper editor is concerned about how many members of the writing team will meet their deadlines. While reviewing alternative car insurance plans, the student driver struggles to decide on the size of the collision insurance deductible. Obviously, no one plans to have an accident, but the risk of an accident is always present. Companies that provide insurance look at the same problem. They come up with pricing strategies for insurance that pool the risk of an individual with large groups of similar people.

> ***Variability*** is a characteristic of data that refers to the recognition that each data value is not the same. For example, there is variability in height of individuals or in their annual income. Two common statistics used to characterize a dataset's variability are variance and standard deviation.

Companies launching a new product or service must deal with uncertain demand. Police patrol supervisors must consider random fluctuations in the demand for service and patterns of crime when planning how many patrol officers are needed on each shift and where to place them. Plant managers and school officials must cope with workers who randomly do not show up for work due to illness.

This text is designed to provide decision-making guidance in the presence of uncertainty. One of the challenges in learning basic concepts of probability is that many of us do not have good intuition about randomness. We will address this problem while developing probabilistic decision-making skills. The text is therefore designed to develop your ability to recognize and understand patterns of random events. We will do so by having you simulate random experiments first with a coin flip, then with the random number function in your calculator, and lastly, using the random number generator in Excel to develop and analyze a large sample.

In introducing basic concepts, we routinely use the concept of ***relative frequency*** as an estimate of probability. Thus, our introductory examples will use data rather than the counting methods that you may have seen in other probability courses.

> ***Randomness*** is a characteristic of a process, experiment, or environment in which outcomes cannot be predicted with certainty. A random experiment, such as rolling a die, can yield different unpredictable outcomes. The gender of newborn is a result of a random process involved in becoming pregnant. The temperature on any day reflects randomness of our environment.

In later chapters we introduce mathematical formulas that can be used to describe different patterns of randomness. These are probability distributions. These formulas will be applied to two different types of variables: discrete and continuous. A ***discrete variable*** is countable. For example, the number of crimes in a day and the number of people absent from work are discrete variables. On the other hand, a ***continuous variable*** cannot be counted and is measured instead. For example, the time to complete an exam and the height of an individual are continuous variables.

Managing and Deciding in an Uncertain World

It is not possible to make 100% accurate predictions as to who will win the Super Bowl. Management of a plant or school cannot know with certainty how many people will show up for work each day. Nor can a dealer know how many cars he will sell each month. However, probability theory and sophisticated probability models enable us to identify individual occurrences that are more or less likely, as well as estimate long-term averages. This knowledge and understanding is essential for planning and making critical decisions in advance as well as adjustments on the spot. A plant or office manager must decide how many spare workers to have on duty to address the problem of absenteeism. The high school newspaper editor will also have to plan what to do if only eight of 10 writers have completed their assignments. Similarly, an automobile dealership needs to know how much inventory to carry. Often we must balance the added costs of having reserves against the impact of running short.

Chapter 1: Basic Probability and Randomness

We all face decisions in our jobs, in our communities, and in our personal lives that involve uncertainty. When making such decisions, there can be no guarantee that the outcome will be favorable. Probability decision models are designed to help an individual to assess the likelihood of various outcomes. The decision maker then uses that information to select a preferred choice that explicitly considers the relative likelihood of both positive and negative results.

In perceiving and responding to the world around us, it is critical to develop an understanding of the difference between patterns of random and on-random events. There is well documented fallacy in probabilistic thinking in which people see patterns when in fact all they have observed is random fluctuations. A decision maker faced with a series of negative outcomes must decide whether or not act. Could this simply be a random fluctuation and no action needs to be taken? Or is there something causing these negative outcomes that he can and should be addressed?

In this chapter we strive to develop a better understanding of random patterns through the use of simulated experiments. We begin with a physical simulation, flipping a coin. This can be used to model randomness with two equally likely events. Random number generators are introduced to represent randomness when two events are not equally likely. These random number generators are found in advanced calculators and Excel spreadsheets.

