Chapter 2

Conditional Probability

2.1 Independence

Intuitively, two events A and B are independent if the occurrence of A has no influence on the probability of occurrence of B. The formal definition is: A and B are **independent** if

$$P(A \cap B) = P(A)P(B)$$

We now give three classic examples of independent events. In each case it should be clear that the intuitive definition is satisfied, so we will only check the conditions of the formal one.

- Flip two coins. $A =$ "The first coin shows Heads," $B =$ "The second coin shows Heads." $P(A) = 1/2$, $P(B) = 1/2$, $P(A \cap B) = 1/4$.
- Roll two dice. $A =$ "The first die shows 5," $B =$ "The second die shows 2." $P(A) = 1/6$, $P(B) = 1/6$, $P(A \cap B) = 1/36$.
- Pick a card from a deck of 52. $A =$ "The card is an ace," $B =$ "The card is a spade." $P(A) = 1/13$, $P(B) = 1/4$, $P(A \cap B) = 1/52$.

Two examples of events that are not independent are

Example 2.1. *Draw two cards from a deck. $A =$ "The first card is a spade," $B =$ "The second card is a spade." $P(A) = 1/4$, $P(B) = 1/4$, but*

$$P(A \cap B) = \frac{C_{13,2}}{C_{52,2}} = \frac{13 \cdot 12}{52 \cdot 51} < \left(\frac{1}{4}\right)^2$$

Intuitively, these two events are not independent, since getting a spade the first time reduces the fraction of spades in the deck and makes it harder to get a spade the second time. Anticipating a result in the next section, note that we have a probability of $13/52$ of getting a spade the first time and, if we succeed, only a $12/51$ chance the second time.

Example 2.2. *Roll two dice. $A =$ "The sum of the two dice is 9," $B =$ "The first die is 2." $A = \{(6,3), (5,4), (4,5), (3,6)\}$, so $P(A) = 4/36$. $P(B) = 1/6$, but $P(A \cap B) = 0$ since $(2,7)$ is impossible.*

In general if A and B are disjoint events that have positive probability, they are not independent since $P(A)P(B) > 0 = P(A \cap B)$.

There are two ways of extending the definition of independence to more than two events. $A_1, \ldots, A_n$ are said to be **pairwise independent** if for each $i \neq j$, $P(A_i \cap A_j) = P(A_i)P(A_j)$, that is, each pair is independent. $A_1, \ldots, A_n$ are said to be **independent** if for any $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ we have

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$$

If we flip n coins and let $A_i =$ "The ith coin shows Heads," then the A_i are independent since $P(A_i) = 1/2$ and $P(A_{i_1} \cap \ldots \cap A_{i_k}) = 1/2^k$. We have already seen an example of events that are pairwise independent but not independent:

Example 2.3 (Birthdays). *Let $A =$ "Alice and Betty have the same birthday" $B =$ "Betty and Carol have the same birthday," $C =$ "Carol and Alice have the same birthday." Each pair of events is independent but the three are not.*

Since there are 365 ways two girls can have the same birthday out of 365^2 possibilities (as in Example 1.6, we are assuming that leap year does not exist and that all the birthdays are equally likely), $P(A) = P(B) = P(C) = 1/365$. Likewise, there are 365 ways all three girls can have the same birthday out of 365^3 possibilities, so

$$P(A \cap B) = \frac{1}{365^2} = P(A)P(B)$$

i.e., A and B are independent. Similarly, B and C, are independent and C and A are independent, so A, B, and C are pairwise independent. The three events A, B, and C are not independent, however, since $A \cap B = A \cap B \cap C$ and hence

$$P(A \cap B \cap C) = \frac{1}{365^2} \neq \left(\frac{1}{365}\right)^3 = P(A)P(B)P(C)$$

The last example is somewhat unusual. However, the moral of the story is that to show several events are independent, you have to check more than just that each pair is independent.

Example 2.4. *Suppose we roll 6 dice. What is the probability of $A =$ "We get exactly two 4's"?*

One way that A can occur is

$$\frac{\times}{1}\,\frac{4}{2}\,\frac{\times}{3}\,\frac{4}{4}\,\frac{\times}{5}\,\frac{\times}{6}$$

where × stands for “not a 4.” Since the six events “die one shows ×,” “die two shows 4,” ..., “die six shows ×” are independent, the indicated pattern has probability

$$\frac{5}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^4$$

Here we have been careful to say “pattern” rather than “outcome” since the given pattern corresponds to 5^4 outcomes in the sample space of 6^6 possible outcomes for 6 dice. Each pattern that results in A corresponds to a choice of 2 of the 6 trials on which a 4 will occur, so the number of patterns is $C_{6,2}$. When we write out the probability of each pattern there will be two 1/6’s and four 5/6’s so each pattern has the same probability and

$$P(A) = \binom{6}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^4$$

The binomial distribution. Generalizing from the last example, suppose we perform an experiment n times and on each trial an event we call “success” has probability p. (Here and in what follows, when we repeat an experiment, we assume that the outcomes of the various trials are independent.) Then the probability of k successes is

$$\binom{n}{k} p^k (1-p)^{n-k} \tag{2.1}$$

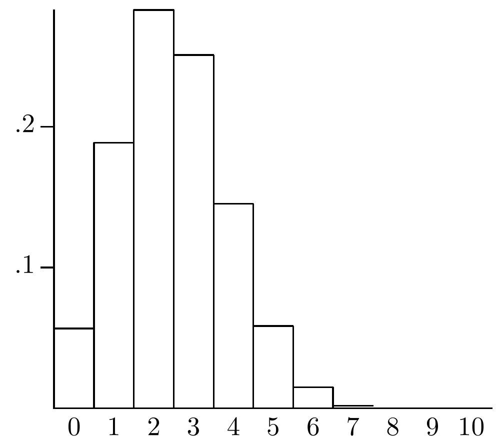
Taking $n = 6$, $k = 2$, and $p = 1/6$ in (2.1) gives the answer in the previous example. The reasoning for the general formula is similar. There are $C_{n,k}$ ways of picking k of the n trials for successes to occur, and each pattern of k successes and $n - k$ failures has probability $p^k(1-p)^{n-k}$.

Example 2.5. *A student takes a test with 10 multiple-choice questions. Since she has never been to class she has to choose at random from the 4 possible answers. What is the probability she will get exactly 3 right?*

The number of trials is $n = 10$. Since she is guessing the probability of success $p = 1/4$, so using (2.1) the probability of $k = 3$ successes and $n - k = 7$ failures is

$$C_{10,3}(1/4)^3(3/4)^7 = \frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3} \frac{3^7}{4^{10}} = 120 \frac{2187}{1,048,576} = 0.250$$

In the same way we can compute the other probabilities. The results are given in the next graph.

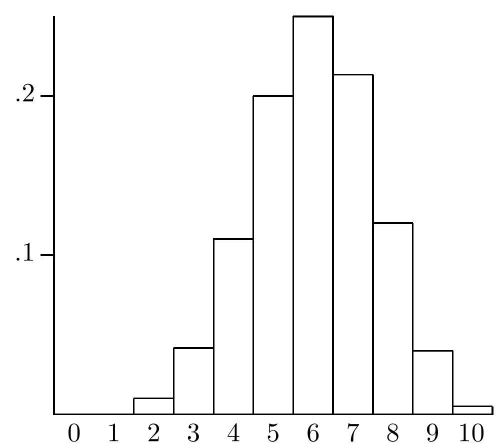


Example 2.6. *A football team wins each week with probability 0.6 and loses with probability 0.4. If we suppose that the outcomes of their 10 games are independent, what is the probability they will win exactly 8 games?*

The number of trials is $n = 10$. We are told that the probability of success $p = 0.6$, so using (2.1) the probability of $k = 8$ successes and $n - k = 2$ failures is

$$C_{10,8}(0.6)^8(0.4)^2 = \frac{10 \cdot 9}{1 \cdot 2}(0.6)^8(0.4)^2 = 0.1209$$

In the same way we can compute the other probabilities. The results are given in the next graph.



