10 Probability Theory I

Employee: I don't need your crummy job, Mr. Employer! I've won the lottery.
Employer: Well, who needs employees? I won the lottery, too!
All: [to camera] Why don't you win the lottery, too.
Announcer: The state lottery, where everybody wins!
(actual odds of winning, one in 380,000,000.)

From: *The Simpsons*

10.1 Introduction

In the last chapter on regression we discussed in somewhat vague terms the use and meaning of the regression line as a predictor. We talked about the value $\hat{y}$ on the regression line as a theoretical average for some unspecified population of data that was similar to the actual data. We even discussed how this hypothetical data would be spread around the theoretical average value as determined by the regression. In scientific work it is common to make statements about an underlying population from information based on a sample. Probability theory is needed to make these statements precise.

It would be nice to begin a discussion of probability by defining the term. This doesn't seem possible in a single way that satisfies a variety of intuitive ideas about probabilities. In this chapter we will examine some aspects of this naïve intuition. Refining this intuition leads to a variety of rules that allow us to quantify and enlarge the scope of our understanding.

10.2 The Frequentist Viewpoint

You have a sore throat so you go to the doctor. A throat culture reveals a bacterial infection for which the doctor prescribes an antibiotic. You expect the principal symptoms will be alleviated in 48 hours and the illness will run its course if you complete the treatment as specified. Until the advent of drug resistance in certain infectious bacteria this scenario involved no uncertainty, as long as a physician and medication were available. We might call this the gold standard for medical treatment.

For most illness, however, we are far from the gold standard. Causative factors may be unknown or only dimly understood. Treatments may only work some of the time and for reasons not predictable in advance or even after the fact. We are now in the realm of probabilities, where the question is "Doctor, what are the chances this will work?" When a physician tells you that a certain medication works about 60% of the time in treating a certain illness, you assume this statement derives from clinical experience and not wishful thinking. From a practical standpoint most of us are "frequentists". This means simply that we interpret probabilities attached to events as reflecting empirical evidence based on relative frequencies.

Of course, not all evidence has the same merit. If the doctor's judgment of 60% efficacy is based on 25 cases, we may not feel so confident that it is universally applicable. Even if it were based on the outcomes in thousands of cases we have seen (recall the *Literary Digest* story from Chapter 7)

that non-random selection can bias a result. Nevertheless, needing some guidance we adopt as a working definition,

---

Definition 10.1: (**Relative Frequency Method**) If an "experiment" is repeated a <u>large number of times</u>, the probability of any outcome E, denoted by $P(E)$, is approximately

$$\frac{\text{\# of times } E \text{ occurs}}{\text{\# of repetitions of the experiment}}. \blacksquare \qquad (10.1)$$


By an "experiment" we mean a set of conditions which, in principle, can be replicated. For example, in evaluating the effectiveness of a medical treatment our experiment consists of selecting patients from the pool of people needing treatment. (Of course the actual medical treatment might also constitute an experiment in the medical sense.) The reliability of the ratio (10.1) as a measure of the true probability $P(E)$ depends heavily on the number of repetitions. We can get some sense of this by conducting simulated experiments using the computer.

Example 10.1: Using a computer simulate the tossing of a coin 50, 100, 500, and 1000 times. Determine the relative frequency of heads (ratio (10.1)) in each experiment.

Solution:

It is not at all obvious that a deterministic machine like a computer can be used to simulate a random process like tossing a coin. We will describe in a later chapter how such a simulation is done. For now you should believe that it adequately reflects an experiment with a real coin. What results do we find?

