SECTION 6.2 Probability Theory

This section introduced several basic concepts from probability. You should be able to apply the definitions to do the kinds of computations shown in the examples. Students interested in further work in probability, especially as it applies to statistics, should consult an elementary probability and statistics textbook. It is advisable (almost imperative) for all mathematics, engineering, and computer science majors to take a good statistics course.

1. We are told that $p(H) = 3p(T)$. We also know that $p(H) + p(T) = 1$, since heads and tails are the only two outcomes. Solving these simultaneous equations, we find that $p(T) = 1/4$ and $p(H) = 3/4$. An interesting example of an experiment in which intuition would tell you that outcomes should be equally likely but in fact they are not is to spin a penny on its edge on a smooth table and let it fall. Repeat this experiment 50 times or so, and you will be amazed at the outcomes. Make sure to count only those trials in which the coin spins freely for a second or more, not bumping into any objects or falling off the table.

3. Let us denote by t the probability that a 2 or 4 appears. Then the given information tells us that $t = 3(1-t)$, since $1-t$ is the probability that some other number appear. Solving this equation gives $t = 3/4$. We assume from the statement of the problem that 2 and 4 are equally likely. Since together they have probability $3/4$, each of them must have probability $3/8$. Similarly, each of the other numbers (1, 3, 5, or 6) must have probability $(1-t)/4$, which works out to $1/16$.

5. There are six ways to roll a sum of 7. We can denote them as $(1,6)$, $(2,5)$, $(3,4)$, $(4,3)$, $(5,2)$, and $(6,1)$, where (i,j) means rolling i on the first die and j on the second. We need to compute the probability of each of these outcomes and then add them to find the probability of rolling a 7. The two dice are independent, so we can argue as follows, using the given information about the probability of each outcome on each die: $p((1,6)) = \frac{1}{7}\cdot\frac{1}{7} = \frac{1}{49}$; $p((2,5)) = \frac{1}{7}\cdot\frac{1}{7} = \frac{1}{49}$; $p((3,4)) = \frac{1}{7}\cdot\frac{1}{7} = \frac{1}{49}$; $p((4,3)) = \frac{2}{7}\cdot\frac{2}{7} = \frac{4}{49}$; $p((5,2)) = \frac{1}{7}\cdot\frac{1}{7} = \frac{1}{49}$; $p((6,1)) = \frac{1}{7}\cdot\frac{1}{7} = \frac{1}{49}$. Adding, we find that the probability of rolling a 7 as the sum is $9/49$.

7. We exploit symmetry in answering many of these.

a) Since 1 has either to precede 4 or to follow it, and there is no reason that one of these should be any more likely than the other, we immediately see that the answer is $1/2$. We could also use brute force here, list all 24 permutations, and count that 12 of them have 1 preceding 4.

b) By the same reasoning as in part **(a)**, the answer is again $1/2$.

c) We could list all 24 permutations, and count that 8 of them have 4 preceding both 1 and 2. But here is a better argument. Among the numbers 1, 2, and 4, each is just as likely as the others to occur first. Thus by symmetry the answer is $1/3$.

d) We could list all 24 permutations, and count that 6 of them have 4 preceding 1, 2, and 3 (i.e., 4 occurring first); or we could argue that there are $3! = 6$ ways to write down the rest of a permutation beginning with 4. But here is a better argument. Each of the four numbers is just as likely as the others to occur first. Thus by symmetry the answer is $1/4$.

e) We could list all 24 permutations, and count that 6 of them have 4 preceding 3, and 2 preceding 1. But here is a better argument. Between 4 and 3, each is just as likely to precede the other, so the probability that 4 precedes 3 is $1/2$. Similarly, the probability that 2 precedes 1 is $1/2$. The relative position of 4 and 3 is independent of the relative position of 2 and 1, so the probability that both happen is the product $(1/2)(1/2) = 1/4$.

