3.1 Introduction

Introduce one of the most important concepts in probability theory: conditional probability.

- The importance of this chapter is twofold.
  1. We are often interested in calculating probabilities when some partial information concerning the result of the experiment is available; in such a situation the desired probabilities are conditional.
  2. Even when no partial information is available, conditional probabilities can often be used to compute the desired probabilities more easily.

3.2 Conditional probabilities

Toss 2 dice and suppose that each of the 36 possible outcomes is equally likely to occur.

- Suppose that the first die is a 3.
- What is the probability that the sum of the 2 dice equals 8?
- $E = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$
- $F = \{(3,j) \,|\, j = 1,2,3,4,5,6\}$
- $P(E|F) = \frac{P(EF)}{P(F)} = \frac{1/36}{6/36} = \frac{1}{6}$

Definition: If $P(F) > 0$, then

$$P(E|F) = \frac{P(EF)}{P(F)} \tag{2.1}$$

Example 3.2a. A coin is flipped twice. If we assume that all four points in the sample $S = \{(H,H),(H,T),(T,H),(T,T)\}$ are equally likely, what is the conditional probability that both flip result in heads, given that the first flip does?

- $E = \{(H,H)\}$: The event that both flips land heads.
- $F = \{(H,H),(H,T)\}$: The event that first flip lands heads.

- The desired probability is given by

$$\begin{aligned} P(E|F) &= \frac{P(EF)}{P(F)} \\ &= \frac{P(\{(H,H)\})}{P(\{(H,H),(H,T)\})} \\ &= \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2} \end{aligned}$$

Example 3.2b. An urn contains 10 white, 5 yellow, and 10 black marbles. A marble is chosen at random from the urn, and it is noted that it is not one of the black marbles. What is the probability that it is yellow?

- Y: The event that the marble selected is yellow.
- B^c: The event that it is not black.
- From Eq. (2.1)

$$P(Y|B^c) = \frac{P(YB^c)}{P(B^c)} = \frac{\frac{5}{25}}{\frac{15}{25}} = \frac{1}{3}$$

- Derive the probability by working directly with the reduced sample space:

$$\frac{5}{15} = \frac{1}{3}$$

Example 3.2c. In the card game bridge the 52 cards are dealt out equally to 4 players (called East, West, North, and South). If North and South have a total of 8 spades among them, what is the probability that East has 3 of the remaining 5 spades?

- Work with the reduced sample space.
- $\frac{\binom{5}{3}\binom{21}{10}}{\binom{26}{13}} \approx .339$

Example 3.2d. The organization for which Ms. Jones works is running a dinner for those employees having at least one son. If Jones is known to have two children, what is the conditional probability that they are both boys, given that she is invited to the dinner? Assume that the sample space S is given by $S = \{(b,b),(b,g),(g,b),(g,g)\}$ and all outcomes are equally likely $[(b,g)$ means, for instance, that the older child is a boy and the younger child is a girl$]$.

- F: Jones has at least one son.

- E: Both children are boys.

$$\begin{aligned} P(E|F) &= \frac{P(EF)}{P(F)} \\ &= \frac{P(\{(b,b)\})}{P(\{(b,b),(b,g),(g,b)\})} \\ &= \frac{\frac{1}{4}}{\frac{3}{4}} \\ &= \frac{1}{3} \end{aligned}$$

Multiplication rule: $P(EF) = P(F)P(E|F)$

Example 3.2e. Celine is undecided as to whether to take a French course or a chemistry course. She estimates that her probability of receiving an A grade would be $\frac{1}{2}$ in a French course, and $\frac{2}{3}$ in a chemistry course. If Celine decides to base her decision on the flip of a fair coin, what is the probability that she gets an A in chemistry?

- C: The event that Celine takes chemistry.
- A: The event that she receives an A in whatever course she takes.

$$\begin{aligned} P(CA) &= P(C)P(A|C) \\ &= \left(\frac{1}{2}\right)\left(\frac{2}{3}\right) = \frac{1}{3} \end{aligned}$$

Example 3.2f. Suppose that an urn contains 8 red balls and 4 white balls. We draw 2 balls from the urn without replacement. If we assume that at each draw each ball in the urn is equally likely to be chosen, what is the probability that both balls drawn are red?

- $R_1(R_2)$: The event that the first (second) ball drawn is red.

$$\begin{aligned} P(R_1R_2) &= P(R_1)P(R_2|R_1) \\ &= \left(\frac{2}{3}\right)\left(\frac{7}{11}\right) = \frac{14}{33} \end{aligned}$$

- Direct computation: $P(R_1R_2) = \frac{\binom{8}{2}}{\binom{12}{2}}$

The multiplication rule:

$$P(E_1E_2E_3\cdots E_n)$$
$$= P(E_1)P(E_2|E_1)P(E_3|E_1E_2)\cdots P(E_n|E_1\cdots E_{n-1})$$

Example 3.2g. An ordinary deck of 52 playing cards is randomly divided into 4 piles of 13 cards each. Compute the probability that each pile has exactly 1 ace.