The figures below shows the outcomes of four separate simulations performed using the file *coin.xls*. The number of "tosses" was 50, 100, 500, and 1000, respectively. The graphs show the variation in the frequency of heads (the ratio (10.1)) throughout the entire simulation as well as the final frequency of heads.
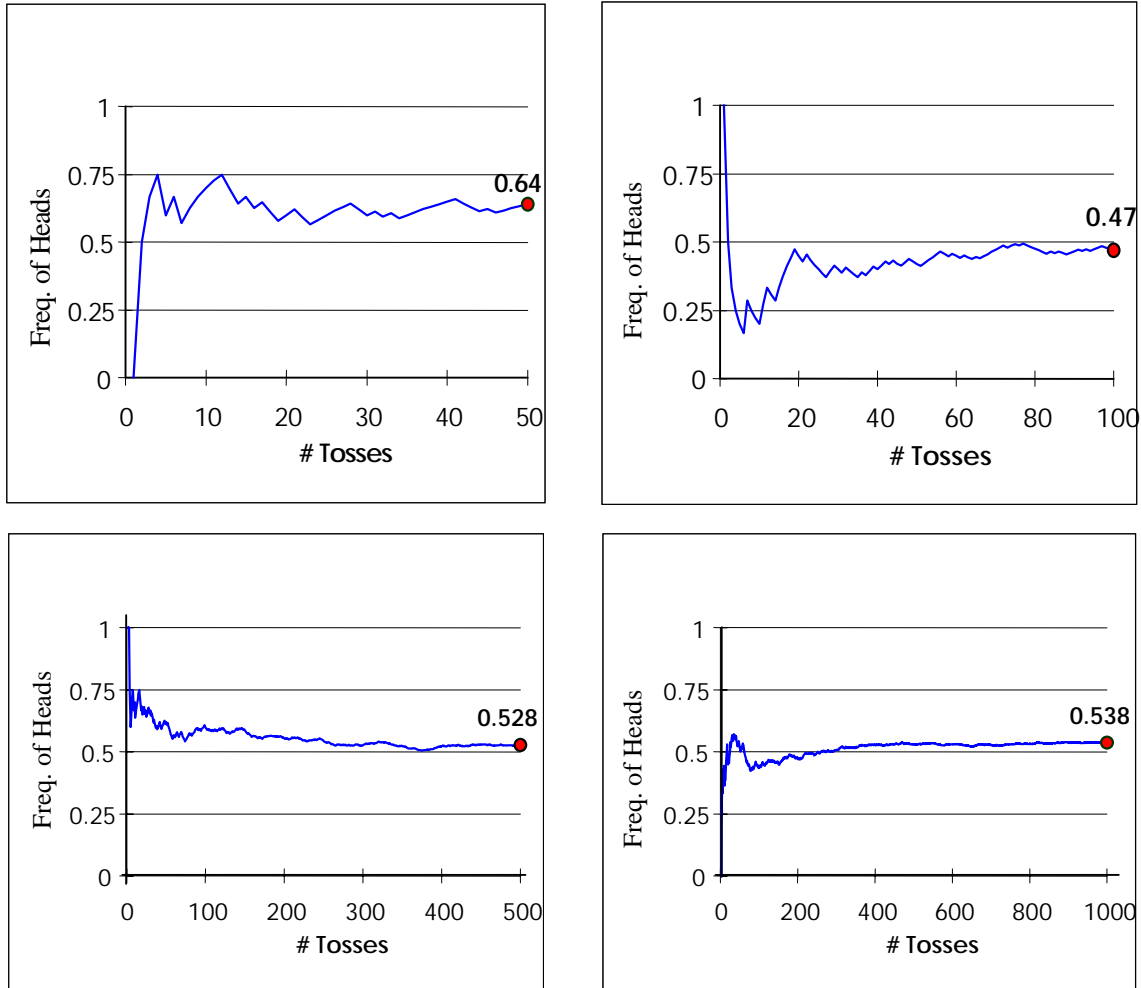
Figure 10.1: Coin Tossing Simulation

Initially there is a good deal of variation in the frequency of heads. The variability appears to diminish as the number of tosses increases.∎

If we wish to assign a probability to a coin landing heads, what number can we use based on the data in Example 10.1? The answer is certainly not obvious, even after 1000 tosses. We might want to extend the series of tosses under the belief that a longer series will give a better approximation to the truth than a shorter one. This is usually true, but it is not guaranteed. In short, we must admit that using Definition 10.1 as a basis for assigning theoretical probabilities seems a hopeless task, although that definition is at the heart of most applications. Example 10.2 below is a typical one of these.

Example 10.2: A survey of 1200 voters showed 500 planning to vote for candidate A, 450 for B and the remainder undecided. What is the probability that a randomly selected voter in the population at large will support candidate A?

Solution:

The relative frequency of A's supporter's among the voters sampled is $500/1200 \approx .42 = 42\%$. According to Definition 10.1 this number (.42) should be an <u>approximation</u> to the actual probability that a randomly selected voter would support A. Of course, since this estimate is based on a sample of voters it will not exactly equal the true probability. Later (Chapter 15) we will learn how to estimate the likelihood that the unknown true probability differs from the estimate by $\pm.01, \pm.05$ etc. For now the best we can say is that the probability in question is 0.42.∎

To obtain a foundation for the subject unencumbered by approximations and estimates we must look at a different approach to constructing probabilities, albeit one which cannot be so easily applied in practical situations (other than card games and the like).