9. Note that there are $26!$ permutations of the letters, so the denominator in all of our answers is $26!$. To find the numerator, we have to count the number of ways that the given event can happen. Alternatively, in some cases we may be able to exploit symmetry.

a) There is only one way for this to happen, so the answer is $1/26!$.

b) There are $25!$ ways to choose the rest of the permutation after the first letter has been specified to be z. Therefore the answer is $25!/26! = 1/26$. Alternatively, each of the 26 letters is equally likely to be first, so the probability that z is first is $1/26$.

c) Since z has either to precede a or to follow it, and there is no reason that one of these should be any more likely than the other, we immediately see that the answer is $1/2$.

d) In effect we are forming a permutation of 25 items—the letters b through y and the double letter combination az. There are $25!$ ways to do this, so the answer is $25!/26! = 1/26$. Here is another way to reason. For a to immediately precede z in the permutation, we must first make sure that z does not occur in the first spot (since nothing precedes it), and the probability of that is clearly $25/26$. Then the probability that a is the letter immediately preceding z given that z is not first is $1/25$, since each of the 25 other letters is equally likely to be in the position in front of z. Therefore the desired probability is $(25/26)(1/25) = 1/26$. Note that this "product rule" is essentially just the definition of conditional probability.

e) We solve this by the first technique used in part (**d**). In effect we are forming a permutation of 24 items, one of which is the triple letter combination amz. There are $24!$ ways to do this, so the answer is $24!/26! = 1/650$.

f) If m, n, and o are specified to be in their original positions, then there are only 23 letters to permute, and there are $23!$ ways to do this. Therefore the probability is $23!/26! = 1/15600$.

11. Clearly $p(E \cup F) \geq p(E) = 0.7$. Also, $p(E \cup F) \leq 1$. If we apply Theorem 2 from Section 6.1, we can rewrite this as $p(E) + p(F) - p(E \cap F) \leq 1$, or $0.7 + 0.5 - p(E \cap F) \leq 1$. Solving for $p(E \cap F)$ gives $p(E \cap F) \geq 0.2$.

13. The items in this inequality suggest that it may have something to do with the formula for the probability of the union of two events given in this section:

$$p(E \cup F) = p(E) + p(F) - p(E \cap F)$$

We know that $p(E \cup F) \leq 1$, since no event can have probability exceeding 1. Thus we have

$$1 \geq p(E) + p(F) - p(E \cap F)\,.$$

A little algebraic manipulation easily transforms this to the desired inequality.

15. Let us start with the simplest nontrivial case, namely $n = 2$. We want to show that

$$p(E_1 \cup E_2) \leq p(E_1) + p(E_2)\,.$$

We know from the formula for the probability of the union of two events given in this section that

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)\,.$$

Since $p(E_1 \cap E_2) \geq 0$, the desired inequality follows immediately. We can use this as the basis step of a proof by mathematical induction. (Technically, we should point out that for $n = 1$ there is nothing to prove, since $p(E_1) \leq p(E_1)$.) For the inductive step, assume the stated inequality for n. Then

$$\begin{aligned} p(E_1 \cup E_2 \cup \cdots \cup E_n \cup E_{n+1}) &\leq p(E_1 \cup E_2 \cup \cdots \cup E_n) + p(E_{n+1}) \\ &\leq p(E_1) + p(E_2) + \cdots + p(E_n) + p(E_{n+1})\,, \end{aligned}$$

as desired. The first inequality follows from the case $n = 2$, and the second follows from the inductive hypothesis.