- $E_1 = \{\text{the ace of spades is in any one of the piles}\}$
- $E_2 = \{\text{the aces of spades and hearts are in different piles}\}$
- $E_3 = \{\text{the aces of spades, hearts, and diamonds are all in different piles}\}$
- $E_4 = \{\text{all 4 aces are in different piles}\}$
- $P(E_1E_2E_3E_4) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)P(E_4|E_1E_2E_3)$
- $P(E_1) = 1$ since E_1 is the sample space.
- $P(E_2|E_1) = \frac{39}{51}$ since the pile containing the ace of spades will receives 12 of the remaining 51 cards.
- $P(E_3|E_1E_2) = \frac{26}{50}$ since the pile containing the ace of spades and hearts will receives 24 of the remaining 50 cards.
- $P(E_4|E_1E_2E_3) = \frac{13}{49}$
- $P(E_1E_2E_3E_4) = \frac{39\cdot26\cdot13}{51\cdot50\cdot49} \approx .105$
- Alternative:

$$\frac{4!\binom{48}{12,12,12,12}}{\binom{52}{13,13,13,13}} = \frac{39\cdot26\cdot13}{51\cdot50\cdot49}$$

3.3 Bayes' formula

- $E = EF\bigcup EF^c$ (Fig. 3.1)

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= P(E|F)P(F) + P(E|F^c)[1 - P(F)] \end{aligned}$$

Example 3.3a. (Part 1). An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. Their statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability .4, whereas this probability decreases to .2 for a non-accident-prone person. If we assume that 30 percent of the population is accident prone, what is the probability that a new policyholder will have an accident within a year of purchasing a policy?

- A_1: The event that the policyholder will have an accident within a year of purchase.
- A: The event that the policyholder is accident prone.

$$\begin{aligned} P(A_1) &= P(A_1|A)P(A) + P(A_1|A^c)P(A^c) \\ &= (.4)(.3) + (.2)(.7) = .26 \end{aligned}$$

Example 3.3a. (Part 2). Suppose that a new policyholder has an accident within a year of purchasing a policy. What is the probability that he or she is accident prone?

- The desired probability is

$$\begin{aligned} P(A|A_1) &= \frac{P(AA_1)}{P(A_1)} \\ &= \frac{P(A)P(A_1|A)}{P(A_1)} \\ &= \frac{(.3)(.4)}{.26} = \frac{6}{13} \end{aligned}$$

Example 3.3b. In answering a question on a multiple-choice test, a student either knows the answer or guesses. Let p be the probability that the student knows the answer and $1 - p$ the probability that the student guesses. Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student know the answer to a question, given that he or she answered it correctly?

- C: The event that the student answers the question correctly.
- K: The event that the student actually knows the answer.

$$\begin{aligned} P(K|C) &= \frac{P(KC)}{P(C)} \\ &= \frac{P(C|K)P(K)}{P(C|K)P(K) + P(C|K^c)P(K^c)} \\ &= \frac{p}{p + (1/m)(1-p)} \\ &= \frac{mp}{1 + (m-1)p} \end{aligned}$$

- If $m = 5, p = 1/2$, then the probability that a student knew the answer to a question he or she correctly answered is $5/6$.

Example 3.3c. A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present. However, the test also yields a "false positive" result for 1 percent of the the healthy person tested. (That is, if a healthy person is tested, then, with probability .01 the test result will imply he or she has the disease.) If .5 percent of the population actually has the disease, what is the probability a person has the disease given that the test result is positive?

- D: The event that the tested person has the disease.
- E: The event that the test result is positive.

$$\begin{aligned} P(D|E) &= \frac{P(DE)}{P(E)} \\ &= \frac{P(E|D)P(E)}{P(E|D)P(D) + P(E|D^c)P(D^c)} \\ &= \frac{(.95)(.005)}{(.95)(.005) + (.01)(.995)} \\ &= \frac{95}{294} \approx .323 \end{aligned}$$

- Alternative: $\dfrac{.95}{.95 + (199)(.01)} = \dfrac{95}{294} \approx .323$

Example 3.3d. Consider a medical practitioner pondering the following dilemma: "If I'm at least 80 percent certain that my patient has this disease, then I always recommend surgery, whereas if I'm not quite as certain, then I recommend additional tests that are expensive and sometime painful. Now, initially I was only 60 percent certain that Jones has the disease, so I ordered the series A test, which always gives a positive result when the patient has the disease and almost never does when he is healthy. The information complicates matters because, although it doesn't change my original 60 percent estimate of his chances of having the disease, it does affect the interpretation of the results of the A test. This is so because the A test, while never yielding a positive result 30 percent of the time in the case of *diabetic* patients not suffering from the disease. Now what do I do? More tests or immediate surgery?"

- D: The event that Jones has the disease.
- E: The event of positive A test result.

$$\begin{aligned} P(D|E) &= \frac{P(DE)}{P(E)} \\ &= \frac{P(E|D)P(E)}{P(E|D)P(D) + P(E|D^c)P(D^c)} \\ &= \frac{(.6)1}{1(.6) + (.3)(.4)} \\ &= .833 \end{aligned}$$

Example 3.3e. At a certain stage of a criminal investigation the inspector in charge is 60 percent convinced of a certain suspect. Suppose now that a new piece of evidence that shows that the criminal has a certain characteristic (such as left-handedness, baldness, or brown hair) is uncovered. If 20 percent of the population possesses this characteristic, how certain of guilt of the suspect should the inspector now be if it turns out that the suspect has this characteristic?