10.3 Classical Probability

Probability theory began in the 17th century, with its roots in questions related to understanding games of chance. The connection of these ideas to statistical problems involving relative frequencies (the frequentist viewpoint) came considerably later. From the classical viewpoint there is no difficulty determining the correct probability of tossing a head with a fair coin. The reader surely knows the answer. Indeed, ignoring the remote chance that the coin will land on an edge, there are two conceivable outcomes, usually designated H for heads and T for tails. The classical or equally likely method of assigning probabilities asserts that, with insufficient information to predict the outcome, either of the two possibilities has the same chance of occurring, which should then be 1 out of 2 or 1/2.

Now this answer is tolerably close to the empirical results found by simulation in Example 10.1, though perhaps the reader is disappointed that the agreement is not better, even with 1000 tosses. This is an instance, one of many in the subject, where untutored intuition as to what should happen is often incorrect. After we learn more about probability theory we shall be able to better understand the connection between the frequentist and classical probability models.

We can extend the idea behind our simple analysis of coin tossing to a general method of assigning probabilities when no single outcome of an experiment appears to have any greater chance of occurring than any other. This method is called the equal likelihood or classical method.

Definition 10.2 (**Equal Likelihood Method**): If an "experiment" results in n distinct "elementary" outcomes, no one of which seems more likely to occur than any other, then we assign probability $1/n$ to each outcome.∎

The elementary outcomes referred to in this definition are collectively called the ***sample space*** of the experiment. Note that since there are n elementary outcomes and we give probability $1/n$ to each, the sum of all the probabilities associated with these elementary outcomes is one.

Example 10.3: What is the probability of tossing a three with a single cubical die?

Solution:

There are six possible elementary outcomes when a cubical die is tossed. We can designate the sample space as the set {1, 2, 3, 4, 5, 6}. Assuming the die has not been tampered with, any outcome should have as much chance of occurring as any other. From Definition 10.2 the probability of each outcome should be 1/6. Thus we may write, $P(3) = 1/6$, where $P(3)$ stands for the probability of obtaining a three.■

10.4 Counting

In order to apply Definition 10.2 we must count the number of outcomes that are possible in an experiment. There are two ways in which to do this. We can either enumerate all the outcomes or figure out how to count them without actually making an explicit list. The latter option is clearly more desirable if there are a large number of possibilities. However, it is often a more difficult conceptual challenge. In this section we develop a few techniques that can be used effectively for both these tasks.

Example 10.4: Suppose we toss a dime and a nickel. Describe the possible outcomes and assign probabilities to each.

Solution:

The dime can land H or T, and following any of these outcomes the nickel can land H or T as well. This gives 4 outcomes, which we can denote by pairs (H, H), (H, T), (T, H) and (T, T), the first entry indicating the outcome for the dime. The outcome designated by (H, T) is not the same as the one denoted by (T, H); why not?

Since the individual outcomes for each coin show no preference for H or T, we should not expect any of the four possibilities listed above to occur with greater likelihood than any other. Thus each is assigned probability 1/4.■

What if we toss three coins and record the sequence of heads and tosses obtained. How many outcomes are possible?

Example 10.5: Enumerate the possible arrangements of heads and tails that can be obtained by tossing three coins.

Solution:

A "tree diagram" is a useful device for systematically obtaining the complete list. Figure 10.2 below shows an example of this construction applied to this problem.
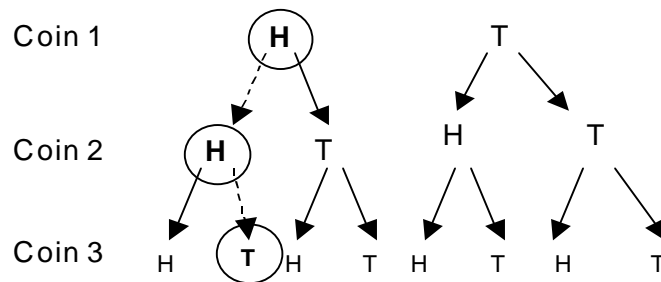
Tree Diagram

Figure 10.2

With this type of diagram, we can enumerate the outcomes for many multi-stage experiments. We construct the diagram as follows: We list the first stage outcomes in the top row. In the second row, each of the first stage outcomes is followed by the 2nd stage outcomes with which it is compatible. The process continues in the third row. The result looks like an inverted tree, hence the name. All possible outcomes can be enumerated by starting at any one of the top nodes and following all possible paths to the bottom. For example, the circled outcomes in the diagram depict the result (H, H, T), where the first two coins land heads and the third lands tails.

There are 8 possible paths from top to bottom and since each component of the path involves a simple head/tail choice, neither of which is more preferable to the other, any of the eight results should have the same probability of occurring. Hence, we assign probability 1/8 to any of the eight outcomes.∎

The tree diagram leads to one of the most important counting principles in probability theory, the multiplication rule.