Example 2.7 (Aces at Bridge). *When we draw 13 cards out of a deck of 52, each ace has a probability 1/4 of being chosen, but the four events are not*

independent. How does the probability of $k = 0, 1, 2, 3, 4$ *aces compare with that of the binomial distribution with* $n = 4$ *and* $p = 1/4$*?*

We first consider the probability of drawing two aces:

$$\frac{C_{4,2}C_{48,11}}{C_{52,13}} = \frac{6\frac{48\cdots 38}{11!}}{\frac{52\cdots 40}{13!}} = 6 \cdot \frac{13 \cdot 12 \cdot 39 \cdot 38}{52 \cdot 51 \cdot 50 \cdot 49} = .2135$$

In contrast the probability for the binomial is

$$C_{4,2}(1/4)^2(3/4)^2 = 0.2109$$

To compare the two formulas note that $13/52 = 1/4$, $12/51 = 0.2352$, $39/50 = 0.78$, $38/51 = 0.745$ versus $(1/4)^2(3/4)^2$ in the binomial formula. Similar computations show that if $D = 52 \cdot 51 \cdot 50 \cdot 49$

	aces	binomial
0	$39 \cdot 38 \cdot 37 \cdot 36/D$	$(3/4)^4$
1	$4 \cdot 13 \cdot 39 \cdot 38 \cdot 37/D$	$4(1/4)(3/4)^3$
2	$6 \cdot 13 \cdot 12 \cdot 39 \cdot 38/D$	$6(1/4)^2(3/4)^2$
3	$4 \cdot 13 \cdot 12 \cdot 11 \cdot 39/D$	$4(1/4)^3(3/4)$
4	$13 \cdot 12 \cdot 11 \cdot 10/D$	$(1/4)^4$

Evaluating these expressions leads to the following probabilities:

	aces	binomial
0	0.3038	0.3164
1	0.4388	0.4218
2	0.2134	0.2109
3	0.0412	0.0468
4	0.00264	0.00390

Example 2.8. *In 8 games of bridge, Harry had 6 hands without an ace. Should he doubt that the cards are being shuffled properly?*

The number of hands with no ace has a binomial distribution with $n = 8$ and $p = 0.3038$. The probability of at least 6 hands without an ace is

$$\sum_{k=6}^{8} C_{8,k}(0.3038)^k(0.6962)^{8-k} = 1 - \sum_{k=0}^{5} C_{8,k}(0.3038)^k(0.6962)^{8-k}$$

We have turned the probability around because on the TI-83 calculator the sum can be evaauted as binomcdf$(8, 0.3038, 5) = 0.9879$. Thus the probability of luck this bad is 0.0121.

The multinomial distribution. The arguments above generalize easily to independent events with more than two possible outcomes. We begin with an example.

Example 2.9. *Consider a die with 1 painted on three sides, 2 painted on two sides, and 3 painted on one side. If we roll this die ten times what is the probability we get five 1's, three 2's and two 3's?*

The answer is

$$\frac{10!}{5!\,3!\,2!}\left(\frac{1}{2}\right)^5\left(\frac{1}{3}\right)^3\left(\frac{1}{6}\right)^2$$

The first factor, by (1.10), gives the number of ways to pick five rolls for 1's, three rolls for 2's, and two rolls for 3's. The second factor gives the probability of any outcome with five 1's, three 2's, and two 3's. Generalizing from this example, we see that if we have k possible outcomes for our experiment with probabilities $p_1, \ldots, p_k$ then the probability of getting exactly n_i outcomes of type i in $n = n_1 + \cdots + n_k$ trials is

$$\frac{n!}{n_1! \cdots n_k!}\, p_1^{n_1} \cdots p_k^{n_k} \tag{2.2}$$

since the first factor gives the number of outcomes and the second the probability of each one.

Example 2.10. *A baseball player gets a hit with probability 0.3, a walk with probability 0.1, and an out with probability 0.6. If he bats 4 times during a game and we assume that the outcomes are independent, what is the probability he will get 1 hit, 1 walk, and 2 outs?*

The total number of trials $n = 4$. There are $k = 3$ categories hit, walk, and out. $n_1 = 1$, $n_2 = 1$, and $n_3 = 2$. Plugging in to our formula the answer is

$$\frac{4!}{1!1!2!}(0.3)(0.1)(0.6)^2 = 0.1296$$

Example 2.11. *The output of a machine is graded excellent 70% of the time, good 20% of the time, and defective 10% of the time. What is the probability a sample of size 15 has 10 excellent, 3 good, and 2 defective items?*

The total number of trials $n = 15$. There are $k = 3$ categories excellent, good, and defective. $n_1 = 10$, $n_2 = 3$, and $n_3 = 2$. Plugging in to our formula the answer is

$$\frac{15!}{10!\,3!\,2!} \cdot (0.7)^{10}(0.2)^3(0.1)^2$$

Genetics (Hardy-Weinberg equilibrium). Most animals and plants are diploid organisms: each cell has two copies of each chromosome, with the exception of the chromosome that determines the individual's sex. In this case, a female has two copies of the X chromosome and a male has one X and one Y. When reproduction occurs, a special cell division process called *meiosis* produces reproductive cells called *gametes* that have one copy of each chromosome. Two gametes are then combined to produce one new individual.

Each hereditary characteristic is carried by a pair of genes, one on each chromosome, so the new offspring gets one gene from its mother and one from its father. We will consider the situation in which each gene can take only two forms, called *alleles*, which we will denote by a and A. An example from the pioneering work of the Czech monk Gregor Mendel is $A =$ "smooth skin" and $a =$ "wrinkled skin" for the pea plants that he used for much of his experimental work. In this case A is *dominant* over a, meaning that Aa individuals (those with one A and one a) will have smooth skin.

Let us start from an idealized infinite population in which individuals are found in the following proportions, where the proportions are nonnegative and sum to 1:

$$\begin{array}{ccc} AA & Aa & aa \\ \alpha_0 & \beta_0 & \gamma_0 \end{array}$$

If we assume that random mating occurs then each new individual picks two parents at random from the population and picks an allele at random from the two carried by each parent. To compute the proportions of the three types in the first generation of offspring, note that (i) since the first allele is picked at random from the population it will be A with probability

$$p_1 = \alpha_0 + (\beta_0/2)$$

and a with probability $1 - p_1 = \gamma_0 + (\beta_0/2)$, and (ii) the second allele will be independent and have the same distribution, so the proportions in the first generation of offspring will be

$$\alpha_1 = p_1^2 \qquad \beta_1 = 2p_1(1-p_1) \qquad \gamma_1 = (1-p_1)^2$$

Something quite remarkable happens when we use these values to compute the fractions in the second generation of offspring. An allele picked at random from the first generation will be A with probability

$$\begin{aligned} p_2 &= \alpha_1 + (\beta_1)/2 \\ &= p_1^2 + 2p_1(1-p_1)/2 = p_1(p_1 + 1 - p_1) = p_1 \end{aligned} \tag{2.3}$$

so the proportions in the second generation of offspring will be

$$\begin{aligned} \alpha_2 &= p_2^2 = p_1^2 = \alpha_1 \\ \beta_2 &= 2p_2(1-p_2) = 2p_1(1-p_1) = \beta_1 \\ \gamma_2 &= (1-p_2)^2 = (1-p_1)^2 = \gamma_1 \end{aligned} \tag{2.4}$$

Since the proportions of AA, Aa, and aa alleles reach equilibrium in one generation of offspring starting from an arbitrary distribution, it follows that if the fraction of A alleles in the population is p then the proportions of the genotypes will be

$$AA \qquad Aa \qquad aa$$

$$p^2 \qquad 2p(1-p) \qquad (1-p)^2$$

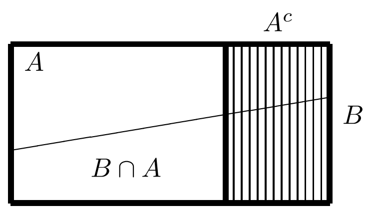
The last result is called the *Hardy-Weinberg Theorem*. To illustrate its use suppose that in a population of pea plants, 91% have smooth skin (AA or Aa) and 9% have wrinkled skin (aa). Since the fractions of AA, Aa, and aa individuals are p^2, $2p(1-p)$, and $(1-p)^2$ and only aa individuals have wrinkled skin, we can infer that $(1-p) = 0.3$ and the three proportions must be 0.49, 0.42, and 0.09.

2.2 Conditional Probability

Suppose we are told that the event A with $P(A) > 0$ occurs. Then the sample space is reduced from Ω to A and the probability that B will occur given that A has occurred is

$$P(B|A) = P(B \cap A)/P(A) \tag{2.5}$$

To explain this formula, note that (i) only the part of B that lies in A can possibly occur, and (ii) since the sample space is now A, we have to divide by $P(A)$ to make $P(A|A) = 1$.



Some examples should help to clarify the definition.

Example 2.12. *Suppose A and B are independent. In this case* $P(A \cap B) = P(A)P(B)$ *so*

$$P(B|A) = \frac{P(A)P(B)}{P(A)} = P(B)$$

Using the words of the intuitive definition of independence, "the occurrence of A has no influence on the probability of the occurrence of B."

Example 2.13. *Suppose we roll two dice and* A = *"The sum is 8,"* B = *"The first die is 3."* $A = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$, *so* $P(A) = 5/36$. $A \cap B = \{(3,5)\}$, *so*

$$P(B|A) = \frac{1/36}{5/36} = \frac{1}{5}$$

The same result holds if B = "The first die is k" and $2 \le k \le 6$. Carrying this reasoning further, we see that given the outcome lies in A, all five possibilities have the same probability. This should not be surprising. The original probability is uniform over the 36 possibilities, so when we condition on the occurrence of A, its five outcomes are equally likely.

As the last example may have suggested, the mapping $B \to P(B|A)$ is a probability. That is, it is a way of assigning numbers to events that satisfies the axioms introduced in Chapter 1. To prove this, we note that

(i) $0 \le P(B|A) \le 1$ since $0 \le P(B \cap A) \le P(A)$.