- G: The event that the suspect is guilty.
- C: The event that he possesses the characteristic of the criminal.

$$\begin{aligned} P(G|C) &= \frac{P(GC)}{P(C)} \\ &= \frac{P(C|G)P(G)}{P(C|G)P(G) + P(C|G^c)P(G^c)} \\ &= \frac{1(.6)}{1(.6) + (.2)(.4)} \approx .882 \end{aligned}$$

Example 3.3f. In the world bridge championships held in Buenos Aires in May 1965 the famous Buenos Aires in May 1965 the famous British bridge partnership of Terrence Reese and Boris Schapiro was accused of cheating by using a system of finger signals that could indicate the number of hearts held by the players. Reese and Schapiro denied the accusation, and eventually a hearing was in the form of a legal proceedings with a prosecuting and defense team, both having the power to call and cross-examine witnesses. During the course of these proceedings the prosecutor examined specific hands played by Reese and Schaprio and claimed that their playing in these hands was consistent with the hypothesis that they were guilty of having illicit knowledge of the heart suit. At this point, the defense attorney pointed out that their play of these hands was also perfectly consistent with their standard line of play. However, the prosecution then argued that as long as their play was consistent with the hypothesis of guilt, then it must be counted as evidence toward this hypothesis. What do you think of the reasoning of the prosecution?

- H: A particular hypothesis (such as the guilt of Reese and Schapiro).
- E: The new evidence.

$$\begin{aligned} P(H|E) &= \frac{P(HE)}{P(E)} \qquad (3.2) \\ &= \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|H^c)[1 - P(H)]} \end{aligned}$$

- $P(H|E) > P(H)$ if $P(E|H) \geq P(E|H)P(H) + P(E|H^c)[1 - P(H)]$ or equivalently $P(E|H) \geq P(E|H^c)$.
- $P(H|E) = \frac{P(H)}{P(H) + [1-P(H)]\frac{P(E|H^c)}{P(E|H)}}$
- The play of the cards can be considered to support the hypothesis of guilt only if such playing would have been more likely if the partnership were cheating than if they were not.

Definition:
The odds ratio of an event A is defined by

$$\frac{P(A)}{P(A^c)} = \frac{P(A)}{1 - P(A)}$$

- Consider now a hypothesis H that is true with probability $P(H)$ and suppose that new evidence is introduced.
- $P(H|E) = \dfrac{P(E|H)P(H)}{P(E)}$
- $P(H^c|E) = \dfrac{P(E|H^c)P(H^c)}{P(E)}$
- $\dfrac{P(H|E)}{P(H^c|E)} = \dfrac{P(H)}{P(H^c)} \dfrac{P(E|H)}{P(E|H^c)}$

Example 3.3g. When coin A is flipped it comes up heads with probability $\frac{1}{4}$, whereas when coin B is flipped it comes up heads with probability $\frac{3}{4}$. Suppose that one of these coins is randomly chosen and is flipped twice. If both flips land heads, what is the probability that coin B was the one flipped?

- B: The event that coin B was the one flipped.
- $P(B) = P(B^c)$
- $\frac{P(B|\text{two heads})}{P(B^c|\text{two heads})} = \frac{\frac{9}{16}}{\frac{1}{16}} = 9$
- The probability is 9/10 that coin B was the one flipped.

Generalized Eq. (3.1):

- Suppose that $F_1, F_2, \ldots, F_n$ are mutually exclusive events such that

$$\bigcup_{i=1}^{n} F_i = S. \qquad \textbf{Partition}$$

- $E = \bigcup_{i=1}^{n} EF_i$
- $E = \sum_{i=1}^{n} P(EF_i) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$

Proposition 3.1 (Bayes's formula):

$$\begin{aligned} P(F_j|E) &= \frac{P(EF_j)}{P(E)} \qquad (3.5) \\ &= \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)} \end{aligned}$$

Example 3.3h. A plane is missing, and it is presumed that it was equally likely to have gone down in any of 3 possible regions. $1-\beta_i$ denote the probability that the plane will be found upon a search of the ith region when the plane is, in fact, in that region, $i = 1, 2, 3$. (The constants β_i are called overlook probabilities because they represent the probability of overlooking the plane; they are generally attributable to the geographical and environmental conditions of the regions.) What is the conditional probability that the plane is in the ith regions, given that a search of region 1 is unsuccessful, $i = 1, 2, 3$?

- R_i: The event that the plane is in region i.
- E: The event that a search region 1 is unsuccessful.

$$\begin{aligned} P(R_1|E) &= \frac{P(ER_1)}{P(E)} \\ &= \frac{P(E|R_1)P(R_1)}{\sum_{i=1}^{3} P(E|R_i)P(R_i)} \\ &= \frac{(\beta_1)\frac{1}{3}}{(\beta_1)\frac{1}{3} + (1)\frac{1}{3} + (1)\frac{1}{3}} \\ &= \frac{\beta_1}{\beta_1 + 2} \end{aligned}$$

- For j =2,3,

$$\begin{aligned} P(R_j|E) &= \frac{P(E|R_j)P(R_j)}{P(E)} \\ &= \frac{(1)\frac{1}{3}}{(\beta_1)\frac{1}{3} + \frac{1}{3} + \frac{1}{3}} \\ &= \frac{1}{\beta_1 + 2} \qquad j = 2, 3 \end{aligned}$$

Example 3.3i. Suppose that we have 3 cards identical in form except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side black. The 3 cards are mixed up in a hat, and 1 card is randomly selected and put down on the ground. If the upper side of the chosen card is colored red, what is the probability that the other side is colored black?