Rule 10.1 (Multiplication Rule): If in a multi-stage experiment there are m options for the first stage, and each of these can be followed by n options for the second stage, and any of these can be followed by p options for the third stage, etc… then the total number of outcomes for the multi-stage experiment is $m \times n \times p \times \cdots$.∎

Indeed, if we were to construct a tree diagram for this process, we would start with m nodes in the top row. Each of these would be followed by n arrows leading to the second row, for a total of $m \times n$ paths. Continuing, any of these paths would have p extensions into the third row, for a total of $m \times n \times p$ paths through stage three, and so on thereafter.

Let's consider some examples that illustrate the application of the Multiplication Rule (Rule 10.1).

Example 10.6: How many outcomes are possible if a coin is tossed 4 times in a row? What probability should be assigned to each outcome?

Solution:

This is a four-stage experiment, where each stage consists of tossing the coin. At every stage, whatever outcome has been obtained up to that point, there are 2 possible outcomes, {H, T}, available at the next stage. The Multiplication Rule (Rule 10.1) implies that the total number of outcomes is $2 \times 2 \times 2 \times 2 = 2^4 = 16$. Each of the 16 outcomes has the same chance of occurring as any other. Hence, following Definition 10.2, we assign probability 1/16 to each.

The result in the last paragraph seems unexceptional, until one thinks about it a little. For example, it says that the outcome (H, H, H, H), in which we get four consecutive heads, has the same chance of occurring as (H, T, T, H) in which the pattern of heads and tails is not particularly interesting. However, whether we find it interesting or not, this <u>particular</u> pattern (and any other) has the same probability of occurrence as the "streak" of four consecutive heads. Note, however, that we have only considered the probability of particular patterns. We have not answered questions such as, "What is the chance of getting two heads when you toss a coin four times?" which we will consider in the next section.

What if, instead of tossing a coin four times, we had tossed four coins? The reader should convince him or herself that the solution above works equally well. From the point of view of probability theory, there is no difference in these scenarios.∎

Example 10.7: If we toss two standard six-sided cubical dice, how many outcomes are possible? What is the probability of each?

Solution:

This is a two-stage experiment. There are six possible outcomes for the toss of the first die, and any of these can be followed by six possible outcomes for the toss of the second die. The Multiplication Rule (Rule 10.1) asserts that there are a total of $6 \times 6 = 36$ possible outcomes for the two-stage experiment. Each outcome occurs with probability 1/36.∎

It will be useful later to have a complete enumeration of the outcomes for this experiment. We could construct a tree diagram for this experiment, although it becomes somewhat unwieldy. For two-stage experiments in which the choices available in stage two are the same regardless of the outcome in stage one, it is often simpler to use a rectangular array to do the enumeration. This technique is used frequently in genetics, where it goes under the name of Punnett Squares. This leads to the array below, in which the body of the table shows the 36 possible outcomes, with the top row and initial column showing the possible outcomes for each individual die.

die 2 → / die 1 ↓	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	**(1,4)**	(1,5)	(1,6)
2	(2,1)	(2,2)	**(2,3)**	(2,4)	(2,5)	(2,6)
3	(3,1)	**(3,2)**	(3,3)	(3,4)	(3,5)	(3,6)
4	**(4,1)**	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Table 10.1: Tossing A Pair of Dice

Example 10.8: Within a given telephone exchange (say 212), how many individual seven digit telephone numbers are possible? Assuming an automatic telephone dialer is programmed to randomly generate such numbers, what is the probability it will select one particular number?

Solution:

Generating a phone number is a 7-stage experiment. The first digit can be selected from the numerals 2, 3, 4 … 9, giving 8 choices. (Why are 0 and 1 excluded?). At each of the remaining six stages the digits can be selected from any of the ten numerals, 0, 1, 2, … 9. Applying the Multiplication Rule (Rule 10.1), the total number of phone numbers is $8 \times 10^6 = 8,000,000$. If the dialer selects numbers randomly, meaning any of the 8 million numbers has the same chance of being selected, the probability of selecting a particular number is $\dfrac{1}{8 \times 10^6} = .125 \times 10^{-6}$. ∎

Example 10.9: (Sampling Without Replacement): A bowl contains ten balls numbered 1 through 10. You select a ball, record its number, place it aside and then repeat the process, drawing altogether three balls. How many outcomes are possible?

Solution:

This is a 3-stage experiment. In the first stage there are 10 possible choices. Whatever ball is selected in the first stage, there are nine choices remaining for the 2nd stage selection. The precise choices available at the second stage will depend on the first ball picked, but the <u>number</u> of choices is always nine. Similarly, following the first two selections there are eight selections possible for the third choice. The Multiplication Rule (Rule 10.1) implies that the total number of choices is $10 \times 9 \times 8 = 720$.