17. There are various ways to prove this algebraically. Here is one. Since $E \cup \overline{E}$ is the entire sample space S, we can break the event F up into two disjoint events, $F = S \cap F = (E \cup \overline{E}) \cap F = (E \cap F) \cup (\overline{E} \cap F)$, using the distributive law. Therefore $p(F) = p((E \cap F) \cup (\overline{E} \cap F)) = p(E \cap F) + p(\overline{E} \cap F)$, since these two events are disjoint. Subtracting $p(E \cap F)$ from both sides, using the fact that $p(E \cap F) = p(E) \cdot p(F)$ (our hypothesis that E and F are independent), and factoring, we have $p(F)(1 - p(E)) = p(\overline{E} \cap F)$. Since $1 - p(E) = p(\overline{E})$, this says that $p(\overline{E} \cap F) = p(\overline{E}) \cdot p(F)$, as desired.

19. As instructed, we are assuming here that births are independent and the probability of a birth in each month is $1/12$. Although this is clearly not exactly true (for example, the months do not all have the same lengths), it is probably close enough for our answers to be approximately accurate.

a) The probability that the second person has the same birth month as the first person (whatever that was) is $1/12$.

b) We proceed as in Example 13. The probability that all the birth months are different is

$$p_n = \frac{11}{12} \cdot \frac{10}{12} \cdots \frac{13-n}{12}$$

since each person after the first must have a different birth month from all the previous people in the group. Note that if $n \geq 13$, then $p_n = 0$ since the 12^{th} fraction is 0 (this also follows from the pigeonhole principle). The probability that at least two are born in the same month is therefore $1 - p_n$.

c) We compute $1 - p_n$ for $n = 2, 3, \ldots$ and find that the first time this exceeds $1/2$ is when $n = 5$, so that is our answer. With five people, the probability that at least two will share a birth month is about 62%.

21. If n people are chosen at random, then the probability that all of them were born on a day other than April 1 is $(365/366)^n$. To compute the probability that exactly one of them is born on April 1, we note that this can happen in n different ways (it can be any of the n people), and the probability that it happens for each particular person is $(1/366)(365/366)^{n-1}$, since the other $n-1$ people must be born on some other day. Putting this all together, the probability that two of them were born on April 1 is $1 - (365/366)^n - n(1/366)(365/366)^{n-1}$. Using a calculator or computer algebra system, we find that this first exceeds $1/2$ when $n = 614$. Interestingly, if the problem asked about *exactly* two April 1 birthdays, then the probability is $C(n,2)(1/366)^2(365/366)^{n-2}$, which never exceeds $1/2$.

23. There are 16 equally likely outcomes of flipping a fair coin five times in which the first flip comes up heads (each of the other flips can be either heads or tails). Of these only four will result in four heads appearing, namely $HHHHT$, $HHHTH$, $HHTHH$, and $HTHHH$. Therefore by the definition of conditional probability the answer is $4/16$, or $1/4$.

25. There are 16 equally likely bit strings of length 4, but only 8 of them start with a 1. Three of these contain at least two consecutive 0's, namely 1000, 1001, and 1100. Therefore by the definition of conditional probability the answer is $3/8$.

27. In each case we need to compute $p(E)$, $p(F)$, and $p(E \cap F)$. Then we need to compare $p(E) \cdot p(F)$ to $p(E \cap F)$; if they are equal, then by definition the events are independent, and otherwise they are not. We assume that boys and girls are equally likely, and that successive births are independent. (Medical science suggests that neither of these assumptions is exactly correct, although both are reasonably good approximations.)

a) If the family has only two children, then there are four equally likely outcomes: BB, BG, GB, and GG. There are two ways to have children of both sexes, so $p(E) = 2/4$. There are three ways to have at most one boy, so $p(F) = 3/4$. There are two ways to have children of both sexes and at most one boy, so $p(E \cap F) = 2/4$. Since $p(E) \cdot p(F) = 3/8 \neq 2/4$, the events are not independent.

b) If the family has four children, then there are 16 equally likely outcomes, since there are 16 strings of length 4 consisting of B's and G's. All but two of these outcomes give children of both sexes, so $p(E) = 14/16$. Only five of them result in at most one boy, so $p(F) = 5/16$. There are four ways to have children of both sexes and at most one boy, so $p(E \cap F) = 4/16$. Since $p(E) \cdot p(F) = 35/128 \neq 4/16$, the events are not independent.

c) If the family has five children, then there are 32 equally likely outcomes, since there are 32 strings of length 5 consisting of B's and G's. All but two of these outcomes give children of both sexes, so $p(E) = 30/32$. Only six of them result in at most one boy, so $p(F) = 6/32$. There are five ways to have children of both sexes and at most one boy, so $p(E \cap F) = 5/32$. Since $p(E) \cdot p(F) = 45/256 \neq 5/32$, the events are not independent.