- RR, BB, RB: The events that the chosen card is all red, all black, or the red-black card.

$$
\begin{aligned}
P(RB|R) &= \frac{P(RB \cap R)}{P(R)} \\
&= \frac{P(R|RB)P(RB)}{P(R|RR)P(RR) + P(R|RB)P(RB) + P(R|BB)P(BB} \\
&= \frac{(\frac{1}{2})(\frac{1}{3})}{(1)(\frac{1}{3}) + (\frac{1}{2})(\frac{1}{3}) + 0(\frac{1}{3})} \\
&= \frac{1}{3}
\end{aligned}
$$

Example 3.3j. A new couple, known to have two children, has just moved into town. Suppose that the mother is encountered walking with one of her children. If this child is a girl, what is the probability that both children are girls?

- G_1: The first child is a girl.
- G_2: The second child is a girl.
- G: The child seen with the mother is a girl.

$$
\begin{aligned}
P(G_1G_2|G) &= \frac{P(G_1G_2G)}{P(G)} \\
&= \frac{P(G_1G_2)}{P(G)}
\end{aligned}
$$

-

$$
\begin{aligned}
P(G) &= P(G|G_1G_2)P(G_1G_2) + P(G|G_1B_2)P(G_1B_2) \\
&\quad + P(G|B_1G_2)P(B_1G_2) + P(G|B_1B_2)P(B_1B_2) \\
&= P(G_1G_2) + P(G|G_1B_2)P(G_1B_2) \\
&\quad + P(G|B_1G_2)P(B_1G_2)
\end{aligned}
$$

-
$$\begin{aligned} P(G_1G_2|G) &= \frac{\frac{1}{4}}{\frac{1}{4} + \frac{P(G|G_1B_2)}{4} + \frac{P(G|B_1G_2)}{4}} \\ &= \frac{1}{1 + P(G|G_1B_2) + P(G|B_1G_2)} \end{aligned}$$

- If we assume that independent of the genders of the children, the child walking with the mother is the elder child with some probability p:
$$\begin{aligned} P(G|G_1B_2) &= p = 1 - P(G|B_1G_2) \\ P(G_1G_2|G) &= \frac{1}{2} \end{aligned}$$

- If we assume that if the children are of different genders, then the mother would choose to walk with the girl with probability q, independent of the birth order of the children:
$$\begin{aligned} P(G|G_1B_2) &= P(G|B_1G_2) = q \\ P(G_1G_2|G) &= \frac{1}{1+2q} \end{aligned}$$

Example 3.3k. At a psychiatric clinic the social workers are so busy that, on the average, only 60 percent of potential new patients that telephone are able to talk immediately with a social worker when they call. The other 40 percent are asked to leave their phone numbers. About 75 percent of the time a social worker is able to return the call on the same day, and the other 25 percent of the time the caller is contacted on the following day. Experience for consultation is .8 if the caller was immediately able to speak to a social worker, whereas it is .6 and .4, respectively, if the patient's call was returned the same day or the following day.

(a) What percentage of people that telephone visit the clinic for consultation?

(b) What percentage of patients that visit the clinic did not have to have their telephone calls returned?

- V: Caller visits the clinic for construction.
- I: Caller immediately speaks to a social worker.
- S: Caller is contacted later on the same day.
- F: Caller is contacted on the following day.

$$\begin{aligned} P(V) &= P(V|I)P(I) + P(V|S)P(S) + P(V|F)P(F) \\ &= (.8)(.6) + (.6)(.4)(.75) + (.4)(.4)(.25) \\ &= .70 \end{aligned}$$

(b)

$$\begin{aligned} P(I|V) &= \frac{P(V|I)P(I)}{P(V)} \\ &= \frac{(.8)(.6)}{.7} \\ &\approx .686 \end{aligned}$$

3.4 Independent events

$P(E|F)$:

- Knowing that F has occurred generally changes the chances of E's occurrence.
- E is independent of F if knowledge that F has occurred does not change the probability that E occurs.

Definition: Two events E and F are said to be independent if $P(EF) = P(E)P(F)$ holds. Two events E and F that are not independent are said to be dependent.

Example 3.4a. A card is selected at random from an ordinary deck of 52 playing cards.

- E: The event that the selected card is an ace.
- F: The event that it is a spade.
- Then E and F are independent.
- This follows because $P(EF) = \frac{1}{52}$, whereas $P(E) = \frac{4}{52}$ and $P(F) = \frac{13}{52}$.

Example 3.4b. Two coins are flipped, and all 4 outcomes are assumed to be equally likely.

- E: The event that the first coin lands heads.
- F: The event that the second lands tails.
- Then E and F are independent.
- $P(E) = P(\{(H,H),(H,T)\}) = \frac{1}{2}$
- $P(F) = P(\{(H,T),(T,T)\}) = \frac{1}{2}$
- $P(EF) = P(\{H,T\}) = \frac{1}{4} = P(E)P(F)$

Example 3.4c. Suppose that we toss 2 fair dice. Let E_1 denote the event that the sum of the dice is 6 and F denote the event that the first die equals 4.