The sampling scheme described in this example is called *sampling without replacement*. The alternative, wherein the selected items are returned to the selection pool before the next sample is

drawn, is called *sampling with replacement*. The reader should show that if we carried out this experiment using the replacement technique then the number of outcomes would be 1000.■

10.5 Probability Rules

In the previous section we described the outcomes of an experiment in terms of what are usually called elementary events. The probabilities of these elementary events can be used to compute probabilities of more complex outcomes.

Example 10.10: What is the probability of tossing a sum of five with two dice?

Solution:

Referring to Table 10.1, the boldface entries (1,4), (2,3), (3,2) and (4,1) give all outcomes with a sum of five. Since each of these outcomes has probability 1/36, we have four chances out of 36 to obtain a sum of five, so the probability of the latter event should be 4/36 = 1/9.

Is our intuition correct? If we toss a pair of dice many times, what happens to the empirical relative frequency for the occurrence of a sum of five? A simulation of 10,000 tosses produced the following results for the sums of the two dice.

Sums	Freq	Rel. Freq	Theory	Sums	Freq.	Rel. Freq	Theory
2	302	0.0302	0.0278	8	1344	0.1344	0.1389
3	561	0.0561	0.0556	9	1114	0.1114	0.1111
4	844	0.0844	0.0833	10	839	0.0839	0.0833
5	*1083*	*0.1083*	*0.1111*	11	576	0.0576	0.0556
6	1387	0.1387	0.1389	12	276	0.0276	0.0278
7	1674	0.1674	0.1667				

Table 10.2: Dice Tossing Simulation

The fourth and last columns contain the theoretical values. For example, in the bold and italicized row the entry in the theory column is $1/9 \approx 0.1111$. By referring to Table 10.1, the reader should work out the exact fractional values for the other entries in the Theory columns and then convert these to the decimal values in the Table 10.2. Generally speaking, there is good agreement between the observed relative frequency and the theoretical probability. If you expected to see three or four place agreement in all categories you are undoubtedly disappointed. Bona fide probability experiments show a surprising amount of variability. Later in the course, we will learn to distinguish between likely and unlikely variation, but to reiterate, one's untutored intuition is not very helpful in making such assessments.■

We can summarize the idea behind Example 10.10 as a rule.

Rule 10.2 (**Equal Likelihood Principle**): Suppose an experiment produces n elementary outcomes each having probability $1/n$. If k of these outcomes are favorable to the occurrence of an event E, then the probability of E is given by $P(E) = \dfrac{k}{n}$. ■

Note that since $0 \le k \le n$ the probability of any event satisfies $0 \le P(E) \le 1$. For reference we place this together with an earlier remark into a separate rule.

Rule 10.3 (Zero-One Principle):

- The probability of any event is a number between 0 and 1.

- An event that cannot occur has probability zero and an event that is certain to occur has probability one.

- The sum of the probabilities of all elementary events totals one.

As we will see shortly, the Equal Likelihood Principle provides a powerful tool for computing probabilities. When we have computed probabilities using this rule or others to be considered later, the relative frequency interpretation of probability given in Definition 10.1 is often cast in a slightly different form. Namely, instead of using Definition 10.1 to estimate the probability as

$$\text{prob} \approx \tfrac{\text{\# of occurences}}{\text{\# of trials}}$$

we can turn this around and use the probability to estimate how often the event will occur when the experiment is repeated a large number of times. We call this the Frequency Rule (Rule 10.4).

Rule 10.4 (Frequency Rule): If an outcome E of an experiment has probability p and the experiment is repeated n times, then E will occur approximately $n \times p$ times. ■

For example, from Table 10.2, the theoretical probability for obtaining a sum of five is 0.1111. Therefore, according to the Frequency Rule (Rule 10.4), in 10,000 tosses we should see approximately 1,111 tosses that produce a sum of five. In the actual experiment, there were 1,083 tosses with a sum of five.

We consider several more examples illustrating the use of the Equal Likelihood Principle (Rule 10.2).

Example 10.11: If a coin is tossed 4 times, what is the probability of obtaining exactly two heads?

Solution:

We considered this experiment in Example 10.6. If E denotes the event in question, we must count how many elementary events produce exactly two heads. A typical example of such an elementary event would be (H, H, T, T). How many examples are there like this? We could write

the entire tree diagram and go through all 16 branches in search of those containing two Hs. More directly, we can begin with the example (H, H, T, T) in which the first toss produced a head, H. Are there other outcomes like this? If the first toss is an H then, if there is exactly one other H, it could appear as the second toss or the 3rd or the 4th. This gives the three outcomes (H, H, T, T), (H, T, H, T) and (H, T, T, H).

