Chapter 2

Combining Probabilities

But to us probability is the very guide of life. (*Bishop Butler*, 1692–1752)

2.1. Either–or Probability

Let us consider the situation where a die is thrown, and we wish to know the probability that the outcome will be *either* a 1 *or* a 6. How do we find this? First, we consider the six possible outcomes, all of equal probability. Two of these, a 1 and a 6 — one-third of the possible outcomes — satisfy our requirement so the probability of obtaining a 1 or a 6 is

$$p_{1\text{ or }6} = \frac{2}{6} = \frac{1}{3} \tag{2.1}$$

Another way of expressing this result is to say that

$$p_{1\text{ or }6} = p_1 + p_6 = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}. \tag{2.2}$$

In words, Eq. (2.2) says that "the probability of getting either a 1 or a 6 is the sum of the probabilities of getting a 1 and getting a 6".

This idea can be extended so that the probability of getting one of 1, 2, or 3 when throwing the die is

$$p_{1\text{ or }2\text{ or }3} = p_1 + p_2 + p_3 = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}. \tag{2.3}$$

Similarly, the probability of getting either a head or a tail when flipping a coin is

$$p_{h \text{ or } t} = p_h + p_t = \frac{1}{2} + \frac{1}{2} = 1, \tag{2.4}$$

which corresponds to certainty since the only possible outcomes are either a head or a tail.

In considering these combinations of probability we are taking alternative outcomes of a single event — e.g., throwing a die. If we are interested in the outcome being a 1, 2, or 3 then, if we obtain a 1, we exclude the possibility of having obtained either of the other outcomes of interest, a 2 or a 3 (Fig. 2.1). Similarly, if we obtained a 2 the outcomes 1 and 3 would have been excluded. The outcomes for which the probabilities are being combined are *mutually exclusive*. It is a general rule that *the probability of having an outcome which is one or other of a set of mutually exclusive outcomes is the sum of the probabilities for each of them taken separately.*

To explore this idea further consider a standard pack of 52 cards. The probability of choosing a particular card by a random selection is 1/52. Four of the cards are jacks so the probability of picking a jack is

$$p_J = \frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52} = \frac{4}{52} = \frac{1}{13}. \tag{2.5}$$

Now, we want to know the probability of picking a court card (i.e. jack, queen or king) from the pack. The separate probabilities of outcomes for jacks, queens, and kings are all 1/13 but the outcomes

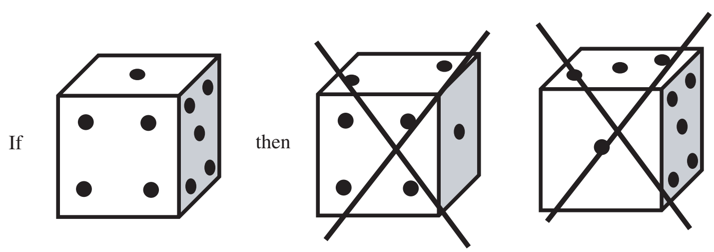

Fig. 2.1. In either–or probability if a 1 is obtained then a 2 or a 3 is excluded.

of obtaining a jack, a queen, or a king are mutually exclusive. Hence the probability of picking a court card is

$$p_{\mathrm{J\ or\ Q\ or\ K}} = p_{\mathrm{J}} + p_{\mathrm{Q}} + p_{\mathrm{K}} = \frac{1}{13} + \frac{1}{13} + \frac{1}{13} = \frac{3}{13}. \qquad (2.6)$$

Of course one could have lumped court cards together as a single category and since there are 12 of them in a pack of 52 the probability of selecting one of them could have been found directly as $\frac{12}{52} = \frac{3}{13}$ but having a mathematically formal way of considering problems is sometimes helpful in less obvious cases.

This kind of combination of probabilities has been called *either–or* and a pedantic interpretation of the English language would give the inference that only two possible outcomes could be involved. However, that is not a mathematical restriction and this type of probability combination can be applied to any number of mutually-exclusive outcomes.