29. We can model this problem using the binomial distribution. We have here $n = 6$ Bernoulli trials (the six coins being flipped), with $p = 1/2$ (the probability of heads, which we will arbitrarily call success). There are two ways in which there could be an odd person out. Either there could be five heads and one tail, or there could be one head and five tails. Thus we want to know the probability that the number of successes is either $k = 5$ or $k = 1$. According to the formula developed in this section,

$$b(5; 6, \frac{1}{2}) = C(6,5)\left(\frac{1}{2}\right)^5\left(1-\frac{1}{2}\right)^1 = \frac{3}{32},$$

and by a similar calculation, $b(1; 6, 1/2) = 3/32$. Therefore the probability that there is an odd man out is $2 \cdot (3/32) = 3/16$, or about one in five.

31. In each case we need to calculate the probability of having five girls. By the independence assumption, this is just the product of the probabilities of having a girl on each birth.

a) Since the probability of a girl is $1/2$, the answer is $(1/2)^5 = 1/32 \approx 0.031$.

b) The is the same as part **(a)**, except that the probability of a girl is 0.49. Therefore the answer is $0.49^5 \approx 0.028$.

c) Plugging in $i = 1,2,3,4,5$, we see that the probability of having boys on the successive births are 0.50, 0.49, 0.48, 0.47, and 0.46. Therefore the probability of having girls on the successive births are 0.50, 0.51, 0.52, 0.53, and 0.54. The answer is thus $0.50 \cdot 0.51 \cdot 0.52 \cdot 0.53 \cdot 0.54 \approx 0.038$.

33. In each case we need to calculate the probability that the first child is a boy (call this $p(E)$) and the probability that the last two children are girls (call this $p(F)$). Then the desired answer is $p(E \cup F)$, which equals $p(E) + p(F) - p(E) \cdot p(F)$. This last product comes from the fact that events E and F are independent.

a) Clearly $p(E) = 1/2$ and $p(F) = (1/2) \cdot (1/2) = 1/4$. Therefore the answer is

$$\frac{1}{2} + \frac{1}{4} - \frac{1}{2} \cdot \frac{1}{4} = \frac{5}{8}.$$

b) Clearly $p(E) = 0.51$ and $p(F) = 0.49 \cdot 0.49 = 0.2401$. Therefore the answer is $0.51 + 0.2401 - 0.51 \cdot 0.2401 = 0.627649$.

c) Plugging in $i = 1,2,3,4,5$, we see that the probability of having boys on the successive births are 0.50, 0.49, 0.48, 0.47, and 0.46. Thus $p(E) = 0.50$ and $p(F) = 0.53 \cdot 0.54 = 0.2862$. Therefore the answer is $0.50 + 0.2862 - 0.50 \cdot 0.2862 = 0.6431$.

35. We need to use the binomial distribution, which tells us that the probability of k successes is

$$b(k; n, p) = C(n,k)p^k(1-p)^{n-k}.$$

a) Here $k = n$, since we want all the trials to result in success. Plugging in and computing, we have $b(n; n, p) = 1 \cdot p^n \cdot (1-p)^0 = p^n$.

b) There is at least one failure if and only if it is not the case that there are no failures. Thus we obtain the answer by subtracting the probability in part **(a)** from 1, namely $1 - p^n$.