- Then
$$P(E_1F) = P(\{(4,2)\}) = \frac{1}{36}$$
whereas
$$P(E_1)P(F) = \left(\frac{5}{36}\right)\left(\frac{1}{6}\right) = \frac{5}{216}$$
Hence E_1 and F are not independent.
- Intuitively, the reason for this is clear because if we are interested in the possibility of throwing a 6 (with 2 dice) we shall be quite happy if the first die lands 4 (or any of the number $1, 2, 3, 4, 5$), for then we shall still have a possibility of getting a total of 6.
- On the other hand, if the first die landed 6, we would be unhappy because we would no longer have a chance of getting a total of 6.
- In other words, our chance of getting a total of six depends on the outcome of the first die; hence E_1 and F cannot be independent.
- E_2: The event that the sum of the dice equal to 7.
- $P(E_2F) = \frac{1}{36}$
- $P(E_2)P(F) = (\frac{1}{6})(\frac{1}{6}) = \frac{1}{36}$
- E_2 and F are independent.

Example 3.4d. If we let E denote the event that the next president is a Republican and F the event that there will be a major earthquake within the next year, then most people would probably be willing to assume that E and F are independent. However, then would probably be some controversy over whether it is reasonable to assume that E is independent of G, where G is the event that there will be a major war within two years after the election.

Proposition 4.1: If E and F are independent, then so are E and F^c.

Example 3.4e. Two fair dice are thrown.

- E: The event that the sum of the dice is 7.
- F: The event that the first die equals 4.
- G: The event that the second die equals 3.
- From Example 4c we know that E is independent of F.
- E is also independent of G.

- E is not independent of FG since $P(E|FG) = 1$.

Definition: The three events E, F, and G are said to be independent if

$$\begin{aligned} P(EFG) &= P(E)P(F)P(G) \\ P(EF) &= P(E)P(F) \\ P(EG) &= P(E)P(G) \\ P(FG) &= P(F)P(G) \end{aligned}$$

- If E, F, and G are independent, then E will be independent of any event formed from F and G.
- The events $E_1, E_2, \ldots, E_n$ are said to be independent if for every subset $E_{1'}, E_{2'}, \ldots, E_{r'}, r \leq n$, of these events

$$P(E_{1'}E_{2'} \cdots E_{r'}) = P(E_{1'})P(E_{2'}) \cdots P(E_{r'})$$

- If each subexperiment has the same sample space and the same probability function on its events, then the subexperiment are called *trials*.

Example 3.4f. An infinite sequence of independent trials is to be performed. Each trial results in a success with probability p and a failure with probability $1 - p$. What is the probability that

(a) at least 1 success occurs in the first n trials;

(b) exactly k successes occur in the first n trials;

*(c) all trials result in successes?

- E_i: The event of a failure on the ith trial.

(a) $P(E_1E_2 \cdots E_n) = P(E_1)P(E_2) \cdots P(E_n) = (1-p)^n$

The desired probability is $1 - (1-p)^n$.

(b) $P\{\text{exactly } k \text{ successes}\} = \binom{n}{k} p^k (1-p)^{n-k}$

(c)

$$\begin{aligned} P(E_1^c E_2^c \cdots E_n^c) &= p^n \\ P\left(\bigcap_{i=1}^{\infty} E_i^c\right) &= P\left(\lim_{n\to\infty} \bigcap_{i=1}^{n} E_i^c\right) \\ &= \lim_{n\to\infty} P\left(\bigcap_{i=1}^{n} E_i^c\right) \\ &= \lim_n p^n = \begin{cases} 0 & \text{if } p < 1 \\ 1 & \text{if } p = 1 \end{cases} \end{aligned}$$

Example 3.4g. A system composed of n separate components is said to be a parallel system if it functions when at least one of the components functions (see Figure 3.2). For such a system, if component i, independent of other components, functions with probability $p_i, i = 1, \ldots, n$, what is the probability that the system functions?

- A_i: The event that component i functions.

$$\begin{aligned} P\{\text{system functions}\} &= 1 - P\{\text{system does not function}\} \\ &= 1 - P\{\text{all components do not function}\} \\ &= 1 - P\left(\bigcap_i A_i^c\right) \\ &= 1 - \prod_{i=1}^{n}(1 - p_i) \quad \text{by independence} \end{aligned}$$

Example 3.4h. Independent trials, consisting of rolling a pair of fair dice, are performed. What is the probability that an outcome of 5 appears before an outcome of 7 when the outcome of a roll is the sum of the dice?

- E_n: The event that no 5 or 7 appears on the first $n-1$ trials and a 5 appears on the nth trial.
- $P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n)$
- $P\{5 \text{ on any trial}\} = \frac{4}{36}$
- $P(E_n) = \left(1 - \frac{10}{36}\right)^{n-1} \frac{4}{36}$

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} E_n\right) &= \frac{1}{9}\sum_{n=1}^{\infty}\left(\frac{13}{18}\right)^{n-1} \\ &= \frac{1}{9}\frac{1}{1 - \frac{13}{18}} \\ &= \frac{2}{5} \end{aligned}$$

- Alternative by using conditional probabilities:
  - E: The event that a 5 occurs before a 7.
  - F: The event that the first trial results in a 5.
  - G: The event that the first trial results in a 7.