2.2. Both–and Probability

Now, we imagine that two events occur — a coin is spun and a die is thrown. We now ask the question "What is the probability that we get *both* a head *and* a 6?". The two outcomes are certainly not mutually exclusive — indeed they are *independent* outcomes. The result of spinning the coin can have no conceivable influence on the result of throwing the die. First, we list all the outcomes that are possible:

h + 1	h + 2	h + 3	h + 4	h + 5	h + 6 ←
t + 1	t + 2	t + 3	t + 4	t + 5	t + 6

There are 12 possible outcomes, each of equal probability, and we are concerned with the one marked with an arrow. Clearly the probability of having *both* a head *and* a 6 is 1/12. This probability can be considered in two stages. First, we consider the probability of obtaining a head — which is 1/2. This corresponds to the outcomes on the top row of our list. Now, we consider the probability that we also have a 6 that restricts us to one in six of the combinations in the top row since the probability of getting a 6 is 1/6. Looked at in two stages, we see that

$$p_{\text{both h and 6}} = p_{\text{h}} \times p_6 = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}. \tag{2.7}$$

This rule can be extended to find the combined probability of any number of *independent* events. Thus, the combined probability that spinning a coin gives a head, throwing a die gives a 6, and picking a card from a pack gives a jack is

$$p_{\text{h and 6 and J}} = p_{\text{h}} \times p_6 \times p_J = \frac{1}{2} \times \frac{1}{6} \times \frac{1}{13} = \frac{1}{156}. \tag{2.8}$$

Once again, our description of this probability combination has done violence to the English language. The combination *both–and* should formally be applied only to two outcomes but we stretch it to describe the combination of any number of independent outcomes.

These rules of "either–or" and "both–and" combinations of probability can themselves be combined together to solve quite complicated probability problems.

2.3. Genetically Inherited Disease — Just Gene Dependent

Within any population there will exist a number of genetically inherited diseases. There are about 4,000 such diseases known and particular diseases tend to be prevalent in particular ethnic groups. For example, sickle-cell anemia is mainly present in people of West African origin, which will include many of the black populations of the Caribbean and North America and also of the United Kingdom. This disease affects the hemoglobin molecules contained within red blood cells, which are responsible for carrying oxygen from the lungs to muscles in the body and carbon dioxide back from the muscles to the lungs. The hemoglobin forms long rods within the cells, distorting them into a sickle shape and making them less flexible so they flow less easily. In addition, the cells live less time than the normal 120 days for a healthy red cell and so the patient suffers from a constant state of anemia. Another genetic disease is Tay–Sachs disease which affects people of Jewish origin. This attacks the nervous system, destroying brain and nerve cells, and is always fatal, usually at the infant stage.

To understand how genetically transmitted diseases are transmitted, we need to know something about the gene structure of living matter, including humans. Contained within each cell of a human being there are a large number of chromosomes, thread like bodies which contain, strung out along them, large numbers of genes. The number of genes controlling human characteristics is somewhere in the range 30,000–40,000. Each gene is a chain of DNA of length anything from 1,000 to hundreds of thousands of the base units that make up DNA. Genes occur in pairs which usually correspond to contrasting hereditary characteristics. For example, one gene pair might control stature so that gene A predisposes toward height while the other member of the pair, gene a, gives a tendency to produce shorter individuals. Each person has two of these stature genes in his cells. If they are both A then the person will have a tendency to be tall and if they are both a then there will be a tendency to be short. One can only talk about tendency in this instance since other factors influence stature, in particular diet. A child inherits one "stature" gene from each parent and which of the two genes he gets from each parent is purely random. Here, we show some of the possibilities for various parental contributions:

Father	Mother	Child (all pairs of equal probability)			
Aa	Aa	AA	Aa	aA	aa
AA	aa	Aa	Aa	Aa	Aa
AA	AA	AA	AA	AA	AA

where an individual has a contrasting gene pair then sometimes the characteristics will combine, so that, for example, Aa will tend to give a medium-height individual, but in other cases one of the genes may be dominant. Thus, if *B* is a "brown-eye" gene and *b* is a "blue-eye" gene then *BB* gives an individual with brown eyes, *bb* gives an individual with blue eyes, and *Bb* (equivalent to *bB*) will give brown eyes because *B* is the dominant gene. Sometimes genes become "fixed" in a population. All Chinese are *BB* so that all Chinese children must inherit the gene pair *BB* from their parents. Thus, all Chinese have brown eyes.

