CSE 312: Foundations of Computing II

Section 3: Conditional Probability, Bayes Theorem Solutions

1. Review of Main Concepts

(a) **Conditional Probability** (only defined when $Pr(B) > 0$) $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

(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 $\mathbb{P}(E) = \mathbb{P}(E|F)$

(c) **Bayes Theorem**: $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$

(d) **Partition:** Nonempty events $E_1, \dots, E_n$ partition the sample space Ω iff

- $E_1, \dots, E_n$ are exhaustive: $E_1 \cup E_2 \cup \dots \cup E_n = \bigcup_{i=1}^n E_i = \Omega$, and
- $E_1, \dots, E_n$ are pairwise mutually exclusive: $\forall i \neq j, E_i \cap E_j = \emptyset$

(e) **Law of Total Probability (LTP)**: Suppose $A_1, \dots, A_n$ partition Ω and let B be any event. Then $\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(B \cap A_i) = \sum_{i=1}^n \mathbb{P}(\mathrm{B} \mid A_i)\mathbb{P}(A_i)$

(f) **Bayes Theorem with LTP**: Suppose $A_1, \dots, A_n$ partition Ω and let B be any event. Then $\mathbb{P}(A_1|B) = \dfrac{\mathbb{P}(B \mid A_1)\mathbb{P}(A_1)}{\sum_{i=1}^n \mathbb{P}(\mathrm{B} \mid A_i)\mathbb{P}(A_i)}$. In particular, $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(\mathrm{B} \mid A)\mathbb{P}(A) + \mathbb{P}(\mathrm{B} \mid A^C)\mathbb{P}(A^C)}$

2. Naive Bayes Presentation

Some parts of section03 may be a presentation on Naive Bayes... pending further updates!

3. Random Grades?

Suppose there are three possible teachers for CSE 312: Aleks Jovcic, Anna Karlin, and Shayan Oveis Gharan. Suppose the probabilities of getting an A in Aleks's class is $\frac{5}{15}$, for Anna's class is $\frac{3}{15}$, and for Shayan's class is $\frac{1}{15}$. Suppose you are assigned a grade randomly according to the given probabilities when you take a class from one of these professors, irrespective of your performance. Furthermore, suppose Shayan teaches your class with probability $\frac{1}{2}$ and Anna and Aleks have an equal chance of teaching if Shayan isn't. What is the probability you had Shayan, given that you received an A? Compare this to the unconditional probability that you had Shayan.

Solution:

Let J, K, O be the events you take 312 from Jovcic, Karlin, and Oveis Gharan, respectively. Let A be the event you get an A. We use Bayes' theorem with LTP, conditioning on each of J, K, O since those events partition our sample space.

$$\mathbb{P}(O|A) = \frac{\mathbb{P}(A|O)\mathbb{P}(O)}{\mathbb{P}(A|J)\mathbb{P}(J) + \mathbb{P}(A|K)\mathbb{P}(K) + \mathbb{P}(A|O)\mathbb{P}(O)} = \frac{1/15 \cdot 1/2}{5/15 \cdot 1/4 + 3/15 \cdot 1/4 + 1/15 \cdot 1/2} = \frac{2}{5+3+2} = \boxed{\frac{1}{5}}$$

Note that we used Bayes' Theorem because we already know the reverse probability $Pr(A|O)$, which makes it easy for us to evaluate the initial probability $Pr(O|A)$.

4. Coin Flipping

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 $\mathbb{P}(X = k)$, for $k = 0, \ldots n$? Verify that $\sum_{k=0}^{n} \mathbb{P}(X = k) = 1$, as it should.

Solution:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

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$\binom{n}{k}p^k(1-p)^{n-k}$.

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

. The middle equality uses the Binomial Theorem.

5. More Coin Flipping

Suppose we have a coin with probability p of heads. Suppose we flip this coin independently 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 $\mathbb{P}(X = k)$, for $k = 1, 2, \ldots$? Verify that$\sum_{k=1}^{\infty} \mathbb{P}(X = k) = 1$, as it should. (You may use the fact that $\sum_{j=0}^{\infty} a^j = \frac{1}{1-a}$ for$|a| < 1$).

Solution:

$$\mathbb{P}(X = k) = (1-p)^{k-1} p$$

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.

$$\sum_{k=1}^{\infty} \mathbb{P}(X = k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{j=0}^{\infty} (1-p)^j = \frac{p}{1-(1-p)} = 1$$

6. Game Show

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. 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. 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?