(ii) $P(\Omega|A) = P(\Omega \cap A)/P(A) = 1$

(iii) and (iv). If B_i are disjoint then $B_i \cap A$ are disjoint and $(\cup_i B_i) \cap A = \cup_i (B_i \cap A)$, so using the definition of conditional probability and parts (iii) and (iv) of the definition of probability we have

$$P(\cup_i B_i|A) = \frac{P(\cup_i (B_i \cap A))}{P(A)} = \frac{\sum_i P(B_i \cap A)}{P(A)} = \sum_i P(B_i|A)$$

From the last observation it follows that $P(\cdot|A)$ has the same properties that ordinary probabilities do, for example

$$P(B^c|A) = 1 - P(B|A) \tag{2.6}$$

Example 2.14. *A person picks 13 cards out of a deck of 52. Let A_1 = "He receives at least one Ace," H = "He has the Ace of hearts," and A_2 = "He receives at least two Aces." Since all Aces are alike, it may at first be surprising that the probability he has two Aces given that he has the Ace of hearts is larger than the probability he has two Aces given that he has one Ace, but this is true.*

Let E_1 = "He has exactly one Ace." Since $E_1 \cap A_2 = \emptyset$ and $E_1 \cup A_2 \supset A_1$ we have

$$1 = \frac{P(E_1 \cap A_1) + P(A_2 \cap A_1)}{P(A_1)} = P(E_1|A_1) + P(A_2|A_1)$$

From this and similar reasoning for H we get

$$P(A_2|A_1) = 1 - P(E_1|A_1) \qquad P(A_2|H) = 1 - P(E_1|H)$$

Let E_0 = "He has no Ace."

$$p_0 = P(E_0) = \frac{C_{48,13}}{C_{52,13}} \qquad p_1 = P(E_1) = \frac{4C_{48,12}}{C_{52,13}}$$

Since $E_1 \subset A_1$ and $A_1 = E_0^c$, $P(E_1|A_1) = P(E_1)/P(A_1) = p_1/(1-p_0)$. On the other hand,

$$P(E_1|B) = \frac{P(E_1 \cap H)}{P(H)} = \frac{C_{48,12}/C_{52,13}}{1/4} = p_1$$

so $P(E_1|A_1) = p_1/(1-p_0) > p_1 = P(E_1|H)$ as claimed. The intuitive explanation of the last result is that it is harder to get the Ace of hearts than to get at least one Ace, so conditioning on having the Ace of hearts gives you a better chance of having two or more Aces than conditioning on having at least one Ace.

Multiplying the definition of conditional probability in (2.1) on each side by $P(A)$ gives the **multiplication rule**

$$P(A)P(B|A) = P(B \cap A) \tag{2.7}$$

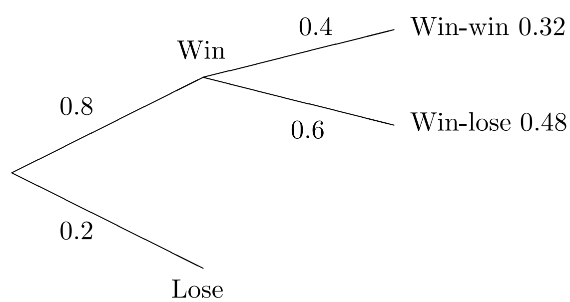
Example 2.15. *Suppose we draw two cards out of a deck of 52. What is the probability both cards are spades?*

Let A = "The first card is a spade," B = "The second card is a spade." $P(A) = 1/13$. To compute $P(B|A)$ we note that if A has occurred then only 12 of the remaining 51 cards are spades, so $P(B|A) = 12/51$ and

$$P(A \cap B) = P(A)P(B|A) = \frac{13}{52} \cdot \frac{12}{51}$$

Note that in this example we computed $P(B|A)$ by thinking about the situation that exists after A has occurred, rather than using the definition $P(B|A) = P(A \cap B)/P(A)$. Indeed, it is more common to use $P(A)$ and $P(B|A)$ to compute $P(A \cap B)$ than to use $P(A)$ and $P(A \cap B)$ to compute $P(B|A)$.

Example 2.16. *The Cornell hockey team is playing in a four team tournament. In the first round they have any easy opponent that they will beat 80% of the time but if they win that game they will play against a tougher team where their probability of success is 0.4. What is the probability that they will win the tournament?*



If A and B are the events of victory in the first and second games then $P(A) = 0.8$ and $P(B|A) = 0.4$, so the probability that they will win the tournament is

$$P(A \cap B) = P(A)P(B|A) = 0.8(0.4) = 0.32$$

The reasoning in the last two examples extends easily to three events:

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)$$

since the right-hand side is equal to

$$P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)}$$

Example 2.17. *In the town of Mythica 90% of students graduate high school, 60% of high school graduates complete college, and 20% of college graduates get graduate or professional degrees. What fraction of students get advanced degrees?*

Answer $= (0.9)(0.6)(0.2) = 0.108$.

The formula for three events generalizes to any number of events.

Example 2.18. *What is the probability of a flush, i.e., all cards of the same suit when we draw 5 cards out of a deck of 52?*

$$1 \cdot \frac{12}{51} \cdot \frac{11}{50} \cdot \frac{10}{49} \cdot \frac{9}{48}$$

The first time we can draw anything. On the second draw we must pick one of the other 12 cards in that suit among the 51 that remain. If we succeed on the second draw then there are 11 good cards out of 50, etc.

Conditional probabilities are the sources of many "paradoxes" in probability. One of these attracted worldwide attention in 1990 when Marilyn vos Savant discussed it in her weekly column in the Sunday *Parade* magazine.

Example 2.19 (The Monty Hall problem). *The problem is named for the host of the television show Let's Make A Deal in which contestants were often placed in situations like the following: Three curtains are numbered 1, 2, and 3. Behind one curtain is a car; behind the other two curtains are donkeys. You pick a curtain, say #1. To build some suspense the host opens up one of the two remaining curtains, say #3, to reveal a donkey. What is the probability you will win given that there is a donkey behind #3? Should you switch curtains and pick #2 if you are given the chance?*

Many people argue that "the two unopened curtains are the same so they each will contain the car with probability 1/2, and hence there is no point in switching." As we will now show, this naive reasoning is incorrect. To compute the answer we have to make an assumption about how the host behaves. Suppose that he always chooses to show you a donkey and picks at random if there are two unchosen curtains with donkeys. Assuming you pick curtain #1, there are three possibilities

	#1	#2	#3	host's action
case 1	donkey	donkey	car	opens #2
case 2	donkey	car	donkey	opens #3
case 3	car	donkey	donkey	opens #2 or #3

Now $P(\text{case 2}, \text{open door \#3}) = 1/3$ and

$$P(\text{case 3}, \text{open door \#3}) = P(\text{case 3})P(\text{open door \#3}|\text{case 3}) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

Adding the two ways door #3 can be opened gives $P(\text{open door \#3}) = 1/2$ and it follows that

$$P(\text{case 3}|\text{open door \#3}) = \frac{P(\text{case 3}, \text{open door \#3})}{P(\text{open door \#3})} = \frac{1/6}{1/2} = \frac{1}{3}$$

Although it took a number of steps to compute this answer, it is "obvious." When we picked one of the three doors initially we had probability 1/3 of picking the car, and since the host can always open a door with a donkey the new information does not change our chance of winning.

The paradox actually predates the game show in the following form. Three prisoners, Al, Bob, and Charlie, are in a cell. At dawn two will be set free and one will be hanged, but they do not know who will be chosen. The guard offers to tell Al the name of one of the other two prisoners who will go free but Al stops him, screaming, "No, don't! That would increase my chances of being hanged to 1/2." Criticize Al's reasoning.

People vs. Collins. In 1964 the purse of an elderly woman shopping in Los Angeles was snatched by a young white female with a blond ponytail. The theif fled on foot but soon after was seen getting into a yellow car driven by a black man who had a mustache and beard. A police investigation subsequently turned up a suspect who was blond, wore a ponytail, and lived with a black man who had a mustache, beard, and a yellow car. None of the eyewitnesses were able to identify the suspects so the police turned to probability

Characteristic	**Probability**
Yellow car	0.1
Man with mustache	0.25
Black man with beard	0.01
Woman with ponytail	0.1
Woman with blond hair	0.33
Interracial couple in car	0.01

Multiplying the probabilities as if they were independent gives an overall probability of $1/12,000,000$. This impressed the jury who handed down a verdict of second degree robbery. The Supreme Court of California later disagreed and reversed the conclusion based on the fact that if there were 1,000,000 couples in Los Angeles the probability of no one else with these six characterisitics was

$$\left(1 - \frac{1}{12,000,000}\right)^{1,000,000}$$

This doesn't work well in a calculator, so to evaluate the probability we have to rexpress it as

$$\exp(1,000,000 \ln(1 - 1/12,000,000)) = 0.92$$

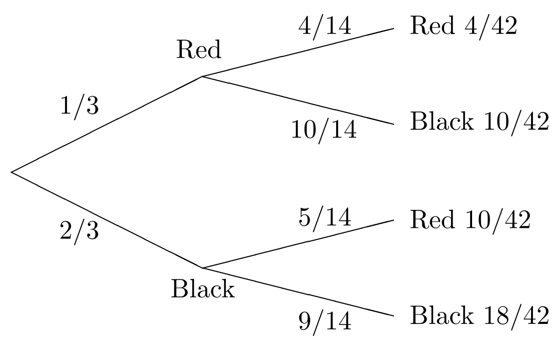
Thus, even if one accepts the dubious assumption of independence, there is an 8% chance of another couple with these characteristics.

2.3 Two-Stage Experiments

We begin with several examples and then describe the collection of problems we will treat in this section.