Now, we consider a genetically related disease linked to the gene pair *D*, *d*. The gene *d* predisposes toward the disease and someone who inherits a pair *dd* will certainly get the disease and die before maturity. However, we take it that *d* is a very rare gene in the community and *D* is dominant. Anyone who happens to be *Dd* will be free of the disease but may pass on the harmful gene *d* to his, or her, children; such a person is known as a *carrier*. Let us suppose that in this particular population the ratio of *d*:*D* is 1:100. What is the probability that, with random mating, i.e., no monitoring of parents, a baby born in that population will have the disease?

We can consider this problem by considering the allocation of the gene pair to the baby one at a time. The probability that the first gene will be *d* is 0.01 because that is the proportion of the *d* gene. The allocation of the second gene of the pair is independent of what the first one happens to be so, again, the probability that this one is *d* is 0.01. Hence the probability that *both* the first gene is *d and* the second gene is *d* is $0.01 \times 0.01 = 0.0001$, or one chance in 10,000. If we were interested in how many babies would be *carriers* of the faulty gene, i.e., possessing the gene pair *Dd*, then, using both–and probability, we note that:

the probability that *both* gene 1 is *D and* gene 2 is *d* is $0.99 \times 0.01 = 0.0099$

the probability that *both* gene 1 is *d and* gene 2 is *D* is $0.01 \times 0.99 = 0.0099$.

Since *Dd* and *dD* are mutually exclusive the probability of the baby carrying the gene pairs *either Dd or dD* is

$$0.0099 + 0.0099 = 0.0198$$

so that about one in 50 babies born is a carrier.

Some genetically related diseases are very rare indeed because the incidence of the flawed gene is low; for $d{:}D = 1{:}1{,}000$ only one child in a million would contract the disease although about one in 500 of them would be carriers of the disease. On the other hand diabetes, which is thought to have a genetically transmitted element, is much more common and the number of carriers is probably quite high.

2.4. Genetically Dependent Disease — Gender Dependent

There are genetically transmitted diseases where the gender of the individual is an important factor. The sex of an individual is determined by two chromosomes, X and Y, a female having the chromosome pair XX and a male XY. The female always contributes the same chromosome to her offspring, X, but the man contributes X or Y with equal probability thus giving a balance between the numbers of males and females in the population. Henry VIII divorced two wives because they did not give him a son but now we know that he was to blame for this! There are some defective genes, for example which leads to hemophilia, which only occurs in the X chromosome. If we call a chromosome with the defective X gene X' then the following situations can occur:

Daughter with chromosome combination XX' will be a carrier of hemophilia but will not have the disease because the presence of X compensates for the X'.

Son with $X'Y$ will have the disease because there is no accompanying X to compensate for the X'.

Now, we can see various outcomes from different parental chromosome compositions

(i)

All children are free of hemophilia and none are carriers

(ii)

The father suffers from the disease but all children are free of hemophilia. Sons are completely unaffected because they just receive a Y chromosome from their fathers. However, all daughters are XX' and so all are carriers.

(iii)

Here, the father is free of the disease but we have a carrier mother. Half the sons are $X'Y$ and so are hemophiliacs. The other half of the sons are XY and so are free of the disease. Half of the daughters are XX' and so are carriers but the other half are XX and are not carriers.

(iv)

Here, we have a hemophiliac father with a carrier mother. Half the sons are $X'Y$and so are hemophiliacs. The other half of the sons are XY and so are free of the disease. Half the daughters are $X'X'$ and so will have the disease while the other half are $X'X$ and so are carriers.

The incidence of hemophilia does not seem to have any strong correlation with ethnicity and about 1 in 5,000 males are born with the disease. The most famous case of a family history of hemophilia is concerned with Queen Victoria who is believed to have been a carrier. She had nine children and 40 grandchildren and several male descendants in Royal houses throughout Europe suffered from the disease. The most notable example was that of Tsarevitch Alexis, the long-awaited male heir to the throne of Russia, born in 1904. A coarse, and rather dissolute, priest, Rasputin, became influential in the Russian court because of his apparent ability to ameliorate the symptoms of the disease in the young Alexis. There are many who believe that his malign influence on the court was an important contributory factor that led to the Russian revolution of 1917.