Solution:

Let S_i be the event that she stayed during the i-th episode. By the Law of Total Probability conditioning on whether the contestant bribed the judges we get,

$$\mathbb{P}(S_1) = \mathbb{P}(\text{Bribe})\, \mathbb{P}(S_1 \mid \text{Bribe}) + \mathbb{P}(\text{No bribe})\, \mathbb{P}(S_1 \mid \text{No bribe}) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{3} = \boxed{\frac{1}{2}}$$

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

Solution:

Let S_i be defined as before. Staying during both episodes is equivalent to the contestant staying in episodes 1 and 2, so the event $S_1 \cap S_2$. By the Law of Total Probability, we get:

$$\mathbb{P}(S_1 \cap S_2) = \mathbb{P}(\text{Bribe})\, \mathbb{P}(S_1 \cap S_2 \mid \text{Bribe}) + \mathbb{P}(\text{No bribe})\, \mathbb{P}(S_1 \cap S_2 \mid \text{No bribe}) \tag{1}$$

We know a contestant is guaranteed to stay on the show, given that they are bribing the judges, hence:

$$\mathbb{P}(S_1 \cap S_2 \mid \text{Bribe}) = 1$$

On the other hand, if they have not been bribing judges, then the probability they stay on the show is $1/3$, independent of what happens on earlier episodes. By conditional independence, we have:

$$Pr(S_1 \cap S_2 \mid \text{No bribe}) = Pr(S_1 \mid \text{No bribe}) Pr(S_2 \mid \text{No bribe}) = \frac{1}{3} \cdot \frac{1}{3}$$

Plugging our results above into equation (1) gives us:

$$\mathbb{P}(S_1 \cap S_2) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{3} \cdot \frac{1}{3} = \boxed{\frac{1}{3}}$$

(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?

Solution:

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

$$\mathbb{P}(\overline{S_2} \mid S_1) = \frac{\mathbb{P}(S_1 \cap \overline{S_2})}{\mathbb{P}(S_1)} = \frac{\mathbb{P}(S_1 \cap \overline{S_2} \mid \text{Bribe})\mathbb{P}(\text{Bribe}) + \mathbb{P}(S_1 \cap \overline{S_2} \mid \text{No bribe})\mathbb{P}(\text{No bribe})}{\mathbb{P}(S_1)}$$

We have already computed $P(S_1)$ in part (a). We compute the numerator term by term. Given that a contestant is bribing the judges, they are guaranteed to stay on the show. As such:

$$\mathbb{P}(S_1 \cap \overline{S_2} \mid \text{Bribe}) = \mathbb{P}(S_1 \mid \text{Bribe}) \cdot \mathbb{P}(\overline{S_2} \mid \text{Bribe}) = 1 \cdot 0 = 0$$

On the other hand, if they have not been bribing judges, the probability they leave the show is $2/3$ (by complementing). We can then write:

$$\mathbb{P}(S_1 \cap \overline{S_2} \mid \text{No bribe}) = \mathbb{P}(S_1 \cap \mid \text{No bribe}) \cdot \mathbb{P}(\overline{S_2} \mid \text{No bribe}) = \frac{1}{3} \cdot \frac{2}{3}$$

We can now evaluate our initial expression:

$$\mathbb{P}(\overline{S_2} \mid S_1) = \frac{0 \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{3}{4}}{\frac{1}{2}} = \frac{1/6}{1/2} = \boxed{\frac{1}{3}}$$

(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?

Solution:

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

$$\mathbb{P}(B \mid S_1) = \frac{\mathbb{P}(S_1 \mid B)\mathbb{P}(B)}{\mathbb{P}(S_1)} = \frac{1 \cdot \frac{1}{4}}{\frac{1}{2}} = \boxed{\frac{1}{2}}$$

7. No More Coins Please

There are three coins, C_1, C_2, and C_3. The probability of "heads" is 1 for C_1, 0 for C_2, and p for C_3. A coin is picked among these three uniformly at random, and then flipped a certain number of times.

(a) What is the probability that the first n flips are tails?

Solution:

We have

$$1/3 \cdot 0 + 1/3 \cdot 1 + 1/3 \cdot (1-p)^n = \frac{1}{3} + \frac{1}{3}(1-p)^n \, .$$

(b) Given that the first n flips were tails, what is the probability that C_1 was flipped / C_2 was flipped / C_3 was flipped?

Solution:

We use Bayes Rule, and obtain

$$C_1 \mid n \text{ tails} = \frac{1/3 \cdot 0}{1/3 + 1/3(1-p)^n} = 0$$

$$C_2 \mid n \text{ tails} = \frac{1/3 \cdot 1}{1/3 + 1/3(1-p)^n} = \frac{1}{1+(1-p)^n}$$

$$C_3 \mid n \text{ tails} = \frac{1/3 \cdot (1-p)^n}{1/3 + 1/3(1-p)^n} = \frac{(1-p)^n}{1+(1-p)^n}$$

8. Parallel Systems

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) What is the probability the system is functioning?

Solution:

Let C_i be the event component i is working, and F be the event that the system is functioning.