Example 2.20. *An urn contains 5 red and 10 black balls. We draw two balls from the urn without replacement. What is the probability that the second ball drawn is black?*

This is easy to see if we draw a picture. The first split in the tree is based on the outcome of the first draw and the second on the outcome of the final. The outcome of the first draw dictates the probabilities for the second one. We multiply the probabilities on the edges to get probabilities of the four endpoints, and then sum the ones that correspond to Red to get the answer: $4/42+10/42 = 1/3$



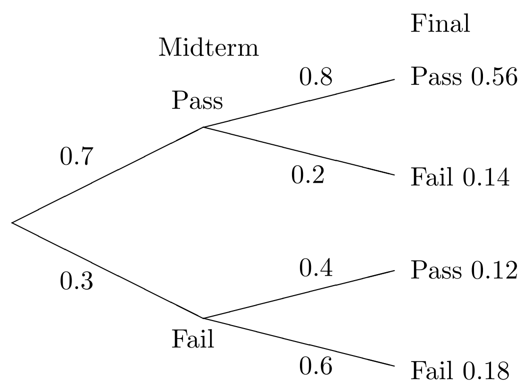
To do this with formulas, let R_i be the event of a red ball on the ith draw and let B_1 be the event of a black ball on the first draw. Breaking things down according to the outcome of the first test, then using the multiplication rule.

$$\begin{aligned} P(R_2) &= P(R_2 \cap R_1) + P(R_2 \cap B_1) \\ &= P(R_2|R_1)P(R_1) + P(R_2|B_1)P(B_1) \\ &= (1/3)(4/14) + (2/3)(5/14) = 14/42 = 1/3 \end{aligned}$$

From this we see that $P(R_2|R_1) < P(R_1) < P(R_2|B_1)$ but the two probabilities average to give $P(R_1)$. This calculation makes the result look like a miracle but it is not. If we number the 15 balls in the urn, then by symmetry each of them is equally likely to be the second ball chosen. Thus the probability of a red on the second, eighth, or fifteenth draw is always the same.

Example 2.21. *Based on past experience, 70% of students in a certain course pass the midterm exam. The final exam is passed by 80% of those who passed the midterm, but only by 40% of those who fail the midterm. What fraction of students pass the final.*

Drawing a tree as before with the first split based on the outcome of the midterm and the second on the outcome of the final, we get the answer: $0.56 + 0.12 = 0.68$



To do this with formulas, let A be the event that the student passes the final and let B be the event that the student passes the midterm. Breaking things down according to the outcome of the first test, then using the multiplication rule.

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \\ &= P(A|B)P(B) + P(A|B^c)P(B^c) \\ &= (0.8)(0.7) + (0.4)(0.3) = 0.68 \end{aligned}$$

Example 2.22. *Al flips 3 coins and Betty flips 2. Al wins if the number of Heads he gets is more than the number Betty gets. What is the probability Al will win?*

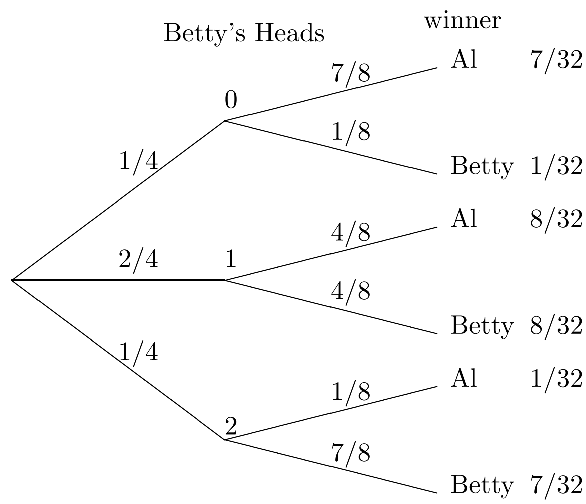
Let A be the event that Al wins. We will break things down according to the number of heads Betty gets. Let B_i be the event that Betty gets i Heads, and let C_j be the event that Al gets j Heads. By considering the four outcomes of flipping two coins it is easy to see that

$$P(B_0) = 1/4 \qquad P(B_1) = 1/2 \qquad P(B_2) = 1/4$$

while considering the eight outcomes for three coins leads to

$$\begin{aligned} &P(A|B_0) = P(C_1 \cup C_2 \cup C_3) = 7/8 \\ &P(A|B_1) = P(C_2 \cup C_3) = 4/8 \\ &P(A|B_2) = P(C_3) = 1/8 \end{aligned}$$

This gives us the raw material for drawing our picture



Adding up the ways Al can win we get $7/32 + 8/32 + 1/32 = 1/2$. To check this draw a line through the middle of the picture note the symmetry between top and bottom.

To do this with formulas, note that $A \cap B_i$, $i = 0, 1, 2$ are disjoint and their union is A, so

$$P(A) = \sum_{i=0}^{2} P(A \cap B_i) = \sum_{i=0}^{2} P(A|B_i)P(B_i)$$

since $P(A \cap B_i) = P(A|B_i)P(B_i)$ by the multiplication rule (2.7). Plugging in the values we computed,

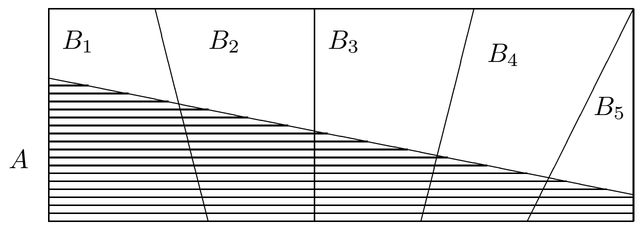
$$P(A) = \frac{1}{4} \cdot \frac{7}{8} + \frac{2}{4} \cdot \frac{4}{8} + \frac{1}{4} \cdot \frac{1}{8} = \frac{7+8+1}{32} = \frac{1}{2}$$

The previous analysis makes it look miraculous that we have a fair game. However it is true in general.

Example 2.23. *Al flips $n+1$ coins and Betty flips n. Al wins if the number of Heads he gets is more than the number Betty gets. What is the probability Al will win?*

Consider the situation after Al has flipped n coins and Betty has flipped n. There are the three possibilities $A > B$, $A = B$, $A < B$, i.e., Al can have more heads, the same number of heads, or fewer heads than Betty. In the first case Al has already won. In the third he cannot win. In the second he wins with probability 1/2. Using symmetry if $p = P(A = B)$ then $P(A > B) = P(A < B) = (1-p)/2$ so the probability Al wins is $(1-p)/2 + p/2 = 1/2$.

Abstracting the structure of the last problem, let $B_1, \ldots, B_k$ be a **partition**, that is, a collection of disjoint sets whose union is Ω.



Using the fact that the sets $A \cap B_i$ are disjoint, and the multiplication rule, we have

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A|B_i)P(B_i) \tag{2.8}$$

a formula that is sometimes called the **law of total probability**.

The name of this section comes from the fact that we think of our experiment as occurring in two stages. The first stage determines which of the B's occur, and when B_i occurs in the first stage A occurs with probability $P(A|B_i)$ in the second. As the next example shows, the two stages are sometimes clearly visible in the problem itself.

Example 2.24. *Roll a die and then flip that number of coins. What is the probability of $A =$ "We get exactly 3 Heads"?*

Let $B_i =$ "The die shows i." $P(B_i) = 1/6$ for $i = 1, 2, \ldots, 6$ and

$$\begin{array}{lll} P(A|B_1) = 0 & P(A|B_2) = 0 & P(A|B_3) = 2^{-3} \\ P(A|B_4) = C_{4,3}\, 2^{-4} & P(A|B_5) = C_{5,3}\, 2^{-5} & P(A|B_6) = C_{6,3}\, 2^{-6} \end{array}$$

So plugging into (2.8),

$$\begin{aligned} P(A) &= \frac{1}{6}\left\{\frac{1}{8} + \frac{4}{16} + \frac{10}{32} + \frac{20}{64}\right\} \\ &= \frac{1}{6}\left\{\frac{8+16+20+20}{64}\right\} = \frac{1}{6} \end{aligned}$$

Example 2.25. *Suppose we roll three dice. What is the probability that the sum is 9?*

Let A = "The sum is 9," B_i = "The first die shows i," and C_j = "The sum of the second and third dice is j." Now $P(A|B_i) = P(C_{9-i})$ and we know the probabilities for the sum of two dice:

j	2	3	4	5	6	7	8	9	10	11	12
$P(C_j)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

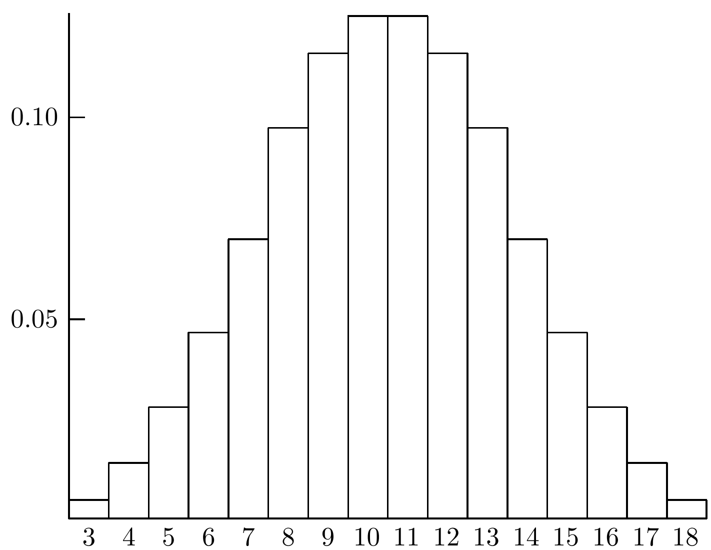
Using (2.8), now we have

$$P(A) = \sum_{i=1}^{6} P(B_i)P(A|B_i) = \frac{1}{6}\left(P(C_8) + P(C_7) + \cdots + P(C_3)\right)$$
$$= \frac{1}{6}\left(\frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36}\right) = \frac{25}{216}$$

In the same way we can compute the probability of A_k = "The sum of three dice is k":

k	3,18	4,17	5,16	6,15	7,14	8,13	9,12	10,11
$P(A_k)$	$\frac{1}{216}$	$\frac{3}{216}$	$\frac{6}{216}$	$\frac{10}{216}$	$\frac{15}{216}$	$\frac{21}{216}$	$\frac{25}{216}$	$\frac{27}{216}$

The next graph shows the shape of the distribution. Note that the triangular shape of the sum of two dice has become a little more rounded.