2.5. A Dice Game — American Craps

American craps is a dice game that is said to originate from Roman times and was probably introduced to America by French colonists. It involves throwing two dice and the progress of the game depends on the sum of the numbers on the two faces. If the thrower gets a sum of either 7 or 11 in his first throw that is called "a natural" and he immediately wins the game. However, if he gets 2 (known as "snake-eyes"), 3, or 12 he immediately loses (Fig. 2.2). Any other sum is the players "point." The player then continues throwing until either he gets his "point" again, in which case he wins, or until he gets a 7, in which case he loses. The probabilities of obtaining various sums are clearly the essence of this game.

(a) (b)

Fig. 2.2. (a) Snake-eyes (b) A *natural* (first throw) but losing throw later.

First, we consider the probability of obtaining 2, 3, or 12. The two dice must show one of

$$1+1 \quad 1+2 \quad 2+1 \quad 6+6$$

and these are mutually exclusive combinations. Since the numbers on each of the dice are independent of each other, each combination has a probability of 1/36, e.g.

$$p_{\text{snake-eyes}} = p_1 \times p_1 = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}. \tag{2.9}$$

Hence the probability of getting 2, 3, or 12 is

$$p_{2,\text{ 3 or 12}} = p_{1+1} + p_{1+2} + p_{2+1} + p_{6+6} = \frac{4}{36} = \frac{1}{9}. \tag{2.10}$$

Getting 7 or 11, as a sum requires one or other of the combinations

$$6+1 \quad 5+2 \quad 4+3 \quad 3+4 \quad 2+5 \quad 1+6 \quad 6+5 \quad 5+6.$$

These combinations are mutually exclusive, each with a probability of 1/36 so that the probability of a "natural" is 8/36 = 2/9.

If the player has to play to his "point" then his chances of winning depends on the value of that "point" and how the probability of obtaining it compares with the probability of obtaining 7. Below, we show for each possible "point," the combinations that can give it and the probability of achieving it per throw.

Point	Combinations					Probability
4	3 + 1	2 + 2	1 + 3			$\frac{3}{36} = \frac{1}{12}$
5	4 + 1	3 + 2	2 + 3	1 + 4		$\frac{4}{36} = \frac{1}{9}$
6	5 + 1	4 + 2	3 + 3	2 + 4	1 + 5	$\frac{5}{36}$
8	6 + 2	5 + 3	4 + 4	3 + 5	2 + 6	$\frac{5}{36}$
9	6 + 3	5 + 4	4 + 5	3 + 6		$\frac{4}{36} = \frac{1}{9}$
10	6 + 4	5 + 5	4 + 6			$\frac{3}{36} = \frac{1}{12}$

The total of 7 can be achieved in six possible ways (see above ways of obtaining a "natural") and its probability of attainment, 1/6, is higher than that of any of the points.

This is an extremely popular gambling game in the United States and those that play it almost certainly acquire a good instinctive feel for probability theory as it affects this game, although few would be able to express their instinctive knowledge in a formal mathematical way. Later, when we know a little more about probabilities, we shall calculate the overall chance of the thrower winning — it turns out that the odds are very slightly against him.

Problems 2

2.1. For the dodecahedron described in problem 1.2 what is the probability that a throw will give a number less than 6?

2.2. If a card is drawn from a pack of cards then what is the probability that it is an ace?

2.3. The dodecahedron from problem 1.2 and a normal six-sided die are thrown together. What is the probability that both the dodecahedron and the die will give a 5?

2.4. Find the combined probability of getting a jack from a pack of cards and getting four heads from spinning four coins. Is this probability greater than or less than that found in problem 2.3?

2.5. The dodecahedron from problem 1.2 and a normal six-sided die are thrown together. What is the probability that the sum of the numbers on the sides is 6?
(*Hint*: Consider the number of ways that a total of six can be obtained and the probability of each of those ways.)

2.6. For a particular gene pair F and f, the latter gene leads to a particular disease although F is the dominant gene. If the ratio of the incidence of the genes is $f{:}F = 1{:}40$ then what proportion of the population is expected to be

(i) affected by the disease, and
(ii) carriers of the disease?