Review: Main Theorems and Concepts

Conditional Probability: $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

Independence: Events $E$ and $F$ are independent iff $\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F)$, or equivalently $\mathbb{P}(F) = \mathbb{P}(F|E)$, or equivalently $\mathbb{P}(E) = \mathbb{P}(E|F)$

Bayes Theorem: $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$

Partition: Nonempty events $E_1, \ldots, E_n$ partition the sample space $\Omega$ iff

- $E_1, \ldots, E_n$ are exhaustive: $E_1 \cup E_2 \cup \cdots \cup E_n = \bigcup_{i=1}^n E_i = \Omega$, and
- $E_1, \ldots, E_n$ are pairwise mutually exclusive: $\forall i \neq j, E_i \cap E_j = \emptyset$
  
  - Note that for any event $A$ (with $A \neq \emptyset, A \neq \Omega$): $A$ and $\overline{A}$ partition $\Omega$

Law of Total Probability (LTP): Suppose $A_1, \ldots, A_n$ partition $\Omega$ and let $B$ be any event. Then

$$\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(B \cap A_i) = \sum_{i=1}^n \mathbb{P}(B | A_i)\mathbb{P}(A_i)$$

Bayes Theorem with LTP: Suppose $A_1, \ldots, A_n$ partition $\Omega$ and let $B$ be any event. Then $\mathbb{P}(A_1|B) = \frac{\mathbb{P}(B | A_1)\mathbb{P}(A_1)}{\sum_{i=1}^n \mathbb{P}(B | A_i)\mathbb{P}(A_i)}$. In particular, $\mathbb{P}(A|B) = \frac{\mathbb{P}(B | A)\mathbb{P}(A)}{\mathbb{P}(B | A)\mathbb{P}(A) + \mathbb{P}(B | \overline{A})\mathbb{P}(A)}$

Chain Rule: Suppose $A_1, \ldots, A_n$ are events. Then

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1)\mathbb{P}(A_3 | A_1 \cap A_2) \ldots \mathbb{P}(A_n | A_1 \cap \ldots \cap A_{n-1})$$

Exercises

1. Suppose we randomly generate a number from the natural numbers $\mathbb{N} = \{1, 2, \ldots\}$. Let $A_k$ be the event we generate the number $k$, and suppose $\mathbb{P}(A_k) = \left(\frac{1}{2}\right)^k$. Once we generate a number, suppose the probability that we win $\$j$ for $j = 1, \ldots, k$ is “uniform”, that is, each has probability $\frac{1}{j}$. Let $B$ be the event we win exactly $\$1$. What is $\mathbb{P}(A_1|B)$? You may use the fact that $\sum_{j=1}^\infty \frac{1}{j} = \ln(\frac{a}{a-1})$ for $a > 1$.

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(B|A_1)\mathbb{P}(A_1)}{\sum_{j=1}^\infty \mathbb{P}(B|A_j)\mathbb{P}(A_j)} = \frac{1 \cdot \frac{1}{2^1}}{\sum_{j=1}^\infty \frac{1}{j} \frac{1}{2^j}} = \frac{1}{2\ln 2} \approx 0.7213$$
2. Suppose there are three possible teachers for CSE 312: Martin Tompa, Anna Karlin, and Larry Ruzzo. Suppose the ratio of grades \(A:B:C:D:F\) for Martin’s class is \(1:2:3:4:5\), for Anna’s class is \(3:4:5:1:2\), and for Larry’s class is \(5:4:3:2:1\). Suppose you are assigned a grade randomly according to the given ratios when you take a class from one of these professors, irrespective of your performance. Furthermore, suppose Martin teaches your class with probability \(\frac{1}{2}\) and Anna and Larry have an equal chance of teaching if Martin isn’t. What is the probability you had Martin, given that you received an \(A\)? Compare this to the unconditional probability that you had Martin.