Example 2.26 (Craps). *In this game, if the sum of the two dice is 2, 3, or 12 on his first roll, the player loses; if the sum is 7 or 11, he wins; if the sum is 4, 5, 6, 8, 9, or 10, this number becomes his "point" and he wins if he "makes his point," i.e., his number comes up again before he throws a 7. What is the probability the player wins?*

The first step in analyzing craps is to compute the probability that the player makes his point. Suppose his point is 5 and let E_k be the event that the sum is k. There are 4 outcomes in E_5 $((1,4), (2,3), (3,2), (4,1))$, 6 in E_7, and hence 26 not in $E_5 \cup E_7$. Letting $\times$ stand for "The sum is not 5 or 7," we see that

$$P(5) = \frac{4}{36} \quad P(\times\, 5) = \frac{26}{36} \cdot \frac{4}{36} \quad P(\times \times 5) = \left(\frac{26}{36}\right)^2 \frac{4}{36}$$

From the first three terms it is easy to see that for $k \geq 0$

$$P(\times \text{ on } k \text{ rolls then } 5) = \left(\frac{26}{36}\right)^k \frac{4}{36}$$

Summing over the possibilities, which represent disjoint ways of rolling 5 before 7, we have

$$P(5 \text{ before } 7) = \sum_{k=0}^{\infty} \left(\frac{26}{36}\right)^k \frac{4}{36} = \frac{4}{36} \cdot \frac{1}{1 - \frac{26}{36}}$$

since

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \tag{2.9}$$

Simplifying, we have $P(5 \text{ before } 7) = (4/36)/(10/36) = 4/10$. Such a simple answer should have a simple explanation, and it does. Consider an urn with four balls marked 5, six marked 7, and twenty-six marked with x. Drawing with replacement until we draw either a 5 or 7 is the same as drawing once from an urn with 10 balls with four balls marked 5 and six marked 7.

5	5	5	5	x	x	x	x	x	x	x	x
7	7	7	7	7	7	x	x	x	x	x	x
x	x	x	x	x	x	x	x	x	x	x	x

Another way of saying this is that if we ignore the outcomes that result in a sum other than 5 or 7, we reduce the sample space from Ω to $E = E_5 \cup E_7$ and the distribution of the first outcome that lands in E follows the conditional probability $P(\cdot|E)$. Since $E_5 \cap E = E_5$ we have

$$P(E_5|E) = \frac{P(E_5)}{P(E)} = \frac{4/36}{10/36} = \frac{4}{10}$$

The last argument generalizes easily to give the probabilities of making any point

k	4	5	6	8	9	10
$\|E_k\|$	3	4	5	5	4	3
P(k before 7)	3/9	4/10	5/11	5/11	4/10	3/9

To compute the probability of $A =$ "He wins," we let $B_k =$ "The first roll is k," and observe that (2.8) implies

$$P(A) = \sum_{k=2}^{12} P(A \cap B_k) = \sum_{k=2}^{12} P(B_k)P(A|B_k)$$

When $k = 2, 3$, or 12 comes up on the first roll we lose, so

$$P(A|B_k) = 0 \quad \text{and} \quad P(A \cap B_k) = 0$$

When $k = 7$ or 11 comes up on the first roll we win, so

$$P(A|B_k) = 1 \quad \text{and} \quad P(A \cap B_k) = P(B_k)$$

When the first roll is $k = 4, 5, 6, 8, 9$, or 10, $P(A|B_k) = P(k$ before 7) and $P(A \cap B_k)$ is

$$\frac{3}{36} \cdot \frac{3}{9} \quad k = 4, 10 \qquad \frac{4}{36} \cdot \frac{4}{10} \quad k = 5, 9 \qquad \frac{5}{36} \cdot \frac{5}{11} \quad k = 6, 8$$

Adding up the terms in the sum in the order in which they were computed,

$$\begin{aligned} P(A) &= \frac{6}{36} + \frac{2}{36} + 2\left(\frac{1}{36} + \frac{4 \cdot 2}{36 \cdot 5} + \frac{5 \cdot 5}{36 \cdot 11}\right) \\ &= \frac{4}{18} + 2\left(\frac{55 + 88 + 125}{36 \cdot 11 \cdot 5}\right) = \frac{220 + 268}{18 \cdot 11 \cdot 5} = \frac{488}{990} = 0.4929 \quad (2.10) \end{aligned}$$

which is not very much less than $1/2 = 495/990$.

Example 2.27. *Al and Bob take turns throwing one dart to try to hit a bullseye. Al hits with probability 1/4 while Bob hits with probability 1/3. If Al goes first what is the probability he will hit the first bullseye?*

Let p be the answer. By considering one cycle of the game we see

$$p = 1/4 + (3/4)(1/3)(0) + (3/4)(2/3)p$$

In words, Al wins if he hits the bullseye on the first try. If he misses and Al hits then he loses. If they both miss then it is Al's turn and the game starts over, so Al's probability of success is p. Solving $p/2 = 1/4$ or $p = 1/2$.

This reasoning can be used to compute the probability q that a player rolls a 5 before 7. By considering the outcome of the first roll $q = 4/36 + (6/36)0 + (26/36)q$ and solving we have $q = 4/10$.

March Madness. A popular example of a four-stage experiment is the NCAA basketball tournament. In 1985 the tournament expanded to 64 teams (four regions with 16 seeded teams). From then up to and including the 2004 season each seeding has had a total of 80 teams. The next table describes the relative success of the various seeds in advancing in the tournament to the rounds of 32, sweet 16, elite 8, the final 4, and the championship game. The numbers are decreasing across each row. For readability one a number becomes 0 the reamining entries are left blank. For reasons that will become clear as you read the table we have listed the seeds in n order dictated by how the games are played.

seed	32	16	8	4	2	winner
1	80	68	56	34	17	11
16	0					
8	37	9	6	3	1	1
9	43	3	1	0		
4	64	36	12	7	2	1
13	16	3	0			
5	54	28	4	3	2	0
12	26	13	1	0		
3	67	38	18	11	7	2
14	13	2	0			
6	56	30	11	3	2	1
11	24	10	3	1	0	
2	76	51	37	18	9	4
15	4	0				
7	48	13	5	0		
10	32	16	6	0		

2.4 Bayes Formula

The title of the section is a little misleading since we will regard Bayes formula as a method for computing conditional probabilities and will only reluctantly give the formula after we have done several examples to illustrate the method.

Example 2.28 (Exit polls). *In the California gubernatorial election in 1982, several TV stations predicted, on the basis of questioning people when they exited the polling place, that Tom Bradley, then mayor of Los Angeles, would win the election. When the votes were counted, however, he lost by a considerable margin. What happened?*

To give our explanation we need some notation and some numbers. Suppose we choose a person at random, let B = "The person votes for Bradley" and suppose that $P(B) = 0.45$. There were only two candidates, so this makes the probability of voting for Deukmejian $P(B^c) = 0.55$. Let A = "The voter stops and answers a question about how she voted" and suppose that $P(A|B) = 0.4$, $P(A|B^c) = 0.3$. That is, 40% of Bradley voters will respond compared to 30% of the Deukmejian voters. We are interested in computing $P(B|A)$ = the fraction of voters in our sample that voted for Bradley. By the definition of conditional probability (2.5),

$$P(B|A) = P(B \cap A)/P(A)$$

To evaluate the numerator we use the multiplication rule (2.7)

$$P(B \cap A) = P(B)P(A|B) = 0.45 \cdot 0.4 = 0.18$$

Similarly,

$$P(B^c \cap A) = P(B^c)P(A|B^c) = 0.55 \cdot 0.3 = 0.165$$

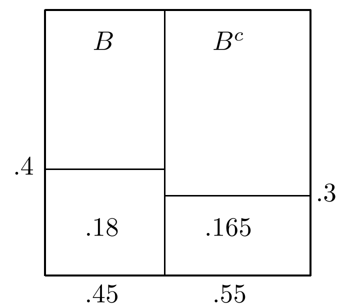
Now $P(A) = P(B \cap A) + P(B^c \cap A)$ so

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{0.18}{0.18 + 0.165} = 0.5217$$

and from our sample it looks as if Bradley will win. The problem with the exit poll is that the difference in the response rates makes our sample not representative of the population as a whole. Turning to the mechanics of the computation, note that 18% of the voters are for Bradley and respond, while 16.5% are for Deukmejian and respond, so the fraction of Bradley voters in our sample is $18/(18 + 16.5)$. In symbols,

$$P(B|A) = \frac{P(A \cap B)}{P(A \cap B) + P(A \cap B^c)}$$

In words, there are two ways an outcome can be in A – it can be in B or in B^c – and the conditional probability is the fraction of the total that comes from the first way.