However, there are more outcomes with exactly two heads. The first head may not appear until the second toss. Using exactly the same reasoning as in the preceding paragraph, there are two elementary outcomes that have this property and also have exactly two heads: (T, H, H, T) and (T, H, T, H). Finally, there is one last possibility. The first of the two heads appears only at the third toss, (T, T, H, H). So altogether, there are 6 elementary events satisfying E and $P(E) = 6/16 = 3/8$.∎

Example 10.12: In Example 10.5 we considered tossing three coins. What is the probability of the event E in which we obtain at least one head?

Solution:

Referring to the tree diagram Figure 10.2, we see that there are eight elementary events, each of which has probability 1/8. In seven of these at least one head is obtained. Thus by the Equal Likelihood Principle (Rule 10.2), the probability of E is 7/8.∎

Example 10.12 suggests another rule that is very helpful when computing probabilities of events described by phrases such as "at least " or "at most." If E is an event, the event E^c will denote the *complementary event*, which is the event that E does not take place. Referring to the event E in the previous example, if this event does not occur, then no heads are obtained in the three tosses. Equivalently, the complement E^c is the event in which all tosses produce a tail. This is a much simpler event to describe than event E. It consists of only one elementary event, (T, T, T) and so has probability 1/8. Using our next rule we can compute either $P(E)$ or $P(E^c)$ if we know the other.

Rule 10.5 (Rule of Complements): For any event E we have $P(E) + P(E^c) = 1$.∎

The rule is true because any elementary event will produce the occurrence of exactly one of the events E or E^c. Thus, the sum $P(E) + P(E^c)$ adds the probabilities of all elementary events, which totals one. As remarked prior to the statement of the rule, we can easily apply the Rule of Complements to Example 10.12. Since $P(E^c) = 1/8$, the rule implies that $P(E) = 1 - 1/8 = 7/8$.

Example 10.13: Ten people are in a room. What is the probability that at least two of them have the same birthday (neglecting the year of birth)?

Solution:

We first need to visualize the multi-stage experiment to be performed. We ask each person to state his or her birthday. Since there are 10 people, this is a 10-stage experiment. How many elementary outcomes are possible? Ignoring a possible birth date of Feb. 29th, the first person's birthday can have any one of 365 possibilities. The same number of choices is possible for the 2nd person, and for the 3rd, and so on. The Multiplication Rule (Rule 10.1) tells us that the number of elementary outcomes is 365^{10}.

Suppose B is the event that at least two people have the same birthday. What is the complementary event, B^c? That's simple. All ten people have different birthdays. How many elementary events meet this condition? There are still 365 possibilities for the birthday of the first person, but since the second person must have a different birthday, only 364 possibilities are available for the second date. Similarly, if the 3rd person is to avoid having a common birthday with either of the first two, there are only 363 choices available for the third position. Applying the Multiplication Rule (Rule 10.1) gives $365 \times 364 \times 363 \times \cdots \times 356$ possible elementary events satisfying B^c. Therefore, by The Equal Likelihood Principle (Rule 10.2)

$$P(B^c) = \frac{365 \times 364 \times \cdots 356}{365^{10}}.$$

A decimal approximation to this fraction may be computed using a simple calculator, if you don't mind entering many numbers. You should find $P(B^c) \approx .88$ and therefore $P(B) = 1 - P(B^c) \approx .12$. Thus, there is about a 12% chance that two people in the group will have the same birthday.∎

10.6 Summary

We have examined the classical case of experiments in which exact probabilities may be assigned using elementary events that are equally probable. This technique applies to many random events generated through physical devices, for example coin tossing, card games, roulette wheels, and lotteries. It also applies to certain problems in genetics, as we will consider in Chapter 12. In these situations we find that the probabilities computed theoretically are in general agreement with the relative frequencies observed in a large number of repetitions of the experiment. For many other random phenomena, probabilities can be estimated from the empirical relative frequencies using the Relative Frequency Method (Definition 10.1). The accuracy achievable with this method will be taken up later in the course (See Chapter 15).

There is no uniform prescription for finding probabilities using the rules discussed in this chapter. If you think that the equal probability model is applicable, you should observe the following steps.

1. Get a clear conception of the possible elementary outcomes of your experiment and convince yourself (and possibly someone else) that these events have equal chance of occurring.

2. Figure out how many or which elementary outcomes produce a particular event E in which you are interested. Then use The Equal Likelihood Principle (Rule 10.2), to find $P(E)$. Remember, $P(E)$ must be a number between 0 and 1 (Zero-One Rule (Rule 10.3)).