- H: The event that the first trial results in neither a 5 nor a 7.
- $P(E) = P(E|F)P(F) + P(E|G)P(G) + P(E|H)P(H)$
- $P(E|F) = 1;\ P(E|G) = 0;\ P(E|H) = P(E)$
- $P(E) = \frac{1}{9} + P(E)\frac{13}{18}$
- $P(E) = \frac{2}{5}$

- If E and F are mutually exclusive events of an experiment, then, when independent trials of this experiment are performed, the event E will occur before the event F with the probability

$$\frac{P(E)}{P(E) + P(F)}$$

Example 3.4i. *The problem of the points.* Independent trials, resulting in a success with probability p and and a failure with probability $1 - p$, are performed. What is the probability that n successes occur before m failure? If we think of A and B as playing a game such that A gains 1 point when a success occurs and B gains 1 point when a failure occurs, then the desired probability is the probability that A would win if the game were to be continued in a position where A needed n and B needed m more points to win.

- $P_{n,m}$: The probability that n successes occur before m failures.

- Conditioning on the outcome of the first trial:

$$P_{n,m} = (1-p)P_{n,m-1} + pP_{n-1,m} \qquad n \geq 1, m \geq 1$$

  $P_{n,0} = 0$ and $P_{0,m} = 1$

- $P_{n,m} = \sum_{k=n}^{m+n-1} \binom{m+n-1}{k} p^k (1-p)^{m+n-1-k}$

Example 3.4j. *The gambler's ruin problem.* Two gamblers, A and B, bet on outcomes of successive flip of a coin. On each flip, if the coin comes up heads, A collects 1 unit from B, whereas if it comes up tails, A pays 1 unit to B. They continue to do this until one of them runs out of money. If it is assumed that the successive flips of the coin are independent and each flip results in a head with probability p, what is the probability that A ends up with all the money if he starts with i units and B starts with $N - i$ units?

- E: The event that A ends up with all the money when he starts with i and B with $N - i$.
- $P_i = P(E)$.
- H: The event that the first flip lands heads.

$$\begin{aligned} P_i = P(E) &= P(E|H)P(H) + P(E|H^c)P(H^c) \\ &= pP(E|H) + (1-p)P(E|H^c) \end{aligned}$$

- $P(E|H) = P_{i+1}$
- $P(E|H^c) = P_{i-1}$
- $P_i = pP_{i+1} + qP_{i-1} \quad i = 1, 2, \ldots, N-1$
- $pP_i + qP_i = pP_{i+1} + qP_{i-1}$
- $P_{i+1} - P_i = \frac{q}{p}(P_i - P_{i-1}),\ i = 1, 2, \ldots, N-1$
- $P_0 = 0$

$$\begin{aligned}
P_2 - P_1 &= \frac{q}{p}(P_1 - P_0) = \frac{q}{p}P_1 \\
P_3 - P_2 &= \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1 \\
&\vdots \\
P_i - P_{i-1} &= \frac{q}{p}\left(P_{i-1} - P_{i-2}\right) = \left(\frac{q}{p}\right)^{i-1} P_1 \\
&\vdots \\
P_N - P_{N-1} &= \frac{q}{p}(P_{N-1} - P_{N-2}) = \left(\frac{q}{p}\right)^{N-1} P_1
\end{aligned}$$

- $P_i - P_1 = P_1 \left[\left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^{i-1}\right]$
- $$P_i = \begin{cases} \frac{1-(q/p)^i}{1-(q/p)} P_1 & \text{if } \frac{q}{p} \neq 1 \\ iP_1 & \text{if } \frac{q}{p} = 1 \end{cases}$$
- $P_N = 1$
$$P_1 = \begin{cases} \frac{1-(q/p)}{1-(q/p)^N} & \text{if } p \neq \frac{1}{2} \\ \frac{1}{N} & \text{if } p = \frac{1}{2} \end{cases}$$
- $$P_i = \begin{cases} \frac{1-(q/p)^i}{1-(q/p)^N} & \text{if } p \neq \frac{1}{2} \\ \frac{i}{N} & \text{if } p = \frac{1}{2} \end{cases}$$
- Q_i: The probability that B winds up with all the money when A starts with i and B with $N-i$.
$$Q_i = \begin{cases} \frac{1-(p/q)^{N-i}}{1-(p/q)^N} & \text{if } q \neq \frac{1}{2} \\ \frac{N-i}{N} & \text{if } q = \frac{1}{2} \end{cases}$$

-
$$\begin{aligned} P_i + Q_i &= \frac{1-(q/p)^i}{1-(q/p)^N} + \frac{1-(p/q)^{N-i}}{1-(p/q)^N} \\ &= \frac{p^N - p^N(q/p^i)}{p^N - q^N} + \frac{q^N - q^N(p/q)^{N-i}}{q^N - p^N} \\ &= \frac{p^N - p^{N-i}q^i - q^N + q^i p^{N-i}}{p^N - q^N} \\ &= 1 \end{aligned}$$

- If A were to start with 5 units and B with 10, then the probability of A's winning would be $\frac{1}{3}$, whereas it would jump to
$$\frac{1-(\frac{2}{3})^5}{1-(\frac{2}{3})^{15}} \approx .87$$
if $p = .6$.