Example 2.29. *Approximately 1% of women aged 40-50 have breast cancer. A woman with breast cancer has a 90% chance of a positive test from a mammogram, while a woman without has a 10% chance of a false positive result. What is the probability a woman has breast cancer given that she just had a positive test?*

Let B = "the woman has breast cancer and A = "a positive reaction." We want to calculate $P(B|A)$. By the definition of conditional probability (2.5),

$$P(B|A) = P(B \cap A)/P(A)$$

To evaluate the numerator we use the multiplication rule (2.7)

$$P(B \cap A) = P(B)P(A|B) = 0.01 \cdot 0.9 = 0.009$$

Similarly,

$$P(B^c \cap A) = P(B^c)P(A|B^c) = 0.99 \cdot 0.1 = 0.099$$

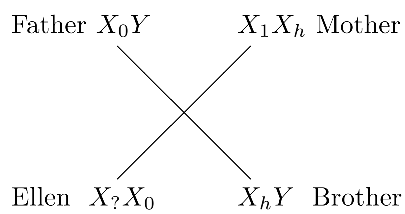
Now $P(A) = P(B \cap A) + P(B^c \cap A)$ so

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{0.009}{0.009 + 0.099} = \frac{9}{108}$$

or a little less than 9%. This situation comes about because it is much easier to have a false positive for a healthy woman which has probability 0.099, than by from a woman with breast cancer having a positive test, which has probability 0.009.

This answer is somewhat surprising. Indeed when ninety-five physicians were asked this question their average answer was 75%. The two statisticians who carried out this survey indicated that physicians were better able to see the answer when the data was presented in frequency format. 10 out of 1000 women have breast canner. Of these 9 will have a positive mammogram. However of the remaining 990 women without breast cancer 99 will have a positive reaction, and again we arrive at the answer $9/(9 + 99)$.

Example 2.30. *Ellen has a brother with hemophilia but two parents who do not have the disease. Since hemophilia is caused by a recessive allele h on the X chromosome, we can infer that her mother is a carrier (that is, the mother has the hemophilia allele h on one of her X chromosomes and the healthy allele H on the other), while her father has the healthy allele on his one X chromosome. Since Ellen received one X chromosome from her father and one from her mother, there is a 50% chance that she is a carrier, and if so, there is a 50% chance that her sons will have the disease. If she has two sons without the disease, what is the probability she is a carrier?*



Let B be the event that she is a carrier and A be the event that she has two healthy sons. We want to compute $P(B|A)$. By the definition of conditional probability (2.5),

$$P(B|A) = P(B \cap A)/P(A)$$

To evaluate the numerator we use the multiplication rule (2.7)

$$P(B \cap A) = P(B)P(A|B) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

since the probability of having two healthy sons when she is a carrier is 1/4. Similarly,

$$P(B^c \cap A) = P(B^c)P(A|B^c) = \frac{1}{2} \cdot 1 = \frac{1}{2}$$

Now $P(A) = P(B \cap A) + P(B^c \cap A)$ so

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{1/8}{1/8 + 1/2} = \frac{1}{5}$$

Example 2.31 (Paternity probabilities). *Before there were sophisticated tests based on DNA samples, the testing of blood type and other hereditary factors was used in paternity cases to infer, using Bayes formula, the probability that a particular man is the father. For a concrete example suppose that the baby's blood type is B, the mother's is A, and that of the suspected father, whom for convenience we will call Bob, is B. Given this information what is the probability Bob is the father?*

To explain how this could happen, we note that the genes that control blood type can be O, A, or B, with A and B dominant over O but neither A nor B dominating the other, so we get the following correspondence between genotypes (the genes on the two chromosomes) and phenotypes (observed blood type):

genotype	OO	AO	AA	BO	BB	AB
phenotype	O	A	A	B	B	AB
proportion	.479	.310	.050	.116	.007	.038

From this table, we see that if the baby's blood type is B then it must be the case that the mother's genotype is AO, she contributed an O gene, and the father contributed a B gene.

Let E (for "evidence") be the event that the baby's blood type is B, and F be the event that Bob is indeed the father. We cannot observe Bob's genotype, but using the proportions of the various genotypes from the table (which were inferred from the observed proportion of phenotypes using a generalization of the Hardy-Weinberg law, discussed above we can compute that

$$P(\text{genotype is } BO|\,\text{phenotype is } B) = 0.116/0.123$$
$$P(E|F) = \frac{(0.116)0.5 + 0.007}{0.123} = \frac{0.065}{0.123} = 0.528$$

There is not too much to argue about in the last computation. When we compute $P(E|F^c)$, we make the first of two questionable assumptions: If Bob is not the father, then the real father is someone chosen at random from the population, so

$$P(E|F^c) = (0.116)0.5 + 0.007 = 0.065$$

To evaluate $P(F|E)$ we thus need to evaluate the **prior probability** $P(F)$ that Bob is the father. It would be natural to make $P(F)$ equal to the fraction of times that the mother had intercourse with Bob near the time of conception. However, it is not unusual for the mother to claim this number is 1 and the alleged father to claim it is 0, so the common practice in these computations is to set $P(F) = 1/2$ (our second questionable assumption). If we do this, then $P(F) = P(F^c) = 1/2$ so

$$\begin{aligned} P(F|E) &= \frac{P(E|F)}{P(E|F) + P(E|F^c)} \\ &= \frac{0.065/0.123}{0.065/0.123 + 0.065} = \frac{1}{1.123} = 0.8904 \end{aligned}$$

Our last example is a situation with more than two events.

Example 2.32. *Three factories make 20%, 30%, and 50% of the computer chips for a company. The probablity of a defective chip is 0.04, 0.03, and 0.02 for the three factories. We have a defective chip. What is the probability it came from Factory 3?*

Let B_i be the event that the chip came from factory i and let A be the event that the chip is defective. We want to compute $P(B_3|A)$. By the definition of conditional probability (2.5),

$$P(B_3|A) = P(B_3 \cap A)/P(A)$$

To evaluate the numerator we use the multiplication rule (2.7)

$$P(B_3 \cap A) = P(B_3)P(A|B_3) = 0.5 \cdot (0.02) = 0.01$$

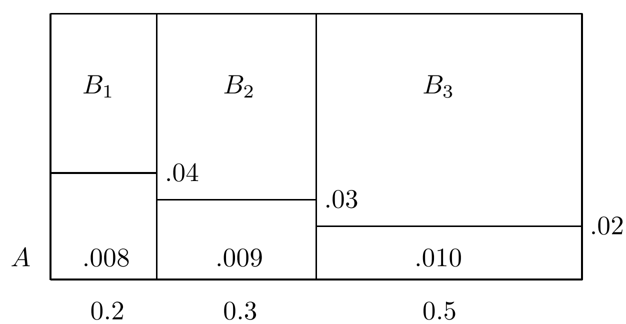
Similarly, $P(B_1 \cap A) = P(B_1)P(A|B_1) = 0.2 \cdot (0.04) = 0.008$ and

$$P(B_2 \cap A) = P(B_2)P(A|B_2) = 0.3 \cdot (0.03) = 0.009$$

Now $P(A) = \sum_i P(B_i \cap A)$ so

$$P(B_3|A) = \frac{P(B_3 \cap A)}{P(A)} = \frac{0.010}{0.008 + 0.009 + 0.010} = \frac{10}{27}$$

The calculation can be summarized by the following picture. The conditional probability $P(B_3|A)$ is the fraction of the event A that lies in B_3.



We are now ready to generalize from our examples and state **Bayes formula**. In each case, we have a **partition** of the probability space $B_1, \dots, B_n$, i.e., a sequence of disjoint sets with $\cup_{i=1}^n B_i = \Omega$. (In the first three examples, $B_1 = B$ and $B_2 = B^c$.) We are given $P(B_i)$ and $P(A|B_i)$ for $1 \le i \le n$ and we want to compute $P(B_1|A)$. By the definition of conditional probability (2.1),

$$P(B_1|A) = P(B_1 \cap A)/P(A)$$

To evaluate the numerator and denominator we observe that

$$P(B_i \cap A) = P(B_i)P(A|B_i)$$

and $P(A) = \sum_i P(B_i \cap A)$ so

$$P(B_1|A) = \frac{P(B_1 \cap A)}{P(A)} = \frac{P(B_1)P(A|B_1)}{\sum_i P(B_i)P(A|B_i)} \tag{2.11}$$

This is Bayes formula. Even though we have numbered it, we advise you not to memorize it. It is much better to remember the procedures we followed to compute the conditional probability.

