$p(E \cap F) + p(E \cap \overline{F})$. We have already shown that $p(E \cap F) = p(E \mid F)p(F)$. Moreover, we have $p(E \mid \overline{F}) = p(E \cap \overline{F})/p(\overline{F})$, which shows that $p(E \cap \overline{F}) = p(E \mid \overline{F})p(\overline{F})$. It follows that

$$p(E) = p(E \cap F) + p(E \cap \overline{F}) = p(E \mid F)p(F) + p(E \mid \overline{F})p(\overline{F}).$$

To complete the proof we insert this expression for $p(E)$ into the equation $p(F \mid E) = p(E \mid F)p(F)/p(E)$. We have proved that

Links

$$p(F \mid E) = \frac{p(E \mid F)p(F)}{p(E \mid F)p(F) + p(E \mid \overline{F})p(\overline{F})}. \quad \triangleleft$$

Bayes' Theorem can be used to assess the probability that someone testing positive for a disease actually has this disease. The results obtained from Bayes' Theorem are often somewhat surprising, as Example 2 shows.

EXAMPLE 2 Suppose that one person in 100,000 has a particular rare disease for which there is a fairly accurate diagnostic test. This test is correct 99% of the time when given to someone with the disease; it is correct 99.5% of the time when given to someone who does not have the disease. Given this information can we find

(a) the probability that someone who tests positive for the disease has the disease?
(b) the probability that someone who tests negative for the disease does not have the disease?

Should someone who tests positive be very concerned that he or she has the disease?

Solution: (a) Let F be the event that a person has the disease, and let E be the event that this person tests positive for the disease. We want to compute $p(F \mid E)$. To use Bayes' Theorem to compute $p(F \mid E)$ we need to find $p(E \mid F)$, $p(E \mid \overline{F})$, $p(F)$, and $p(\overline{F})$.

We know that one person in 100,000 has this disease, so that $p(F) = 1/100{,}000 = 0.00001$ and $p(\overline{F}) = 1 - 0.00001 = 0.99999$. Because someone who has the disease tests positive 99% of the time, we know that $p(E \mid F) = 0.99$; this is the probability of a true positive, that someone with the disease tests positive. We also know that $p(\overline{E} \mid F) = 0.01$; this is the probability of a false negative, that someone who has the disease tests negative. Furthermore, because someone who does not have the disease tests negative 99.5% of the time, we know that $p(\overline{E} \mid \overline{F}) = 0.995$. This is the probability of a true negative, that someone without the disease tests negative. Finally, we know that $p(E \mid \overline{F}) = 0.005$; this is the probability of a false positive, that someone without the disease tests positive.

The probability that someone who tests positive for the disease actually has the disease is $p(F \mid E)$. By Bayes' Theorem, we know that

$$\begin{aligned} p(F \mid E) &= \frac{p(E \mid F)p(F)}{p(E \mid F)p(F) + p(E \mid \overline{F})p(\overline{F})} \\ &= \frac{(0.99)(0.00001)}{(0.99)(0.00001) + (0.005)(0.99999)} \approx 0.002. \end{aligned}$$

This means that only 0.2% of people who test positive for the disease actually have the disease. Because the disease is extremely rare, the number of false positives on the diagnostic test is far greater than the number of true positives, making the percentage of people who test positive