Chapter 5

Probability

Probability is a very powerful tool. It starts with the assumption that the world is consistent, that is, that knowledge we have gained by observing some portion of the world can be used to understand the portion we have not directly observed. Probability is the means by which we make inferences about populations from samples. It also allows us to make predictions about individual variates based on a sample. While probability is very important in making archaeological inferences, its abstract nature is better first illustrated through hypothetical examples using coins, dice, and cards. The observant reader will notice that we are reneging on our promise not to rely heavily on such "Vegas-style" discussions. We promise that this is merely a brief interlude. Archaeological illustrations of the application of probability will be presented in subsequent chapters.

Theoretical Determinations of Probability

To begin, let us consider the simple flip of an unbiased coin. We know that we have two possible outcomes: a head or a tail. The specific outcomes of heads or tails are called *events*. As heads and tails are our only possible outcomes, they constitute the finite *probability space*, i.e., a limited number of possible outcomes.

We can symbolize the probability of heads as P(H), and the probability of tails as P(T). The finite probability space is {H,T}. Being part of a world of coin flippers, we all know that we have a one out of two chance of a coin flip resulting in the event "heads" or the event "tails." So, the probability of heads is ½, or expressed as a proportion, .50. All probabilities are expressed in one or the other forms.

Let us express these probabilities with the appropriate symbols: P(H) = ½ =0.50; P(T) = ½ =0.50. Notice that probabilities within the finite sample space can vary between zero and 1, and must sum to one, e.g., P(H) + P(T) = 0.50 + 0.50 = 1.

It is important to note that the probabilities of H and T were derived theoretically without a single flip of a coin. Given our experience however, we realize that this theoretical expectation regarding probabilities will be realized empirically within an expectable range of error.

Let us consider another theoretical example, the case of an event of the outcome of rolling two dice. Each face of each die is designated with one through six dots. There are a total of 36 unique combinations between the two die (e.g., if the first dice is 1, the second dice can be 1 to 6; if the first dice is 2, the second dice can be 1 through 6, etc…) There is only one-way of rolling a sum of two: one on each dice. The probability of rolling a sum of two is therefore 1/36 or .03, which is derived by dividing the number of ways that an even can occur (in this case 1) by the total number of possible outcomes (in this case 36). We can roll a three two ways. The first die may indicate the numeral one, the second die the numeral two, and vice versa. And so on for four, five, etc.

These probabilities are listed on Table 5.1 and can be graphed as in Figure 5.1. Note the perfect symmetry of the distribution of probabilities. Given these probabilities, we can determine the likelihood of any given outcome of a roll of dice.

Empirical Determinations of Probabilities

As much as we would like, we cannot always theoretically determine probabilities. We can, however, substitute probabilities determined empirically. Consider the following hypothetical example.

The University of New Mexico has a severe parking problem. We have noticed that while it is quite easy to find a parking place early in the morning, as the day progresses locating a parking space becomes more difficult, then nearly impossible as noon approaches. By mid-afternoon parking spaces become easier to find and by late afternoon they are abundant until students and faculty begin to arrive for evening classes and open spaces become infrequent.

Assume that we had the patience, the free time, and the desire to conduct the following experiment. For each of the following hours of the day, one of your authors tried ten times on ten different days to successfully obtain a parking space. These results are presented in Table 5.2.

Column1 depicts the time of day, Column 2 the number of successful attempts at obtaining a parking place, Column 3 the number of failures, Column 4 the number of successes divided by the number of trials, Column 5 the number of failures divided by the number of trials, and Column 6 the sums of Column 4 and Column 5.

As empirically derived probabilities are calculated by equation 5.1:

Equation 5.1. P(event) = *f* of an event/number of trials.

Column 4 is the probability of successfully finding a parking place at a given time, while Column 5 is the probability of being unsuccessful. Therefore, one has a .30 probability or 30% chance of successfully finding a parking place between 11-12, as P(parking space 11-12) = 3/10. With empirical cases, larger numbers of trials increases the accuracy of the prediction. One hundred trials would certainly yield a more reliable predictor. The distribution of the probability of a success is depicted in Figure 5.2.

Also, the assumption of continuity (the idea that the world will behave consistently) is sometimes problematic when dealing with empirically derived probabilities. In the preceding case of the theoretically determined probability, the probability of rolling a sum of two on two die will hold so long as the die are unbiased. Assuming that we have unbiased die, we can calculate the probability of rolling any sum from two through twelve, and these probabilities are expected to hold in the past and in the future, a fact upon which the gambling houses in Las Vegas rely (they wouldn't be able to stay in business if your probability of rolling a seven while playing craps changed every day). We do not need to roll the die a hundred times and then compute the actual probability of rolling the sum of two. Instead, we need only ensure that someone has not tampered with them.