O.J. Simpson trial. DNA testing has considerably more power than blood tests. RFLP's (restriction fragment length polymorphisms) are typed by digesting DNA with "restriction" enzymes and then determining the lengths of the fragments. These lengths are highly variable (polymorphic) in humans, so the use of eight or nine such markers results in incredibly small probabilities. For example, Robin Cotton of Cellmark Diagnostics testified that blood found on a sock near Simpson's bed had the genetic type of Nicole Brown Simpson and the chances of another person having the exact same RFLP alleles were 1 in 9.7 billion.

For this problem we will concentrate on a less dramatic example that involves blood found at the murder scence (item #49). Three blood factors were recorded that matched Simpson's blood types. The next table which comes from p.10 of Vol. 7, No. 4 of *Chance* gives the frequencies estimated from the overall population.

System	Item #49	Frequency
ABO	A	0.347
EsD	1	0.79
PGM	2+,2−	0.016

Here 2+, 2− indicates that two alleles were present one inherited from each parent.

Multiplying the probabilities together gives 0.00438 or 1/227, a number that was approximated in the trial and quoted in the press as 1/200. Letting E denote this evidence and G the event that Simpson is guilty, and sticking with the simpler fraction, we see that $P(E|G^c) = 1/200$. It is an error known as the "Prosecutor's Fallacy" to think of this as $P(G^c|E)$, i.e., the probability that Simpson is innocent given this evidence is 1/200. A second error known as the "Defendant's Fallacy" is to note that 1/200 of the population of Los Angeles is 40,000, so the probability that it is O.J. Simpson's blood is 1/40,000.

Both of these fallacies are based on assuming that unknown probabilities are uniform on the set of possibilities. The correct way to compute $P(G|E)$ is

$$P(G|E) = \frac{P(E|G)P(G)}{P(E|G)P(G) + P(E|G^c)P(G^c)}$$

but of course this requires giving a value to $P(G)$. It is perhaps for this reason that Bruce Weir (Nature Genetics, Vol 11, pages 365–368) argues for the use of the likelihood ratio $P(E|G)/P(E|G^c) = 200$, i.e., the evidence is 200 times more likely if O.J. Simpson is guilty than if the murdered is a randomly chosen person.

2.5 Exercises

Independence

1. Suppose we draw two cards out of a deck of 52. Let $A =$ " The first card is an Ace," $B =$ "The second card is a spade." Are A and B independent?

2. A family has three children, each of whom is a boy or a girl with probability 1/2. Let $A =$ "There is at most 1 girl," $B =$ "The family has children of both sexes." (a) Are A and B independent? (b) Are A and B independent if the family has four children?

3. Suppose we roll a red and a green die. Let $A =$ "The red die shows a 2 or a 5," $B =$ "The sum of the two dice is at least 7." Are A and B independent?

4. Roll two dice. Let $A =$ the sum is even, $B =$ the sum is divisible by 3, i.e., $B = \{3, 6, 9, 12\}$. Are A and B independent?

5. Roll two dice. Let $A =$ "The first die is odd," $B =$ "The second die is odd," and $C =$ "The sum is odd." Show that these events are pairwise independent but not independent.

6. Nine children are seated at random in three rows of three desks. Let $A =$ "Al and Bobby sit in the same row," $B =$ "Al and Bobby both sit at one of the four corner desks." Are A and B independent?

7. Two students, Alice and Betty, are registered for a statistics class. Alice attends 80% of the time, Betty 60% of the time, and their absences are independent. On a given day, what is the probability (a) at least one of these students is in class (b) exactly one of them is there?

8. Let A and B be two independent events with $P(A) = 0.4$ and $P(A \cup B) = 0.64$. What is $P(B)$?

9. Three students each have probability 1/3 of solving a problem. What is the probability at least one of them will solve the problem?

10. Three independent events have probabilities 1/4, 1/3, and 1/2. What is the probability exactly one will occur?

11. Three missiles are fired at a target. They will hit it with probabilities 0.2, 0.4, and 0.6. Find the probability that the target is hit by (a) three, (b) two, (a) one, (d) no missiles.

12. Three couples that were invited to dinner will independently show up with probabilities 0.9, 8/9, and 0.75. Let N be the number of couples that show up. Calculate the probability $N = 3, 2, 1, 0$.

13. When Al and Bob play tennis, Al wins a set with probability 0.7 while Bob wins with probability 0.3. What is the probability Al will be the first to win (a) two sets, (b) three sets?

14. Suppose that in the World Series, team B has probability 0.6 of winning each game. Assuming that the outcomes of the games are independent, what is the probability B will win the series?

15. Chevalier de Mere made money betting that he could "roll at least one 6 in four tries." When people got tired of this wager he changed it to "roll at least one double 6 in 24 tries" but then he started losing money. Compute the probabilities of winning these two bets.

16. Samuel Pepys wrote to Isaac Newton: "What is more likely, (a) one 6 in 6 rolls of one die or (b) two 6's in 12 rolls?" Compute the probabilities of these events.

17. A die is rolled 8 times. What is the probability we will get exactly two 3's?

18. In 1997, 10.8% of female smokers smoked cigars. In a sample of size 20 (a) What is the probability that exactly 2 of the women smoke cigars? (b) Using your calculator determine the probability that in the sample at most 2 smoke cigars.

19. A 1994 report revealed that 32.6% of U.S. births were to unmarried women. A parenting magazine selected 30 women who gave birth in 1994 at random. (a) What is the probability that exactly 10 of the women were unmarried? (b) Using your calculator determine the probability that in the sample at most 10 are unmarried.

20. 20% of all students are left handed. A class of size 20 meets in a room with 5 left-handed and 18 right handed chairs. Use your calculator to find the probability that each student will have a chair to match their needs.

21. David claims to be able to distinguish brand B beer from brand H but Alice claims that he just guesses. They set up a taste test with 10 small glasses of beer. David wins if he gets 8 or more right. What is the probability he will win (a) if he is just guessing? (b) if he gets the right answer with probability 0.9?

22. The following situation comes up the game of Yahtzee. We have three rolls of five dice and want to get three sixes or more. On each turn we reroll any dice that are not 6's. What is the probability we succeed?

23. A baseball pitcher throws a strike with probability 0.5 and a ball with probability 0.5. He is facing a batter who never swings at a pitch. What is the probability that he strikes out, i.e., gets three strikes before four balls.

24. A baseball player is said to "hit for the cycle" if he has a single, a double, a triple, and a home run all in one game. Suppose these four types of hits have probabilities 1/6, 1/20, 1/120, and 1/24. What is the probability of hitting for the cycle if he gets to bat (a) four times, (b) five times? (c) Using $P(\cup_i A_i) \leq \sum_i P(A_i)$ shows that the answer to (b) is at most 5 times the answer to (a). What is the ratio of the two answers?

Conditional probability

25. A friend flips two coins and tells you that at least one is Heads. Given this information, what is the probability that the first coin is Heads?

26. A friend rolls two dice and tells you that there is at least one 6. What is the probability the sum is at least 9?

27. Suppose we roll two dice. What is the probability that the sum is 7 given that neither die showed a 6?

28. Suppose you draw five cards out of a deck of 52 and get 2 spades and 3 hearts. What is the probability the first card drawn was a spade?

29. Two people, whom we call South and North, draw 13 cards out of a deck of 52. South has two Aces. What is the probability that North has (a) none? (b) one? (c) the other two?

30. An urn contains 8 red, 7 blue, and 5 green balls. You draw out two balls and they are different colors. Given this, what is the probability the two balls were red and blue?

31. Suppose 60% of the people subscribe to newspaper A, 40% to newspaper B, and 30% to both. If we pick a person at random who subscribes to at least one newspaper, what is the probability she subscribes to newspaper A?

32. In a town 40% of families have a dog and 30% have a cat. 25% of families with a dog also have a cat. (a) What fraction of people have a dog or cat? (b) What is the probability a family with a cat has a dog?

33. Plumber Bob does 40% of the plumbing jobs in a small town. 30% of the people in town are unhappy with their plumbers but 50% of Bob's customers are unhappy with his work. If your neighbor is not happy with his plumber, what is the probability it was Bob?

34. An ectopic preganancy is twice as likely if a woman smokes cigarettes. If 25% of women of childbearing age are smokers, what fraction of ectopic preganancies occur to smokers?

35. Brown eyes are dominant over blue. That is, there are two alleles B and b. bb individuals have blue eyes but other combinations has brown eyes. Your parents and you have brown eyes but your brother has blue. So you can infer that both of your parents are heterozygotes, i.e., have genetic type Bb. Given this information what is the probability you are a homozygote.

36. Suppose that the probability a married man votes is 0.45, the probability a married woman votes is 0.4, and the probability a woman votes given that her husband does is 0.6. What is the probability (a) both vote, (b) a man votes given that his wife does?

37. Two events have $P(A) = 1/4$, $P(B|A) = 1/2$, and $P(A|B) = 1/3$. Compute $P(A \cap B)$, $P(B)$, $P(A \cup B)$.

38. A, B, and C are events with $P(A) = 0.3$, $P(B) = 0.4$, $P(C) = 0.5$, A and B are disjoint, A and C are independent, and $P(B|C) = 0.1$. Find $P(A \cup B \cup C)$.

Two-stage experiments

39. From a signpost that says MIAMI two letters fall off. A friendly drunk puts the two letters back into the two empty slots at random. What is the probability that the sign still says MIAMI?

40. Two balls are drawn from an urn with balls numbered from 1 up to 10. What is the probability that the two numbers will differ by more (>) than three?