Let \(T, K, R\) be the events you take 312 from Tompa, Karlin, and Ruzzo, respectively. Let the letter grades be events themselves.

\[
P(T|A) = \frac{P(A|T)P(T)}{P(A|T)P(T) + P(A|K)P(K) + P(A|R)P(R)} = \frac{\frac{1}{15} \cdot \frac{1}{2}}{\frac{1}{15} \cdot \frac{1}{2} + \frac{3}{15} \cdot \frac{1}{4} + \frac{5}{15} \cdot \frac{1}{4}} = \frac{2}{10} = \frac{1}{5}
\]

3. Suppose we have a coin with probability \(p\) of heads. Suppose we flip this coin \(n\) times independently. Let \(X\) be the number of heads that we observe. What is \(P(X = k)\), for \(k = 0, \ldots, n\)? Verify that \(\sum_{k=0}^{n} P(X = k) = 1\), as it should.

\[
P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}
\]

For a given sequence with exactly \(k\) heads, the probability of that sequence is \(p^k(1 - p)^{n-k}\). However, there are \(\binom{n}{k}\) such sequences, so the probability of exactly \(k\) heads is \(\binom{n}{k}p^k(1 - p)^{n-k}\).

\[
\sum_{k=0}^{n} P(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1
\]

The middle equality uses the Binomial Theorem.

4. Suppose we have a coin with probability \(p\) of heads. Suppose we flip this coin until we flip a head for the first time. Let \(X\) be the number of times we flip the coin up to and including the first head. What is \(P(X = k)\), for \(k = 1, 2, \ldots\)? Verify that \(\sum_{k=1}^{\infty} P(X = k) = 1\), as it should.

\[
P(X = k) = (1 - p)^{k-1} p
\]

If the \(k\)th flip is our first head, the first \(k - 1\) must be tails. Then the \(k\)th flip must be a head.
\[
\sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} (1 - p)^{k-1} p = p \sum_{j=0}^{\infty} (1 - p)^j = \frac{p}{1 - (1 - p)} = 1
\]

5. Corrupted by their power, the judges running the popular game show America’s Next Top Mathematician have been taking bribes from many of the contestants. During each of two episodes, a given contestant is either allowed to stay on the show or is kicked off. If the contestant has been bribing the judges, she will be allowed to stay with probability 1, independent of what happens in earlier episodes. If the contestant has not been bribing the judges, she will be allowed to stay with probability 1/3, independent of what happens in earlier episodes. Notice that this is a “conditional independence”: conditioned on not bribing the judges, the outcomes of the two episodes are independent. Suppose that 1/4 of the contestants have been bribing the judges. The same contestants bribe the judges in both rounds.

(a) If you pick a random contestant, what is the probability that she is allowed to stay during the first episode?

Let \( S_i \) be the event that she stayed during the \( i \)-th episode. By the Law of Total Probability,

\[
P(S_1) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{2}
\]

(b) If you pick a random contestant, what is the probability that she is allowed to stay during both episodes?

Let \( B \) be the event that she is bribing the judges. By the Law of Total Probability and conditional independence,

\[
P(S_1 \cap S_2) = P(B)P(S_1 \cap S_2 | B) + P(\overline{B})P(S_1 \cap S_2 | \overline{B})
\]

\[
= P(B)P(S_1 | B)P(S_2 | B) + P(\overline{B})P(S_1 | \overline{B})P(S_2 | \overline{B})
\]

\[
= \frac{1}{4} \cdot 1 \cdot 1 + \frac{3}{4} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{3}
\]

(c) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that she gets kicked off during the second episode?

By the definition of conditional probability and the Law of Total Probability and conditional independence,

\[
P(\overline{S}_2 | S_1) = \frac{P(S_1 \cap \overline{S}_2)}{P(S_1)} = \frac{\frac{1}{4} \cdot 1 \cdot 0 + \frac{3}{4} \cdot \frac{1}{3} \cdot \frac{2}{3}}{\frac{1}{2}} = \frac{1}{6} \cdot \frac{3}{1} = \frac{1}{2}
\]
(d) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that she was bribing the judges?