In contrast, the use of empirically derived probabilities requires that we create an argument justifying the application of the probability to future (and past) events. For example, we expect that the probability of finding a parking spot on a Thursday during the Fall semester is 0.30, but will this probability hold during Spring Break? Probably not, given that many of UNM's students escape Albuquerque and travel to somewhere that has water for the week. Many UNM's staff members also take vacations during Spring Break. As a result, we expect that the probability of finding a parking spot between 11 and 12 is higher during spring break than is typical of the rest of the semester.

Our empirically derived probability is consequently based on underlying factors that affect the variable of parking space availability. Ideally we could determine what these factors are and adjust our probabilities to take them into account. For example, we could further refine our probability estimates by establishing probabilities for different days of the week, the time of year, the presence of special events such as High School Band competitions, and student enrollment, among other variables. The manifestation of each of these variables could then be factored into our probabilities.

In general, theoretical determinations are preferred, as they are not subject to sampling vagaries, but empirical determinations are common and useful in archaeological contexts. When using probabilities determined empirically, one should be very clear on the sample size and conditions under which the observations were taken. In general, large numbers of observations increase our knowledge of the probability of an event (we will return to this issue in a future chapter), and one's research design must justify the use of the probability in the context in which it is used.

Complex Events

The examples presented above determined probabilities for *simple events*. At times, we seek probabilities for more than one event. Such situations are called *complex events*. To illustrate the difference between complex and simple events, let us consider a deck of playing cards.

We know that each deck contains 52 unique cards (toss the jokers). These cards are divided into four suits: hearts, spades, clubs and diamonds. Each suit contains 13 cards: ace, king, queen, jack, ten, nine, eight, seven, six, five, four, three, and two. We know that the probability of obtaining any unique card is 1/52. For example, the P(ace of spades)=1/52, the P(three of diamonds)=1/52, etc. These events are *simple events*. We also know that the probability of drawing a card of a specific suit is 13/52, as there are 13 cards in each suit. So, the P(heart)=13/52, the P(spade)=13/52, etc. These too are *simple events*.

For a particular game, we might be interested in the probability of drawing a heart *or* a diamond. This constitutes a complex event. The probability of such a complex event is P(heart)+P(diamond), or 13/52 + 13/52=26/52=.50.

This complex event is appropriately symbolized by $P(H \cup D) = 0.50$, and read "probability of H *union* D is equal to .50." This union is simply the sum of independent probabilities. Note that the categories “hearts” and “diamonds” are mutually exclusive. A card cannot be both a heart *and* a diamond. This is illustrated with the Venn diagram in Figure 5.3.

What if our probability spaces overlap? We can also create a Venn Diagram of the possibilities of the *intersection* of events. Figure 5.4 illustrates one such intersection, that of obtaining a heart *or* an ace. The states of being a heart and being an ace are not independent in that the ace of hearts is both. As a result we cannot simply add the probability of hearts (p[H]=13/52) with the probability of aces (p[aces]=4/52) because doing so would count the ace of hearts twice (once as an ace and once as a heart). Yet there is only one ace of hearts in the deck. We must therefore eliminate the probability of the co-occurrence of events.

This set of non-mutually exclusive events is symbolized as:

$$P(H \cup A) = P(H) + P(A) - P(H \cap A) = 13/52 + 4/52 - 1/52 = 16/52$$

where we have to consider the intersection of the two events (H and A): the ace of hearts. If we were simply to sum the probabilities of hearts and aces, we would be including the ace of hearts twice, and our calculation of probability thus in error. Venn diagrams should always be drawn, as they effectively allow us to consider the mutual exclusivity or intersection of events in a graphical form.

This discussion of probability now allows us to proceed to a consideration of the normal distribution, and examples of archaeological inference.

Table 5.1. Probability of rolling a given sum using two die.

Y	P(Y) as a ratio	P(Y) as a probability
2	1/36	.03
3	2/36	.05
4	3/36	.08
5	4/36	.11
6	5/36	.14
7	6/36	.17
8	5/36	.14
9	4/36	.11
10	3/36	.08
11	2/36	.05
12	1/36	.03
	36/36	1.0

Figure 5.1. The distribution of the probabilities of rolling a sum with two die.