For the system to function, it is sufficient for any component to be working. This means that the only case in which the system does not function is when none of the components work. We can then use complementing to compute $\mathbb{P}(F)$, knowing that $\mathbb{P}(C_i) = p$. We get:

$$\mathbb{P}(F) = 1 - \mathbb{P}(F^C) = 1 - \mathbb{P}(\bigcap_{i=1}^{n} C_i^C) = 1 - \prod_{i=1}^{n} \mathbb{P}(C_i^C) = 1 - \prod_{i=1}^{n} (1 - \mathbb{P}(C_i)) = 1 - \prod_{i=1}^{n} (1-p) = \boxed{1-(1-p)^n}$$

Note that $\mathbb{P}(\bigcap_{i=1}^{n} C_i^C) = \prod_{i=1}^{n} \mathbb{P}(C_i^C)$ due to independence of C_i (components working <u>independently</u> of each other). Note also that $\prod_{i=1}^{n} a = a^n$ for any constant a.

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

Solution:

We know that for the system to function only one component needs to be working, so for all i, we have $\mathbb{P}(F \mid C_i) = 1$. Using Bayes Theorem, we get:

$$\mathbb{P}(C_1|F) = \frac{\mathbb{P}(F|C_1)\mathbb{P}(C_1)}{\mathbb{P}(F)} = \frac{1 \cdot p}{1-(1-p)^n} = \boxed{\frac{p}{1-(1-p)^n}}$$

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

Solution:

$$\mathbb{P}(C_1|C_2, F) = \mathbb{P}(C_1|C_2) = \mathbb{P}(C_1) = p$$

where the first equality holds because knowing C_2 and F is just as good as knowing C_2 (since if C_2 happens, F does too), and the second equality holds because the components working are independent of each other.

More formally, we can use the definition of conditional probability along with a careful application of the chain rule to get the same result. We start with the following expression:

$$\mathbb{P}(C_1 \mid C_2, F) = \frac{\mathbb{P}(C_1, C_2, F)}{\mathbb{P}(C_2, F)} = \frac{\mathbb{P}(F \mid C_1, C_2) \cdot P(C_1 \mid C_2)\mathbb{P}(C_2)}{\mathbb{P}(F \mid C_2) \cdot \mathbb{P}(C_2)}$$

We note that the system is guaranteed to work if any one component is working, so $\mathbb{P}(F \mid C_1, C_2) = \mathbb{P}(F|C_2) = 1$. We also note that components work independently of each other, hence $\mathbb{P}(C_1|C_2) = \mathbb{P}(C_1)$. With that in mind, we can rewrite our expression so that:

$$\mathbb{P}(C_1 \mid C_2, F) = \frac{1 \cdot \mathbb{P}(C_1) \cdot \mathbb{P}(C_2)}{1 \cdot \mathbb{P}(C_2)} = \mathbb{P}(C_1) = \boxed{p}$$

9. Marbles in Pockets

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?

Solution:

Let W, B denote the event that we choose a white marble or a blue marble respectively, with subscripts L, R indicating from which pocket we are picking – left and right, respectively.

We know that we will pick from the left pocket first, and right pocket second. We can then use the Law of Total Probability conditioning on the color of the transferred marble so that:

$$\mathbb{P}(B_R) = \mathbb{P}(W_L) \cdot \mathbb{P}(B_R|W_L) + \mathbb{P}(B_L) \cdot \mathbb{P}(B_R|B_L) = \frac{3}{8} \cdot \frac{4}{9} + \frac{5}{8} \cdot \frac{5}{9} = \boxed{\frac{37}{72}}$$

10. Allergy Season

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.

number of colds	no drug or ineffective	drug effective
0	0.2	0.4
1	0.2	0.3
2	0.2	0.2
3	0.2	0.1
4	0.2	0.0

(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?

Solution:

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,

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

(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?

Solution:

Let the reduced sample space for part (b) be C_1 from part (a), so that $\mathbb{P}_{C_1}(E) = \mathbb{P}_{\Omega}(E|C_1)$. Let D_i be the event that he gets i colds the second winter. By Bayes' Theorem,

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

(c) Why is the answer to (b) the same as the answer to (a)?

Solution:

The probability of two colds whether or not the drug was effective is the same. Hence knowing that Sneezy got two colds does not change the probability of the drug's effectiveness.

11. Infinite Lottery

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 $\mathbb{P}(A_k) = (\frac{1}{2})^k$. Once we generate a number k, that is the maximum we can win. That is, after generating a value k, we can win any number in $[k] = \{1, ..., k\}$ dollars. Suppose the probability that we win $\$j$ for $j \in [k]$ is "uniform", that is, each has probability $\frac{1}{k}$. Let B be the event we win exactly $\$1$. Given that we win exactly one dollar, what is the probability that the number generated was also 1? You may use the fact that $\sum_{j=1}^{\infty} \frac{1}{j \cdot a^j} = \ln(\frac{a}{a-1})$for $a > 1$.

Solution:

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(B|A_1)\mathbb{P}(A_1)}{\sum_{j=1}^{\infty} \mathbb{P}(B|A_j)\mathbb{P}(A_j)} = \frac{\frac{1}{1}\frac{1}{2^1}}{\sum_{j=1}^{\infty} \frac{1}{j}\frac{1}{2^j}} = \frac{1}{2\ln 2} \approx \boxed{0.7213}$$