Let $B$ be the event that she bribed the judges. By Bayes’ Theorem,

$$P(B | S_1) = \frac{P(S_1 | B)P(B)}{P(S_1)} = \frac{1 \cdot \frac{1}{2}}{\frac{1}{2}} = \frac{1}{2}$$

6. A parallel system functions whenever at least one of its components works. Consider a parallel system of $n$ components and suppose that each component works with probability $p$ independently.

(a) If the system is functioning, what is the probability that component 1 is working?

$$P = \frac{p}{1 - (1 - p)^n}$$

(b) If the system is functioning and component 2 is working, what is the probability that component 1 is working?

$$P$$

7. A girl has 5 blue and 3 white marbles in her left pocket, and 4 blue and 4 white marbles in her right pocket. If she transfers a randomly chosen marble from left pocket to right pocket without looking, and then draws a randomly chosen marble from her right pocket, what is the probability that it is blue?

By the Law of Total Probability,

$$\frac{5}{8} \cdot \frac{5}{9} + \frac{3}{8} \cdot \frac{4}{9} = \frac{37}{72}$$

8. In a certain population, everyone is equally susceptible to colds. The number of colds suffered by each person during each winter season ranges from 0 to 4, with probability 0.2 for each value (see table below). A new cold prevention drug is introduced that, for people for whom the drug is effective, changes the probabilities as shown in the table. Unfortunately, the effects of the drug last only the duration of one winter season, and the drug is only effective in 20% of people, independently.

<table>
<thead>
<tr>
<th>number of colds</th>
<th>no drug or ineffective</th>
<th>drug effective</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>0.0</td>
</tr>
</tbody>
</table>
(a) Sneezy decides to take the drug. Given that he gets 1 cold that winter, what is the probability that the drug is effective for Sneezy?

Let $E$ be the event that the drug is effective for Sneezy, and $C_i$ be the event that he gets $i$ colds the first winter. By Bayes’ Theorem,

$$P(E | C_1) = \frac{P(C_1 | E)P(E)}{P(C_1 | E)P(E) + P(C_1 | \overline{E})P(\overline{E})} = \frac{0.3 \times 0.2}{0.3 \times 0.2 + 0.2 \times 0.8} = \frac{3}{11}$$

(b) The next year he takes the drug again. Given that he gets 2 colds in this winter, what is the updated probability that the drug is effective for Sneezy?

Let the reduced sample space for part (b) be $C_1$ from part (a). Let $D_i$ be the event that he gets $i$ colds the second winter. By Bayes’ Theorem,

$$P(E | D_2) = \frac{P(D_2 | E)P(E)}{P(D_2 | E)P(E) + P(D_2 | \overline{E})P(\overline{E})} = \frac{0.2 \times \frac{3}{11}}{0.2 \times \frac{3}{11} + 0.2 \times \frac{8}{11}} = \frac{3}{11}$$

(c) The third winter he decides not to bother taking the drug and gets 2 colds. He argues that the drug must not have been effective for him, since he got the same number of colds last year as this year. Comment on his logic.

The posterior probability that the drug is effective is $3/11$. This is greater than the prior probability $1/5$, so the drug probably was effective.

9. Guildenstern has three coins $C_1, C_2, C_3$ in a bag. $C_1$ has $P($heads$) = 1$, $C_2$ has $P($heads$) = 0$, and $C_3$ has $P($heads$) = p$. He takes a random coin from the bag, each coin equally probable, and flips this same coin some number of times.

(a) Suppose $q$ is the conditional probability that he flipped coin $C_1$, given that the flip came up heads. Determine $p$ as a function of $q$.

Let $F_i$ be the event that he flipped coin $C_i$ and $H$ be the event that the flip came up heads. By Bayes’ Theorem,

$$q = P(F_1 | H) = \frac{P(H | F_1)P(F_1)}{P(H | F_1)P(F_1) + P(H | F_2)P(F_2) + P(H | F_3)P(F_3)}$$

$$= \frac{1 \times \frac{1}{3}}{1 \times \frac{1}{3} + 0 \times \frac{1}{3} + p \times \frac{1}{3}}$$

$$= \frac{1}{1 + p}$$

$$p = \frac{1}{q} - 1$$

5
(b) What is the probability that the first \( n \) flips come up tails?

