Problem 1. After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease and that the test is 99% accurate (i.e., the probability of testing positive when you do have the disease is 0.99, as is the probability of testing negative when you don’t have the disease). The good news is that this is a rare disease, striking only 1 in 10,000 people of your age. Why is it good news that the disease is rare? What are the chances that you actually have the disease?

Solution .. We are given the following information:

\[
\begin{align*}
P(\text{test}|\text{disease}) &= 0.99 \\
P(\neg\text{test}|\neg\text{disease}) &= 0.99 \\
P(\text{disease}) &= 0.0001
\end{align*}
\]

and the observation test. What the patient is concerned about it \(P(\text{disease}|\text{test})\). Roughly speaking, the reason is a good thing that the disease is rare is that \(P(\text{disease}|\text{test})\) is proportional to \(P(\text{disease})\), so a lower prior for disease will mean a lower value for \(P(\text{disease}|\text{test})\).

Roughly speaking, if 10,000 people take the test, we expect 1 to actually have the disease, and most likely test positive, while the rest do not have the disease, but 1% of them (about 100 people) will test positive anyway, so \(P(\text{disease}|\neg\text{test})\) will be about 1 in 100. More precisely, using the normalization equation:

\[
P(\text{disease}|\text{test}) = \frac{P(\text{test}|\text{disease})P(\text{disease})}{P(\text{test}|\text{disease})P(\text{disease}) + P(\text{test}|\neg\text{disease})P(\neg\text{disease})} = \frac{0.99 \times 0.0001}{0.99 \times 0.0001 + 0.01 \times 0.9999} = 0.009804
\]

The moral is that when the disease is much rarer than the test accuracy, a positive test result does not mean the disease is likely. A false positive reading remains much more likely. Here is an alternative exercise along the same lines: A doctor says that an infant who predominantly turns the head to the right while lying on the back will be right-handed, and one who turns to the left will be left-handed. Isabella predominantly turned her head to the left. Given that 90% of the population is right-handed, what is Isabellas probability of being right-handed if the test is 90% accurate? If it is 80% accurate? The reasoning is the same, and the answer is 50% right-handed if the test is 90% accurate, 69% right-handed if the test is 80% accurate.

Problem 2. Considering the set of all possible five-card poker hands dealt fairly from a standard deck of fifty-two card.

Solution .. This is a classic combinatorics question that could appear in a basic text on discrete mathematics. The point here is to refer to the relevant axioms of probability. The
question also helps students to grasp the concept of the joint probability distribution as the distribution over all possible states of the world.

(a) How many atomic events are there in the joint probability distribution (i.e., how many five-cards hands are there)?

Solution .. There are \( \binom{52}{5} = \frac{52 \times 51 \times 50 \times 49 \times 48}{1 \times 2 \times 3 \times 4 \times 5} = 2,598,960 \) possible five-cards hands.

(b) What is the probability of each atomic event?

Solution .. By the fair-dealing assumption, each of there is equally likely. By axioms 2 and 3, each hand therefore occurs with probability \( \frac{1}{2,598,960} \).

(c) What is the probability of being dealt a royal straight flush? Four of a kind?

Solution .. There are four hands that are royal straight flushes (one in each suit). By axiom 3, since the events are mutually exclusive, the probability of a royal straight flush is just the sum of the probabilities of the atomic events, i.e., \( \frac{4}{2,598,960} = \frac{1}{649,740} \). For four of a kind events, there are 13 possible kinds and for each, the fifth card can be one of 48 possible other cards. The total probability is therefore \( \frac{13 \times 48}{2,598,960} = \frac{1}{165} \).

Problem 3 (Optional). Suppose you are given a coin that lands heads with probability \( x \) and tails with probability \( 1 - x \). Are the outcomes of successive flips of the coin independent of each other given that you know the value of \( x \)? Are the outcomes of successive flips of the coin independent of each other if you do not know the value of \( x \)? Justify your answer.

Solution .. If the probability \( x \) is known, then successive flips of the coin are independent of each other, since we know that each flip of the coin will land heads with probability \( x \). Formally, if \( F_1 \) and \( F_2 \) represent the results of two successive flips, we have \( P(F_1 = \text{heads}, F_2 = \text{heads} | x) = x \times x = P(F_1 = \text{heads} | x)P(F_2 = \text{heads} | x) \). Thus, the events \( F_1 = \text{heads} \) and \( F_2 = \text{heads} \) are independent. If we do not know the value of \( x \), however, the probability of each successive flip is dependent on the result of all previous flips. The reason for this is that each successive flip gives us information to better estimate the probability \( x \) (i.e., determining the posterior estimate for \( x \) given our prior probability and the evidence we see in the most recent coin flip). This new estimate of \( x \) would then be used as our best guess of the probability of the coin coming up heads on the next flip. Since this estimate for \( x \) is based on all the previous flips we have seen, the probability of the next flip coming up heads depends on how many heads we saw in all previous flips, making them dependent. For example, if we had a uniform prior over the probability \( x \), then one can show that after \( n \) flips if \( m \) of them come up heads then the probability that the next one comes up heads is \( \frac{m + 1}{n + 2} \), showing dependence on previous flips.

Problem 4 (Optional). Show that the three forms of independence in Equation 1 are equivalent.

\[
P(a|b) = P(a) \text{ or } P(b|a) = P(b) \text{ or } P(a \wedge b) = P(a)P(b) \tag{1}
\]
Solution .. Independence is symmetric (that is, $a$ and $b$ are independent iff $b$ and $a$ are independent) so $P(a|b) = P(a)$ is the same as $P(b|a) = P(b)$. So we need only prove that $P(a|b) = P(a)$ is equivalent to $P(a \land b) = P(a)P(b)$. The product rule, $P(a \land b) = P(a|b)P(b)$, can be used to rewrite $P(a \land b) = P(a)P(b)$ as $P(a|b)P(b) = P(a)P(b)$, which simplifies to $P(a|b) = P(a)$.

Problem 5 (Optional). We wish to transmit an $n$-bit message to a receiving agent. The bits in the message are independently corrupted (flipped) during transmission with $\epsilon$ probability each. With an extra parity bit send along with the original information, a message can be corrected by the receiver if at most one bit in the entire message (including the parity bit) has been corrupted. Suppose we want to ensure that the correct message is received with probability at least $1 - \delta$. What is the maximum feasible value of $n$? Calculate this value for the case $\epsilon = 0.001, \delta = 0.01$.

Solution .. The correct message is received if either zero or one of the $n + 1$ bits are corrupted. Since corruption occurs independently with probability $\epsilon$, the probability that zero bits are corrupted is $(1 - \epsilon)^{n+1}$. There are $n + 1$ mutually exclusive ways that exactly one bit can be corrupted, one for each bit in the message. Each has probability $\epsilon(1 - \epsilon)^n$, so the overall probability that exactly one bit is corrupted is $n\epsilon(1 - \epsilon)^n$. Thus, the probability that the correct message is received is $(1 - \epsilon)^{n+1} + n\epsilon(1 - \epsilon)^n$. The maximum feasible value of $n$, therefore, is the largest $n$ satisfying the inequality $(1 - \epsilon)^{n+1} + n\epsilon(1 - \epsilon)^n \geq 1 - \delta$.

Numerically solving this for $\epsilon = 0.001, \delta = 0.01$ , we find $n = 147$.  

3