This chapter introduces the basic concepts and notation of probability. The multiplication rule is used to calculate the probability of occurrence of two independent random events. The principle of complementarity is used to indirectly compute the likelihood of an event. Missing from this chapter is any discussion of counting methods, combinations and permutations. Historically, these were developed primarily to apply probability to games of chance. In our experience we have never had to use counting rules to tackle non-gambling decisions involving uncertainty.

Chapter 2: Conditional Probability

Chapter 2 introduces the concept of conditional probability, P(B|A). Given that event A has occurred, what is the probability that event B will occur? Unlike most probability textbooks, we dedicate an entire chapter to explore conditional probability. Research has shown that many people have poor intuition with regard to conditional probability concepts.

This chapter illustrates how to determine conditional probability from data tables as well as from a problem context. It illustrates the application of the multiplication rule to calculate joint probabilities of two events that are not independent. The chapter proceeds to develop the formula for a probabilistic partition. In this chapter we introduce the concept of a random variable. This is a function that translates the outcome of a random experiment into a unique numeric value. The measure of central tendency of this random variable is a probabilistic weighted sum called the expected value.

Chapter 3: Decision Trees

Decision trees provide a structure for determining the alternative that optimizes the expected value. Decisions are represented by rectangles and uncertain events by ovals. Branches emanating from a rectangle correspond to distinct decision alternatives. Branches from a random event correspond to different possible outcomes. The decision tree also provides the probability distribution for each of the alternatives.

The first example simply introduces the concept of tradeoff between more or less conservative decision alternatives and possible outcomes. Two subsequent decision contexts maximize the expected profit. The last example minimizes the expected cost of collision insurance.

Chapter 4 Binomial Distribution and Geometric Distribution

In Chapter 1 we used basic probability and simulation to study random events with two mutually exclusive outcomes: answer or do not answer phone, at work or absent from work, on time or late. In chapter 4 we introduce the concept of a probability distribution function. When repeated random events follow the same assumptions, there may be a formula that summarizes the probabilistic pattern. This formula will have parameters whose values vary from context to context.

The binomial distribution can be applied to a situation of identical independent repetitions of the same random experiment. The four common elements to the random experiment are:

1. only two possible outcomes: success and failure
2. p, the probability of success, is the same for each repetition
3. each repetition is independent of every other repetition
4. n identical repetitions of the random event

The parameters of the binomial distribution are n and p. The formula for the binomial distribution calculates the probability of X successes out of n trials. The random variable X has a range from zero to one. Statistical calculators and Excel spreadsheets include formulas for the binomial distribution and the cumulative binomial distribution. The binomial distribution examples in this chapter are all extensions of those that appeared in Chapter 1.

In the binomial distribution we count the number of success out of n trials. The geometric distribution includes the same first three elements as the binomial. However, the random variable that is tracked is the number of repetitions until the first success. This random variable can take on the values of one to infinity. One interesting application of the geometric distribution involves the number of shuttle flights until the first shuttle disaster.

Chapter 5 Poisson Distribution

The Poisson distribution is used to characterize a probabilistic environment in which random events occur totally independent of one another. Emergency calls to 911 or calls to a telephone helpline are examples of situations that have been modeled with the Poisson distribution. This distribution has one parameter, λ, the average number of incidents per unit of time. The distribution is used to estimate the probability that there will be X events in a unit time. Alternatively, it can be used to model an extended time period t. In that case the average or expected value of the random variable is λt. The random variable X is discrete and has a range of one to infinity. Managers face the difficult task of scheduling the necessary resources to handle the fluctuations in workload caused by the unpredictability of these random events.

Chapter 6 Normal Distribution

The Normal distribution is used to describe a continuous random variable with a mean of μ and a standard deviation of σ. The sum of identically distributed independent random variables is known to approach the normal distribution. For that reason the randomness of the sample average of a random variable can be approximated with the normal distribution. In the first chapter example, the normal distribution is used to model the random error associated with cutting panels of material for use in constructing a parachute. The normal distribution can help determine the proportion of panels that are within specifications. The second example uses the normal distribution to help establish an appropriate warranty replacement policy. In the final example, a manager has to decide how many units of a new toy to stock when the demand is normally distributed. The decision maker has to balance the risk of having too few items and losing out on sales against the risk of overstocking and having to steeply discount toys that are left over after the buying season passes.