3. If you can, perform a computer simulation of the experiment to test your answer. (In Chapter 13 we will discuss how to use *Excel* to carry out many simulations.) If, after many repetitions, the relative frequency with which E occurs in the simulation does not seem near your theoretical value, rethink your solution.

10.7 Exercises

1. The frozen yogurt machine has three yogurt flavors and four toppings. If you partake once a day and always select both a yogurt and a topping, how many days can you go before you have to repeat a combination?

2. A certain airline has 10 flights a day from N.Y. to Chicago. From Chicago there are 8 flights a day to L.A. Can you conclude that there are 80 possible ways to book a flight from N.Y. to L.A. via Chicago? Explain.

3. a) The DNA in the cell nucleus contains chains of four nucleotide bases adenine (A), thymine (T), cytosine (C), and guanine (G). Via the so-called genetic code, a unique triplet of bases, called a codon, determines the synthesis of an amino acid. How many amino acids can be coded for in this scheme?

   b) Twenty amino acids are needed to synthesize any protein. Given this fact, could the genetic code have worked with a codon of two bases?

4. a) A man and woman wish to have three children. Make a tree diagram showing all possibilities for the sexes of the children.

   b) Using your answer to a) show that the probability that the couple will have a child of each sex is 3/4.

5. Show how to obtain the probabilities listed under the "Theory" column in Table 10.2.

6. There are approximately 131,600 cases of colorectal cancer in the U.S. each year (data as of 1998). The U.S population is approximately 270 million. What is the probability that a randomly selected person in the U.S. will develop this type of cancer during a year? What method of assigning probabilities are you using to make this determination?

7. a) According to the American Cancer Society (www.cancer.org) 62% of patients treated for colorectal cancer survive more than 5 years; the ten-year survival rate is 51%. Express this data in terms of probabilities.

   b) What is the probability that a patient with colorectal cancer will survive between 5 and 10 years?

8. The weather forecast states that there is a 70% chance of rain tomorrow. Which of the "definitions" of probability is relevant to this statement and explain why?

9. Suppose the probability of being killed in an auto accident in the U.S. during a given year is .00016. Given the current population of 270 million, how many auto fatalities would be expected in a year? Describe the probability principles involved.

10. In a clinical trial of a new drug 5% of the users reported significant side effects. How would you interpret this information in terms of probabilities? What aspect of your answer might depend on how many people participated in the trial?

11. Some board games use dice in the shape of a tetrahedron, with the four faces numbered 1, 2, 3, and 4, as shown below. The side facing down is counted. In the figure below this would be the numeral four.
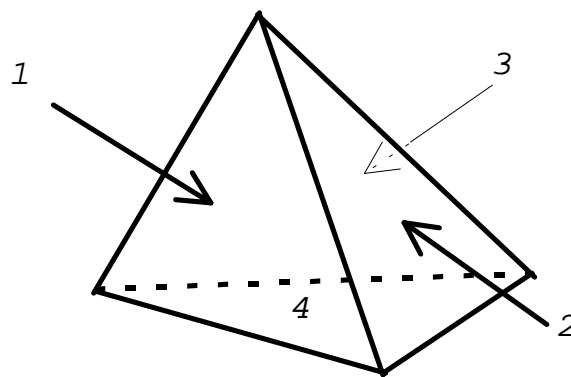
Figure 10.3: A Tetrahedral Die

   a) How many elementary outcomes are possible when two such dice are tossed? Display the results as a tree diagram and in a rectangular array.

   b) Using a) determine the probabilities for obtaining all possible sums for two such dice.

12. a) A bowl contains three red balls and two black balls. Explain in terms of equally likely elementary events why the probability of selecting a red ball is 3/5 and the probability of selecting a black ball is 2/5?

   b) Suppose we select two balls <u>without replacement</u>. Construct a tree diagram showing the possible outcomes as 20 elementary events of equal probability. (Hint: You should distinguish the balls from each other, say as R_1, R_2, R_3 and B_1, B_2.)

   c) Using b) find the probabilities for each of the following events:

   i)   Both balls selected are black.

   ii) At least one of the selected balls is black.

   iii) The two selected balls have different colors.

