SECTION 6.3 Bayes' Theorem

2. We know that $p(E \mid F) = p(E \cap F)/p(F)$, so we need to find those two quantities. We are given $p(F) = 3/4$. To compute $p(E \cap F)$, we can use the fact that $p(E \cap F) = p(E)p(F \mid E)$. We are given that $p(E) = 2/3$ and that $p(F \mid E) = 5/8$; therefore $p(F \cap E) = (2/3)(5/8) = 5/12$. Putting this together, we have $p(E \mid F) = (5/12)/(3/4) = 5/9$.

4. Let F be the event that Ann picks the second box. Thus we know that $p(F) = p(\overline{F}) = 1/2$. Let B be the event that Frida picks an orange ball. Because of the contents of the boxes, we know that $p(B \mid F) = 5/11$ (five of the eleven balls in the second box are orange) and $p(B \mid \overline{F}) = 3/7$. We are asked for $p(F \mid B)$. We use Bayes' Theorem:

$$p(F \mid B) = \frac{p(B \mid F)p(F)}{p(B \mid F)p(F) + p(B \mid \overline{F})p(\overline{F})} = \frac{(5/11)(1/2)}{(5/11)(1/2) + (3/7)(1/2)} = \frac{35}{68}$$

6. Let S be the event that a randomly chosen soccer player uses steroids. We know that $p(S) = 0.05$ and therefore $p(\overline{S}) = 0.95$. Let P be the event that a randomly chosen person tests positive for steroid use. We are told that $p(P \mid S) = 0.98$ and $p(P \mid \overline{S}) = 0.12$ (this is a "false positive" test result). We are asked for $p(S \mid P)$. We use Bayes' Theorem:

$$p(S \mid P) = \frac{p(P \mid S)p(S)}{p(P \mid S)p(S) + p(P \mid \overline{S})p(\overline{S})} = \frac{(0.98)(0.05)}{(0.98)(0.05) + (0.12)(0.95)} \approx 0.301$$

8. Let D be the event that a randomly chosen person has the rare genetic disease. We are told that $p(D) = 1/10000 = 0.0001$ and therefore $p(\overline{D}) = 0.9999$. Let P be the event that a randomly chosen person tests positive for the disease. We are told that $p(P \mid D) = 0.999$ ("true positive") and that $p(P \mid \overline{D}) = 0.0002$ ("false positive"). From these we can conclude that $p(\overline{P} \mid D) = 0.001$ ("false negative") and $p(\overline{P} \mid \overline{D}) = 0.9998$ ("true negative").

a) We are asked for $p(D \mid P)$. We use Bayes' Theorem:

$$p(D \mid P) = \frac{p(P \mid D)p(D)}{p(P \mid D)p(D) + p(P \mid \overline{D})p(\overline{D})} = \frac{(0.999)(0.0001)}{(0.999)(0.0001) + (0.0002)(0.9999)} \approx 0.333$$

b) We are asked for $p(\overline{D} \mid \overline{P})$. We use Bayes' Theorem:

$$p(\overline{D} \mid \overline{P}) = \frac{p(\overline{P} \mid \overline{D})p(\overline{D})}{p(\overline{P} \mid \overline{D})p(\overline{D}) + p(\overline{P} \mid D)p(D)} = \frac{(0.9998)(0.9999)}{(0.9998)(0.9999) + (0.001)(0.0001)} \approx 1.000$$

(This last answer is exactly $49985001/49985006 \approx 0.99999989997$.)

10. Let A be the event that a randomly chosen person in the clinic is infected with avian influenza. We are told that $p(A) = 0.04$ and therefore $p(\overline{A}) = 0.96$. Let P be the event that a randomly chosen person tests positive for avian influenza on the blood test. We are told that $p(P \mid A) = 0.97$ and $p(P \mid \overline{A}) = 0.02$ ("false positive"). From these we can conclude that $p(\overline{P} \mid A) = 0.03$ ("false negative") and $p(\overline{P} \mid \overline{A}) = 0.98$.

a) We are asked for $p(A \mid P)$. We use Bayes' Theorem:

$$p(A \mid P) = \frac{p(P \mid A)p(A)}{p(P \mid A)p(A) + p(P \mid \overline{A})p(\overline{A})} = \frac{(0.97)(0.04)}{(0.97)(0.04) + (0.02)(0.96)} \approx 0.669$$

b) In part (a) we found $p(A \mid P)$. Here we are asked for the probability of the complementary event (given a positive test result). Therefore we have simply $p(\overline{A} \mid P) = 1 - p(A \mid P) \approx 1 - 0.669 = 0.331$.

c) We are asked for $p(A \mid \overline{P})$. We use Bayes' Theorem:

$$p(A \mid \overline{P}) = \frac{p(\overline{P} \mid A)p(A)}{p(\overline{P} \mid A)p(A) + p(\overline{P} \mid \overline{A})p(\overline{A})} = \frac{(0.03)(0.04)}{(0.03)(0.04) + (0.98)(0.96)} \approx 0.001$$

d) In part (c) we found $p(A \mid \overline{P})$. Here we are asked for the probability of the complementary event (given a negative test result). Therefore we have simply $p(\overline{A} \mid \overline{P}) = 1 - p(A \mid \overline{P}) \approx 1 - 0.001 = 0.999$.