Let \( T \) be the event that he flips \( n \) tails in a row. By the Law of Total Probability,

\[
P(T) = P(T \mid F_1)P(F_1) + P(T \mid F_2)P(F_2) + P(T \mid F_3)P(F_3)
\]

\[
= 0 \times \frac{1}{3} + 1 \times \frac{1}{3} + (1 - p)^n \times \frac{1}{3}
\]

\[
= \frac{1}{3}(1 + (1 - p)^n)
\]

(c) Given that the first \( n \) flips come up tails, what is the probability he flipped \( C_1 \)? \( C_2 \)? \( C_3 \)?

By Bayes’ Theorem,

\[
P(F_1 \mid T) = \frac{P(T \mid F_1)P(F_1)}{P(T)}
\]

\[
= \frac{0 \times \frac{1}{3}}{\frac{1}{3}(1 + (1 - p)^n)}
\]

\[
= 0
\]

\[
P(F_2 \mid T) = \frac{P(T \mid F_2)P(F_2)}{P(T)}
\]

\[
= \frac{1 \times \frac{1}{3}}{\frac{1}{3}(1 + (1 - p)^n)}
\]

\[
= \frac{1}{1 + (1 - p)^n}
\]

\[
P(F_3 \mid T) = \frac{P(T \mid F_3)P(F_3)}{P(T)}
\]

\[
= \frac{(1 - p)^n \times \frac{1}{3}}{\frac{1}{3}(1 + (1 - p)^n)}
\]

\[
= \frac{(1 - p)^n}{1 + (1 - p)^n}
\]

10. Guildenstern has a fair coin and a “magic” coin that comes up heads with probability \( p_1 > \frac{1}{2} \). Suppose he picks a coin at random, with probability \( p_2 \) of choosing the magic coin and \( 1 - p_2 \) of choosing the fair coin, and tosses it \( n \) times. All of the tosses come up heads. He would like to convince Rosencrantz that he flipped the magic coin. Rosencrantz only believes him if the conditional probability that it is the magic coin, given the \( n \) heads, is at least 99%. Derive a function \( n = f(p_1, p_2) \) that gives the minimum number of consecutive heads \( n \) to convince Rosencrantz that Guildenstern flipped the magic coin. Remember that \( n \) must be a positive integer.
Let $M$ be the event that Guildenstern picked the magic coin, and $H$ be the event that he flipped $n$ heads in a row. By Bayes’ Theorem,

$$
P(M \mid H) = \frac{P(H \mid M)P(M)}{P(H \mid M)P(M) + P(H \mid \overline{M})P(\overline{M})}
$$

$$
= \frac{p_1^n p_2}{p_1^n p_2 + (\frac{1}{2})^n (1 - p_2)}
\geq 0.99
$$

$$
0.01 p_1^n p_2 \geq 0.99(1 - p_2)/2^n
$$

$$
(2p_1)^n \geq 99 \cdot \frac{1 - p_2}{p_2} = \frac{99}{p_2} - 99
$$

$$
n = \left\lceil \frac{\log(\frac{99}{p_2} - 99)}{\log(2p_1)} \right\rceil
$$

11. This problem demonstrates that independence can be “broken” by conditioning. Let $D_1$ and $D_2$ be the outcomes of two independent rolls of a fair die. Let $E$ be the event “$D_1 = 1$”, $F$ be the event “$D_2 = 6$”, and $G$ be the event “$D_1 + D_2 = 7$”. Even though $E$ and $F$ are independent, show that

$$
P(E \cap F \mid G) \neq P(E \mid G) P(F \mid G).
$$

$$
P(E \mid G) = P(D_1 = 1 \mid D_1 + D_2 = 7) = 1/6
$$

$$
P(F \mid G) = P(D_2 = 6 \mid D_1 + D_2 = 7) = 1/6
$$

$$
P(E \cap F \mid G) = P(D_1 = 1 \cap D_2 = 6 \mid D_1 + D_2 = 7) = 1/6
$$