13. In each of the following decide whether the stated outcomes should be assigned equal probabilities. If not, how would you determine the correct probabilities?

   a) Passing or failing an exam for a randomly selected student from a class.

b) A patient is given a TB screening test ("patch" test) as part of a routine physical examination. The events are whether the test shows a positive reaction (indicative of illness) or a negative one.

c) A randomly selected person's blood is classified into one of the four categories, O, A, B, AB.

d) You spin the arrow on the circular diagram below and record the number of the sector on which it lands (The spin is repeated if the arrow lands on a boundary line).
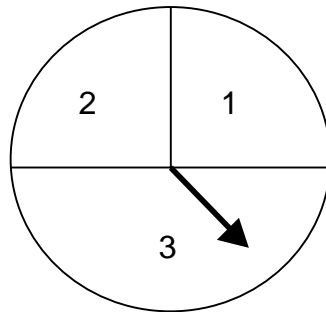
e) A large bowl contains both red and black balls. You select a ball at random and observe its color.

14. A bowl contains four balls numbered 1 through 4. You select the four balls successively, without replacement, until none remain in the bowl.

a) What is the probability that the two even numbered balls are selected before the odd numbered balls?

b) What is the probability that the balls are not selected in size order (i.e. 1, 2, 3, 4 or 4, 3, 2, 1)?

15. a) In Example 10.13 we determined the probability that in a group of 10 people at least two will have a common birthday. Find the probability that in such a group, at least one other person has the same birthday as yours. (You are included as one of the 10 people.)

b) Referring again to Example 10.13, suppose the group consists of 25 people. Compute the probability that there are at least two people with the same birthday.

c) Referring to part b), if there are 25 people, what is the probability that at least one other has the same birthday as yours?

N.B.: Many "incredible" coincidences are situations in which some event actually has a reasonable probability of occurring to somebody, but a very low probability of occurring to a specific individual. For example, we may know amongst our family and friends more than a dozen birthdays. Based on the calculations above, it is not surprising if one of these matches the birthday of a complete stranger.

16. Three people are eating dinner in a restaurant and ask the waiter for separate checks. The waiter leaves the checks face down in the middle of the table and each person grabs one at random. Using a tree diagram to enumerate all possible outcomes find:

    a) The probability that each person picks his own check.

    b) The probability that no one gets his own check.

    c) The probability that at least one person gets his own check.

17. A coin is tossed five times in succession.

    a) What is the probability that at least one head and at least one tail will appear in the five tosses?

    b) Find the probability that a streak of three or more consecutive heads will appear in the five tosses?

18. Based on public health statistics (from Canada) there is a chance of .024 that a 50-year-old woman will develop breast cancer sometime during the next 10 years of her life.

    a) How many cases of breast cancer would you expect to develop over a ten-year period in a population of 1000 fifty -year old women? In a population of 10,000?

    b) Run the program *coin.xls* with the probability of heads equal to .024. This simulation models the incidence of breast cancer described in a). Run the program 10 times with the number of tosses (i.e. number of women studied) equal to 1000, and 10 times again with the number of tosses equal to 10,000. Record the relative frequency of heads (incidence of breast cancer) in each replicate.

    c) Enter your frequency data from a) (you should have two sets of 10 observations) in a new spreadsheet in two separate labeled columns. Make box plots on the same axes comparing the two data sets. Near what value are the centers of the plots located? Which plot shows the widest spread? Is this what you expected? Why or why not?

19. a) You toss three cubical dice. Describe the equally probable elementary outcomes and determine their number.

    b) Using Table 10.1, which gives information on the sum of two dice, find the probability for each of the possible sums for three dice. Show clearly how each probability is obtained. (Hint: To count the number of outcomes in which the sum of the three dice equals say 10, consider the possible values of one of the die and then use Table 10.1 to count the number of ways of obtaining the remaining two dice.)

    c) Run the program *dice.xls*. In the dialog box choose three dice and run the simulation for 10,000 tosses. Fill in the column marked "Theory" using your answer to b). (You can type a fraction; *Excel* will automatically convert it to decimal.) How would you assess the agreement between your theoretical answer and the empirical relative frequency? Print out the simulation sheet. Make sure to include the theoretical values.

20. The game of "craps" is played with two dice as follows. If the player tossing the dice tosses a sum of 7 or 11, the player wins. If the toss sums to a 2, 3, or 12 the player loses. For any other sum (i.e. 4, 5, 6, 8, 9, or 10) the sum obtained is called the player's "point". Tosses continue until the player throws a sum equal to the point, in which case the player wins, or tosses a sum of 7, in which case the player loses.

   a) Using Table 10.1 determine the probabilities that the first toss results in

      i)   the player winning

      ii) the player losing

      iii) play continuing.

   b) How can you compute the answer to part a)iii) knowing the answers to the first two parts? Explain.

   c) Run the simulation program *craps.xls*, completing 2000 games. Record the estimated probability of winning based on your 2000 games. Based on your result, how much confidence do you place in the statement that the probability of winning at craps is less than 0.5.