12. Let E be the event that a 0 was received; let F_1 be the event that a 0 was sent; and let F_2 be the event that a 1 was sent. Note that $F_2 = \overline{F}_1$. Then we are told that $p(F_2) = 1/3$, $p(F_1) = 2/3$, $p(E \mid F_1) = 0.9$, and $p(E \mid F_2) = 0.2$.

a) $p(E) = p(E \mid F_1)p(F_1) + p(E \mid F_2)p(F_2) = 0.9 \cdot (2/3) + 0.2 \cdot (1/3) = 2/3$.

b) We use Bayes' Theorem:

$$p(F_1 \mid E) = \frac{p(E \mid F_1)p(F_1)}{p(E \mid F_1)p(F_1) + p(E \mid F_2)p(F_2)} = \frac{0.9 \cdot (2/3)}{0.9 \cdot (2/3) + 0.2 \cdot (1/3)} = \frac{0.6}{2/3} = 0.9$$

14. By Generalized Bayes' Theorem,

$$\begin{aligned} p(F_2 \mid E) &= \frac{p(E \mid F_2)p(F_2)}{p(E \mid F_1)p(F_1) + p(E \mid F_2)p(F_2) + p(E \mid F_3)p(F_3)} \\ &= \frac{(3/8)(1/2)}{(2/7)(1/6) + (3/8)(1/2) + (1/2)(1/3)} = \frac{7}{15}. \end{aligned}$$

16. Let L be the event that Ramesh is late, and let B, C, and O (standing for "omnibus") be the events that he went by bicycle, car, and bus, respectively. We are told that $p(L \mid B) = 0.05$, $p(L \mid C) = 0.50$, and $p(L \mid O) = 0.20$. We are asked to find $p(C \mid L)$.

a) We are to assume here that $p(B) = p(C) = p(O) = 1/3$. Then by Generalized Bayes' Theorem,

$$\begin{aligned} p(C \mid L) &= \frac{p(L \mid C)p(C)}{p(L \mid B)p(B) + p(L \mid C)p(C) + p(L \mid O)p(O)} \\ &= \frac{(0.50)(1/3)}{(0.05)(1/3) + (0.50)(1/3) + (0.20)(1/3)} = \frac{2}{3}. \end{aligned}$$

b) Now we are to assume here that $p(B) = 0.60$, $p(C) = 0.30$, $p(O) = 0.10$. Then by Generalized Bayes' Theorem,

$$\begin{aligned} p(C \mid L) &= \frac{p(L \mid C)p(C)}{p(L \mid B)p(B) + p(L \mid C)p(C) + p(L \mid O)p(O)} \\ &= \frac{(0.50)(0.30)}{(0.05)(0.60) + (0.50)(0.30) + (0.20)(0.10)} = \frac{3}{4}. \end{aligned}$$

18. We follow the procedure in Example 3. We first compute that $p(\text{exciting}) = 40/500 = 0.08$ and $q(\text{exciting}) = 25/200 = 0.125$. Then we compute that

$$r(\text{exciting}) = \frac{p(\text{exciting})}{p(\text{exciting}) + q(\text{exciting})} = \frac{0.08}{0.08 + 0.125} \approx 0.390\,.$$

Because $r(\text{exciting})$ is less than the threshold 0.9, an incoming message containing "exciting" would not be rejected.

20. **a)** We follow the procedure in Example 3. In Example 4 we found $p(\text{undervalued}) = 0.1$ and $q(\text{undervalued}) = 0.025$. So we compute that

$$r(\text{undervalued}) = \frac{p(\text{undervalued})}{p(\text{undervalued}) + q(\text{undervalued})} = \frac{0.01}{0.01 + 0.025} \approx 0.286\,.$$

Because $r(\text{undervalued})$ is less than the threshold 0.9, an incoming message containing "undervalued" would not be rejected.

b) This is similar to part (a), where $p(\text{stock}) = 0.2$ and $q(\text{stock}) = 0.06$. Then we compute that

$$r(\text{stock}) = \frac{p(\text{stock})}{p(\text{stock}) + q(\text{stock})} = \frac{0.2}{0.2 + 0.06} \approx 0.769\,.$$

Because $r(\text{stock})$ is less than the threshold 0.9, an incoming message containing "stock" would not be rejected. Notice that each event alone was not enough to cause rejection, but both events together were enough (see Example 4).

22. **a)** Out of a total of $s + h$ messages, s are spam, so $p(S) = s/(s+h)$. Similarly, $p(\overline{S}) = h/(s+h)$.

b) Let W be the event that an incoming message contains the word w. We are told that $p(W \mid S) = p(w)$ and $p(W \mid \overline{S}) = q(w)$. We want to find $p(S \mid W)$. We use Bayes' Theorem:

$$p(S \mid W) = \frac{p(W \mid S)p(S)}{p(W \mid S)p(S) + p(W \mid \overline{S})p(\overline{S})} = \frac{p(w)\frac{s}{(s+h)}}{p(w)\frac{s}{(s+h)} + q(w)\frac{h}{(s+h)}} = \frac{p(w)s}{p(w)s + q(w)h}$$

The assumption made in this section was that $s = h$, so those factors cancel out of this answer to give the formula for $r(w)$ obtained in the text.