Example 3.4k. The complete graph having n vertices is defined to be a set of n points (called vertices) in the plane and the $\binom{n}{2}$ lines (called edges) connecting each pair of vertices. The complete graph having 3 vertices is shown in Figure 3.3. Suppose now that each edge in a complete graph on n vertices is to be colored either red or blue. For a fixed integer k, a question of interest is whether there is a way of coloring the edges so that no set of k vertices has all of its $\binom{k}{2}$ connecting edges the same color. It can be shown, by a probabilistic argument, that if n is not too large, then the answer is yes.

- Suppose that each edge is equally likely to be colored either red or blue.
- Number the $\binom{n}{k}$ sets of k vertices.
  E_i = {all of the connecting edges of the
  ith set of k vertices are the same color}
- The probability that they are all the same color is
$$P(E_i) = 2\left(\frac{1}{2}\right)^{k(k-1)/2}$$
- $P\left(\bigcup_i E_i\right) \leq \binom{n}{k}\left(\frac{1}{2}\right)^{k(k-1)/2-1}$
- If $\binom{n}{k}(\frac{1}{2})^{k(k-1)/2-1} < 1$ or $\binom{n}{k} < 2^{k(k-1)/2-1}$ then the probability that at least one of the $\binom{n}{k}$ sets of k vertices has all of its connecting edges of the same color is less than 1.
- There is a positive probability that no set of k vertices has all of its connecting edges the same color.

*3.5 $P(\cdot|F)$ is a probability

Proposition 5.1:

(a) $0 \le P(E|F) \le 1$.

(b) $P(S|F) = 1$.

(c) If E_i, $i = 1, 2, \ldots$ are mutually exclusive events, then

$$P\left(\bigcup_1^\infty E_i|F\right) = \sum_1^\infty P(E_i|F)$$

- $Q(E) = P(E|F)$
- $Q(E_1 \bigcup E_2) = Q(E_1) + Q(E_2)$
- $P(E_1 \bigcup E_2|F) = P(E_1|F) + P(E_2|F) - P(E_1E_2|F)$
- $Q(E_1|E_2) = \dfrac{Q(E_1E_2)}{Q(E_2)}$
- $Q(E_1) = Q(E_1|E_2)Q(E_2)Q(E_1|E_2^c)Q(E_2^c)$
- $P(E_1|F) = P(E_1|E_2F)P(E_2|F) + P(E_1|E_2^cF)P(E_2^c|F)$

Example 3.5a. Consider Example 3a, which is concerned with an insurance company that believes that people can be divided into two distinct classes-those who are accident prone and those who are not. During any given year an accident-prone person will have an accident with probability .4, whereas the corresponding figure for a non-accident-prone person is .2. What is the conditional probability that a new policyholder will have an accident in his or her second year of policy ownership, given that the policyholder has had an accident in the year?

- A: The event that the policyholder is accident prone.
- A_i: The event that he or she has had an accident in the ith year.

- The desired probability:

$$P(A_2|A_1) = P(A_2|AA_1)P(A|A_1) + P(A_2|A^cA_1)P(A^c|A_1)$$

- Now,

$$P(A|A_1) = \frac{P(A_1A)}{P(A_1)} = \frac{P(A_1|A)P(A)}{P(A_1)}$$

- $P(A|A_1) = \frac{(.4)(.3)}{.26} = \frac{6}{13}$ and
  $P(A^c|A_1) = 1 - P(A|A_1) = \frac{7}{13}$
- Since $P(A_2|AA_1) = .4$ and $P(A_2|A^cA_1) = .2$, we see that
  $P(A_2|A_1) = (.4)\frac{6}{13} + (.2)\frac{7}{13} \approx .29$

Example 3.5b. Independent trials, each resulting in a success with probability p or a failure with probability $q = 1 - p$ are performed. We are interested in computing the probability that a run of n consecutive successes occurs before a run of m consecutive failures.

- E: The event that a run of n consecutive successes occurs before a run of m consecutive failures.
- H: The event that the first trial results in a success.
- $P(E) = pP(E|H) + qP(E|H^c)$
- F: The event that trials 2 through n all are successes.

$$P(E|H) = P(E|FH)P(F|E) + P(E|F^cH)P(F^c|H)$$

- $P(E|F^cH) = P(E|H^c)$
- $P(E|H) = p^{n-1} + (1 - p^{n-1})P(E|H^c)$
- G: The event that trials 2 through m all are failures.

$$P(E|H^c) = P(E|GH^c)P(G|H^c) + P(E|G^cH^c)P(G^c|H^c)$$

- $P(E|G^cH^c) = P(E|H)$
- $P(E|H^c) = (1 - q^{m-1})P(E|H)$
- $P(E|H) = \frac{p^{n-1}}{p^{n-1}+q^{m-1}-p^{n-1}q^{m-1}}$
- $P(E|H^c) = \frac{(1-q^{m-1})p^{n-1}}{p^{n-1}+q^{m-1}-p^{n-1}q^{m-1}}$

$$\begin{aligned} P(E) &= pP(E|H) + qP(E|H^c) \\ &= \frac{p^n + qp^{n-1}(1 - q^{m-1})}{p^{n-1} + q^{m-1} - p^{n-1}q^{m-1}} \\ &= \frac{p^{n-1}(1 - q^m)}{p^{n-1} + q^{m-1} - p^{n-1}q^{m-1}} \end{aligned}$$

- P{run of m failures before a run of n successes}
  $= \frac{q^{m-1}(1-p^n)}{q^{m-1}+p^{n-1}-q^{m-1}p^{n-1}}$

Example 3.5c. At a party n men take off their hats. The hats are then mixed up, and each man randomly selects one. We say that a match occurs if a man selects his own hat. What is the probability of

(a) no matches;

(b) exactly k matches?