41. How can 5 black and 5 white balls be put into two urns to maximize the probability a white ball is drawn when we draw from a randomly chosen urn?

42. Suppose we draw k cards out of a deck. What is the probability that we do not draw an Ace? Is the answer larger or smaller than $(3/4)^k$?

43. You and a friend each roll two dice. What is the probability you will both have the same two numbers?

44. In a dice game the "dealer" rolls two dice, the player rolls two dice, and the player wins if his total is larger (>) than the dealer's. What is the probability the player wins?

45. What is the most likely total for the sum of four dice and what is its probability?

46. Charlie draws five cards out of a deck of 52. If he gets at least three of one suit, he discards the cards not of that suit and then draws until he again has five cards. For example, if he gets three hearts, one club, and one spade, he throws the two nonhearts away and draws two more. What is the probability he will end up with five cards of the same suit?

47. Suppose 60% of the people in a town will get exposed to flu in the next month. If you are exposed and not inoculated then the probability of your getting the flu is 80%, but if you are inoculated that probability drops to 15%. Of two executives at Beta Company, one is inoculated and one is not. What is the probability at least one will not get the flu? Assume that the events that determine whether or not they get the flu are independent.

48. John takes the bus with probability 0.3 and the subway with probability 0.7. He is late 40% of the time when he takes the bus but only 20% of the time when he takes the subway. What is the probability he is late for work?

49. The population of Cyprus is 70% Greek and 30% Turkish. 20% of the Greeks and 10% of the Turks speak English. What fraction of the people of Cyprus speak English?

50. You are going to meet a friend at the airport. Your experience tells you that the plane is late 70% of the time when it rains, but is late only 20% of the time when it does not rain. The weather forecast that morning calls for a 40% chance of rain. What is the probability the plane will be late?

51. Two boys have identical piggy banks. The older boy has 18 quarters and 12 dimes in his; the younger boy, 2 quarters and 8 dimes. One day the two banks get mixed up. You pick up a bank at random and shake it until a coin comes out. What is the probability you get a quarter? Note that there are 20 quarters and 20 dimes in all.

52. Suppose that the number of children in a family has the following distribution

number of children	0	1	2	3	4
probability	0.15	0.25	0.3	0.2	0.1

Assume that each child is independently a girl or a boy with probability 1/2 each. If a family is picked at random what is the chance it has exactly two girls.

53. A student is taking a multiple-choice test in which each question has four possible answers. She knows the answers to 50% of the questions, can narrow the choices down to two 30% of the time, and does not know anything about 20% of the questions. What is the probability she will correctly answer a question chosen at random from the test?

54. A student is taking a multiple-choice test in which each question has four possible answers. She knows the answers to 5 of the questions, can narrow the choices down to 2 in 3 cases, and does not know anything about 2 of the questions. What is the probability she will correctly answer (a) 10, (b) 9, (c) 8, (d) 7, (e) 6, (f) 5 questions?

55. Two boys, Charlie and Doug, take turns rolling two dice with Charlie going first. If Charlie rolls a 6 before Doug rolls a 7 he wins. What is the probability Charlie wins?

56. Three boys take turns shooting a basketball and have probabilities 0.2, 0.3, and 0.5 of scoring a basket. Compute the probabilities for each boy to get the first basket.

57. Change the second and third probabilities in the last problem so that each boy has an equal chance of winning.

Bayes' formula

58. 5% of men and 0.25% of women are colorblind. What is the probability a colorblind person is a man?

59. The alpha fetal protein test is meant to detect spina bifida in unborn babies, a condition that affects 1 out of 1000 children who are born. Let B be the event that the baby has spina bifida and B^c be the event that it does not. The literature on the test indicates that 5% of the time a healthy baby will cause a positive reaction. We will assume that the test is positive 100% of the time when spina bifida is present. Your doctor has just told you that your alpha fetal protein test was positive. What is the probability that your baby has spina bifida?

60. Binary digits, i.e., 0's and 1's, are sent down a noisy communications channel. They are received as sent with probability 0.9 but errors occur with probability 0.1. Assuming that 0's and 1's are equally likely, what is the probability that a 1 was sent given that we received a 1?

61. To improve the reliability of the channel described in the last example, we repeat each digit in the message three times. What is the probability that 111 was sent given that (a) we received 101? (b) we received 000?

62. Two hunters shoot at a deer, which is hit by exactly one bullet. If the first hunter hits his targets with probability 0.3 and the second with probability 0.6, what is the probability the second hunter killed the deer? The answer is not 2/3. Do you think the answer is larger or smaller?

63. A cab was involved in a hit and run accident at night. Two cab companies green and blue operate 85% and 15% of the cabs in the city respectively. A witness identified the cab as blue. However, in a test 80% of witnesses were able to correctly identify the cab color. Given this what is the probability that the cab involved in the accident was blue?

64. A student goes to class on a snowy day with probability 0.4, but on a nonsnowy day attends with probability 0.7. Suppose that 20% of the days in February are snowy. What is the probability it snowed on February 7th given that the student was in class on that day?

65. A golfer hits his drive in the fairway with probability 0.7. When he hits his drive in the fairway he makes par 80% of the time. When he doesn't he makes par only 30% of the time. He just made par on a hole. What is the probability he hit his drive in the fairway?

66. You are about to have an interview for Harvard Law School. 60% of the interviewers are conservative and 40% are liberal. 50% of the conservatives smoke cigars but only 25% of the liberals do. Your interviewer lights up a cigar. What is the probability he is a liberal?

67. Five pennies are sitting on a table. One is a trick coin that has Heads on both sides, but the other four are normal. You pick up a penny at random and flip it four times, getting Heads each time. Given this, what is the probability you picked up the two-headed penny?

68. One slot machine pays off 1/2 of the time, while another pays off 1/4 of the time. We pick one of the machines and play it six times, winning 3 times. What is the probability we are playing the machine that pays off only 1/4 of the time?

69. A student is taking a multiple choice exam in which each question has four possible answers. She knows the answers to 60% of the questions and guesses at the others. What is the probability she guessed given that she got question #12 right?

70. 20% of people are "accident-prone" and have a probability 0.15 of having an accident in a one-year period in contrast to a probability of 0.05 for the other

80% of people. (a) If we pick a person at random, what is the probability he will have an accident this year? (b) What is the probability a person is accident-prone if he had an accident last year? (c) What is the probability he will have an accident this year if he had one last year?

71. One die has 4 red and 2 white sides; a second has 2 red and 4 white sides. (a) If we pick a die at random and roll it, what is the probability the result is a red side? (b) If the first result is a red side and we roll the same die again, what is the probability of a second red side?

72. A particular football team is known to run 40% of its plays to the left and 60% to the right. When the play goes to the right, the right tackle shifts his stance 80% of the time, but does so only 10% of the time when the play goes to the left. As the team sets up for the play the right tackle shifts his stance. What is the probability that the play will go to the right?

73. A company gives a test to 100 salesmen, 80 with good sales records and 20 with poor records. 60% of the good salesmen pass the test, but only 30% of the poor salesmen do. A new applicant takes the test and passes. What is the probability he is a good salesman?

74. You are a serious student who studies on Friday nights but your roommate goes out and has a good time. 40% of the time he goes out with his girlfriend; 60% of the time he goes to a bar. 30% of the times when he goes out with his girlfriend he spends the night at her apartment. 40% of the times when he goes to a bar he gets in a fight and gets thrown in jail. You wake up on Saturday morning and your roomate is not home. What is the probability he is in jail?

75. Two masked robbers try to rob a crowded bank during the lunch hour but the teller presses a button that sets off an alarm and locks the front door. The robbers, realizing they are trapped, throw away their masks and disappear into the chaotic crowd. Confronted with 40 people claiming they are innocent, the police give everyone a lie detector test. Suppose that guilty people are detected with probability 0.95, and innocent people appear to be guilty with probability 0.01. What is the probability Mr. Jones is guilty given that the lie detector says he is?

76. Three bags lie on the table. One has two gold coins, one has two silver coins, and one has one silver and one gold. You pick a bag at random, and pick out one coin. If this coin is gold, what is the probability you picked from the bag with two gold coins?

77. In a certain city 30% of the people are Conservatives, 50% are Liberals, and 20% are Independents. In a given election, 2/3 of the Conservatives voted, 80% of the Liberals voted, and 50% of the Independents voted. If we pick a voter at random what is the probability she is Liberal?

78. An undergraduate student has asked a professor for a letter of recommendation. He estimates that the probability he will get the job is 0.8 with a strong letter, 0.4 with a medium letter, and 0.1 with a weak letter. He also believes

that the probabilities that the letter will be strong, medium, or weak are 0.5, 0.3, and 0.2. What is the probability that the letter was strong given that he got the job.

79. 1 out of 1000 births results in fraternal twins; 1 out of 1500 births results in identical twins. Identical twins must be the same sex but the sexes of fraternal twins are independent. If two girls are twins, what is the probability they are fraternal twins?

80. Consider the following data on traffic accidents

age group	% of drivers	accident probability
16 to 25	15	.10
26 to 45	35	.04
46 to 65	35	.06
over 65	15	.08

Calculate (a) the probability a randomly chosen driver will have an accident this year, and (b) the probability a driver is between 46 and 65 given that she had an accident.