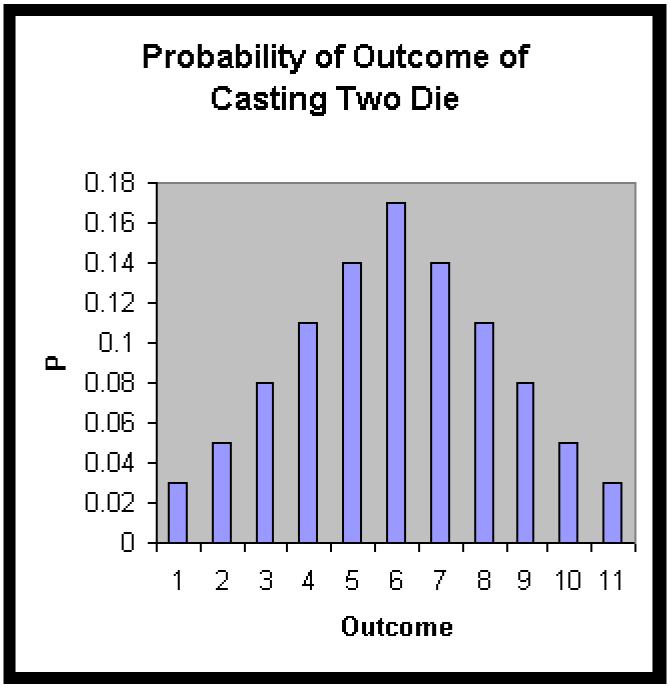

Table 5.2. Probabilities of Success or Failure of finding a parking space at UNM during different times of the Day.

1	2	3	4	5	6
Time of day	S	F	S/n trials, where n trials = 10	F/n trials, where n trials = 10	∑ S/n trials + F/n trials
7-8	10	0	1.0	0	1.0
8-9	8	2	.80	.20	1.0
9-10	7	3	.70	.30	1.0
10-11	2	8	.20	.80	1.0
11-12	3	7	.30	.70	1.0
12-1	0	10	0	1.0	1.0
1-2	1	9	.10	.90	1.0
2-3	2	8	.20	.80	1.0
3-4	4	6	.40	.60	1.0
4-5	7	3	.70	.30	1.0

Figure 5.2. Distribution of the probability of finding a parking spot on the University of New Mexico's overcrowded parking lot.

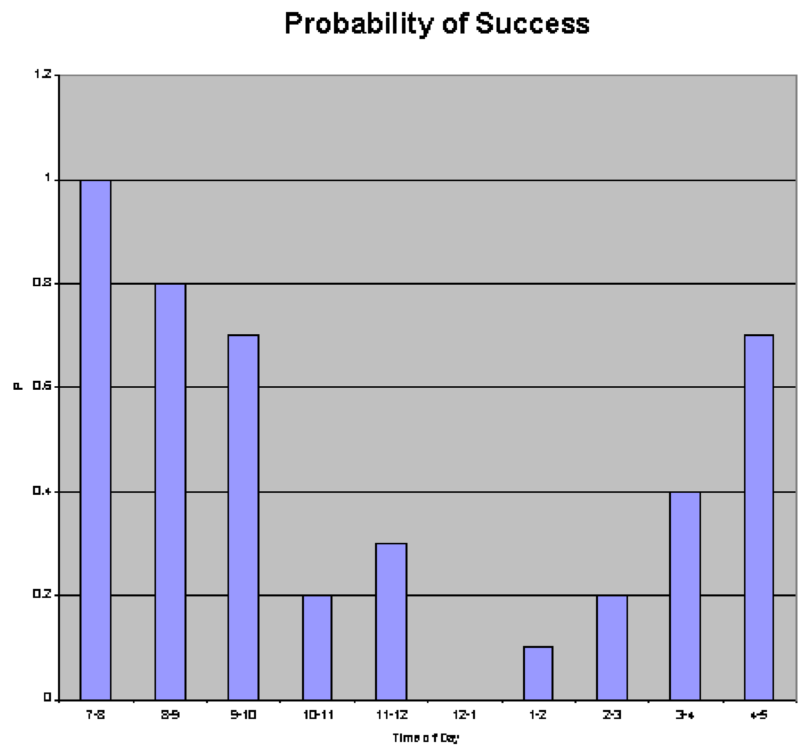

Figure 5.3. Venn Diagram of probabilities of hearts and diamonds.

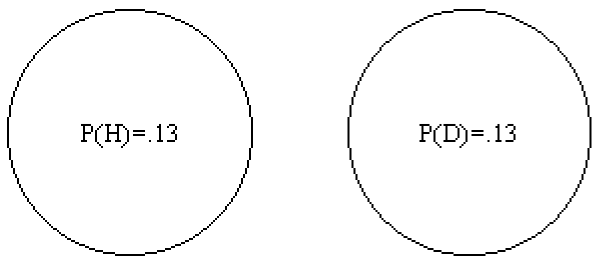

Figure 5.4 Intersection of probabilities of obtaining a heart or an ace.