(a)

- E: The event that no matches occur.
- $P_n = P(E)$
- M: A man selects his own hat.
- $P_n = P(E) = P(E|M)P(M) + P(E|M^c)P(M^c)$
- $P(E|M) = 0$
- $P_n = P(E|M^c)\frac{n-1}{n}$
- $P(E|M^c) = P_{n-1} + \frac{1}{n-1}P_{n-2}$
- $P_n = \frac{n-1}{n}P_{n-1} + \frac{1}{n}P_{n-2}$
- $P_n - P_{n-1} = -\frac{1}{n}(P_{n-1} - P_{n-2})$
- We have $P_1 = 0$, $P_2 = \frac{1}{2}$
  $P_3 - P_2 = -\frac{(P_2-P_1)}{3} = -\frac{1}{3!}$
  or $P_3 = \frac{1}{2!} - \frac{1}{3!}$
  $P_4 - P_3 = -\frac{(P_3-P_2)}{3} = \frac{1}{4!}$
  or $P_4 = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}$
- $P_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + \frac{(-1)^n}{n!}$

(b)

- $\frac{1}{n}\frac{1}{n-1}\cdots\frac{1}{n-(k-1)}P_{n-k} = \frac{(n-k)!}{n!}P_{n-k}$
- $\frac{P_{n-k}}{k!} = \frac{\frac{1}{2!}-\frac{1}{3!}+\cdots+\frac{(-1)^{n-k}}{(n-k)!}}{k!}$
- $P(E_1|E_2F) = P(E_1|F)$
- $P(E_1E_2|F) = P(E_1|F)P(E_2|F)$

Example 3.5d. *Laplace's rule of succession.* There are $k+1$ coins in a box. The ith coin will, when flipped, turn up heads with probability i/k. A coin is randomly selected from the box and is then repeatedly flipped. If the first n flips all result in heads, what is the conditional probability that the $(n+1)$st flip will do likewise?

- E_i: The event that the ith coin is initially selected.
- F_n: The event that the first n flips all result in heads.
- F: The event that the $(n+1)$st flip is a head.

- The desired probability

$$P(F|F_n) = \sum_{i=0}^{k} P(F|F_nE_i)P(E_i|F_n)$$

- $P(F|F_nE_i) = P(F|E_i) = \frac{i}{k}$
  Also,

$$\begin{aligned} P(E_i|F_n) &= \frac{P(E_iF_n)}{P(F_n)} \\ &= \frac{P(F_n|E_i)P(E_i)}{\sum_{j=0}^{k} P(F_n|E_j)P(E_j)} \\ &= \frac{(\frac{i}{k})^n \left[\frac{1}{k+1}\right]}{\sum_{j=0}^{k}(\frac{j}{k})^n \left[\frac{1}{k+1}\right]} \end{aligned}$$

- Hence

$$P(F|F_n) = \frac{\sum_{i=0}^{k} \left(\frac{i}{k}\right)^{n+1}}{\sum_{j=0}^{k} \left(\frac{j}{k}\right)^{n}}$$

- If k is large, use the integral approximations

$$\frac{1}{k}\sum_{i=0}^{k}\left(\frac{i}{k}\right)^{n+1} \approx \int_0^1 x^{n+1}dx = \frac{1}{n+2}$$

$$\frac{1}{k}\sum_{j=0}^{k}\left(\frac{j}{k}\right)^{n} \approx \int_0^1 x^{n}dx = \frac{1}{n+1}$$

- For k large,

$$P(F|F_n) \approx \frac{n+1}{n+2}$$

Summary

- Conditional probability of E given F:

$$P(E|F) = \frac{P(EF)}{P(F)}$$

- Multiplication rule of probability:

$$P(E_1E_2\cdots E_n) = P(E_1)P(E_2|E_1)\cdots P(E_n|E_1E_2\cdots E_{n-1})$$

- Compute $P(E)$ by conditioning on F:

$$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$$

- Odds ratio of event H:

$$\frac{P(H)}{P(H^c)}$$

- Odds ratio of event H given E:

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

- Bayes' formula:

$$P(F_j|E) = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^n P(E|F_i)P(F_i)}$$

- Independence of E and F:

$$P(EF) = P(E)P(F)$$

- Independence of $E_i, i = 1, 2, \ldots, n$:

$$P(E_{i_1}E_{i_2}\cdots E_{i_r}) = P(E_{i_1})P(E_{i_2})\cdots P(E_{i_r}) \quad r = 2, 3, \ldots, n$$