c) There are two ways in which there can be at most one failure: no failures or one failure. We already computed that the probability of no failures is p^n. Plugging in $k = n - 1$ (one failure means $n - 1$ successes), we compute that the probability of exactly one failure is $b(n-1; n, p) = n \cdot p^{n-1} \cdot (1-p)$. Therefore the answer is $p^n + np^{n-1}(1-p)$. This formula only makes sense if $n > 0$, of course; if $n = 0$, then the answer is clearly 1.

d) Since this event is just that the event in part **(c)** does not happen, the answer is $1 - [p^n + np^{n-1}(1-p)]$. Again, this is for $n > 0$; the probability is clearly 0 if $n = 0$.

37. By Definition 2, the probability of an event is the sum of the probabilities of the outcomes in that event. Thus $p(\bigcup_{i=1}^{\infty} E_i)$ is the sum of $p(s)$ for each outcome s in $\bigcup_{i=1}^{\infty} E_i$. Since the E_i's are pairwise disjoint, this is the sum of the probabilities of all the outcomes in any of the E_i's, which is what $\sum_{i=1}^{\infty} p(E_i)$ is. Thus the issue is really whether one can rearrange the summands in an infinite sum of positive numbers and still get the same answer. Note that the series converges absolutely, because all terms are positive and all partial sums are at most 1. From calculus, we know that rearranging terms is legitimate in this case. An alternative proof could be based on the hint, since $\lim_{i\to\infty} p(E_i)$ is necessarily 0.

39. **a)** Since E is the event that for every set S with k players there is a player who has beaten all k of them, $\overline{E}$ is the event that for some set S with k players there is no player who has beaten all k of them. Thus for $\overline{E}$ to happen, some F_j must happen, so $\overline{E} = \bigcup_{j=1}^{C(m,k)} F_j$. The given inequality now follows from Boole's inequality (Exercise 15).

b) The probability that a particular player not in the j^{th} set beats all k of the players in the j^{th} set is $(1/2)^k = 2^{-k}$. Therefore the probability that this player does not have such a perfect record is $1 - 2^{-k}$, so the probability that all $m - k$ of the players not in the j^{th} set are unable to boast of a perfect record is $(1 - 2^{-k})^{m-k}$. That is precisely $p(F_j)$.

c) The first inequality follows immediately, since all the summands are the same and there are $C(m, k) = \binom{m}{k}$ of them. If this probability is less than 1, then it must be possible that $\overline{E}$ fails, i.e., that E happens. So there is a tournament that meets the conditions of the problem as long as the second inequality holds.

d) We ask a computer algebra system to compute $C(m, 2)(1-2^{-2})^{m-2}$ and $C(m, 3)(1-2^{-3})^{m-3}$ for successive values of m to determine which values of m make the expression less than 1. We conclude that such a tournament exists for $k = 2$ when $m \geq 21$, and for $k = 3$ when $m \geq 91$. In fact, however, these are not the smallest values of m for which such a tournament exists. Indeed, for $k = 2$, we can take the tournament in which Barb beats Laurel, Laurel beats David, and David beats Barb ($m = 3$), and according to *The Probabilistic Method*, second edition, by Noga Alon and Joel Spencer (Wiley, 2000), there is a tournament with seven players meeting the condition when $k = 3$.

41. The input to this algorithm is the integer $n > 1$ to be tested for primality and the number k of iterations desired. If the output is "composite" then we know for sure that n is composite. If the output is "probably prime" then we do not know whether or not n is prime, and of course it makes no sense to talk about the probability that n is prime (either it is or it isn't!—there is no chance involved). What we do know is that the chance that a composite number would produce the output "probably prime" is at most $1/4^k$.

```
procedure probabilistic prime(n, k)
composite := false
i := 0
while composite = false and i < k
begin
        i := i + 1
        choose b uniformly at random with 1 < b < n
        apply Miller's test to base b
        if n fails the test then composite = true
end
if composite = true then print ("composite")
else print ("probably prime")
```