1 Let's Talk Probability

(a) When is $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ true? What is the general expression for $\mathbb{P}[A \cup B]$ that is always true?

(b) When is $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$ true? What is the general expression for $\mathbb{P}[A \cap B]$ that is always true?

(c) If A and B are disjoint, does that imply they're independent?

Solution:

(a) In general, we know $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}(A \cap B)$. This is the Inclusion-Exclusion Principle. Therefore if A and B are disjoint, such that $\mathbb{P}[A \cap B] = 0$, then $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ holds.

(b) $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$ holds if and only if A and B are independent (by definition). The general rule that always holds is $\mathbb{P}[A \cap B] = \mathbb{P}[B|A]\mathbb{P}[A]$ (if $\mathbb{P}[A] \neq 0$) and $\mathbb{P}[A \cap B] = \mathbb{P}[A|B]\mathbb{P}[B]$ (if $\mathbb{P}[B] \neq 0$).

(c) No, if two events are disjoint, we cannot conclude they are independent, unless at least one of them has zero probability.

As an example, let A, B be disjoint non-zero probability events. In other words, $A \cap B = \emptyset$, $\mathbb{P}[A] > 0$, and $\mathbb{P}[B] > 0$. We have $\mathbb{P}[A \cap B] = \mathbb{P}[\emptyset] = 0$, but $\mathbb{P}[A]\mathbb{P}[B] > 0$. Thus, A and B are dependent.

Since disjoint events have $\mathbb{P}[A \cap B] = 0$, we can see that the only time when disjoint A and B are independent is when either $\mathbb{P}[A] = 0$ or $\mathbb{P}[B] = 0$.

2 Independent Complements

Let Ω be a sample space, and let $A, B \subseteq \Omega$ be two independent events.

(a) Prove or disprove: $\overline{A}$ and $\overline{B}$ must be independent.

(b) Prove or disprove: A and $\overline{B}$ must be independent.

(c) Prove or disprove: A and $\overline{A}$ must be independent.

(d) Prove or disprove: It is possible that $A = B$.

Solution:

(a) True. $\overline{A}$ and $\overline{B}$ must be independent:

$$\begin{aligned}
\mathbb{P}[\overline{A} \cap \overline{B}] &= \mathbb{P}[\overline{A \cup B}] && \text{(by De Morgan's law)} \\
&= 1 - \mathbb{P}[A \cup B] && \text{(since } \mathbb{P}[\overline{E}] = 1 - \mathbb{P}[E] \text{ for all } E\text{)} \\
&= 1 - (\mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]) && \text{(union of overlapping events)} \\
&= 1 - \mathbb{P}[A] - \mathbb{P}[B] + \mathbb{P}[A]\mathbb{P}[B] && \text{(using our assumption that } A \text{ and } B \text{ are independent)} \\
&= (1 - \mathbb{P}[A])(1 - \mathbb{P}[B]) \\
&= \mathbb{P}[\overline{A}]\mathbb{P}[\overline{B}] && \text{(since } \mathbb{P}[\overline{E}] = 1 - \mathbb{P}[E] \text{ for all } E\text{)}
\end{aligned}$$

(b) True. A and $\overline{B}$ must be independent:

$$\begin{aligned}
\mathbb{P}[A \cap \overline{B}] &= \mathbb{P}[A - (A \cap B)] \\
&= \mathbb{P}[A] - \mathbb{P}[A \cap B] \\
&= \mathbb{P}[A] - \mathbb{P}[A]\mathbb{P}[B] \\
&= \mathbb{P}[A](1 - \mathbb{P}[B]) \\
&= \mathbb{P}[A]\mathbb{P}[\overline{B}]
\end{aligned}$$

(c) False in general. If $0 < \mathbb{P}[A] < 1$, then $\mathbb{P}[A \cap \overline{A}] = \mathbb{P}[\varnothing] = 0$ but $\mathbb{P}[A]\mathbb{P}[\overline{A}] > 0$, so $\mathbb{P}[A \cap \overline{A}] \neq \mathbb{P}[A]\mathbb{P}[\overline{A}]$; therefore A and $\overline{A}$ are not independent in this case.

(d) True. To give one example, if $\mathbb{P}[A] = \mathbb{P}[B] = 0$, then $\mathbb{P}[A \cap B] = 0 = 0 \times 0 = \mathbb{P}[A]\mathbb{P}[B]$, so A and B are independent in this case. (Another example: If $A = B$ and $\mathbb{P}[A] = 1$, then A and B are independent.)

3 Conditional Practice

(a) Suppose you have 3 bags. Two of them contain a \$10 bill and a \$5 bill, and the third contains two \$5 bills. Suppose you pick one of these bags uniformly at random, you draw a bill from the bag without looking uniformly at random. Suppose it turns out to be a \$5 bill. If a you draw the remaining bill from the bag, what is the probability that it, too, is a \$5 bill? Show your calculations.

(b) Now suppose that you have a large number of bags, and that each of them contain either a gold or a silver coin (every bag contains exactly one coin). Moreover, these bags are either colored red, blue, or yellow (every bag is exactly one of these colors). Half of the bags are red and a third of the bags are blue. Moreover, two thirds of the red bags and one fourth of the blue bags contain gold coins. Lastly, a randomly chosen bag has a $\frac{1}{2}$ probability of containing a silver coin. Suppose that you pick a bag at random and find that it contains a silver coin. What is the probability that the bag you picked was yellow?

Solution:

(a) Let A denote the event that the first bill we draw is a 5 dollar bill, and let B denote the event that the second bill we draw is also a 5 dollar bill. We wish to compute $P[B \mid A]$. By Bayes' Rule, we wish to compute $\frac{P[A\cap B]}{P[A]}$.

We start by computing $P[A\cap B]$. This happens with probability $\frac{1}{3}$, as this can only happen if we pick the one bag out of the three that contains the 2 five dollar bills.

Next, $P[A]$ is $\frac{4}{6}$; there are 6 dollar bills that are equally likely to be chosen first, and 4 of them are a \$5 bill.

We conclude that $P[B \mid A] = \frac{\frac{1}{3}}{\frac{4}{6}} = \boxed{\frac{1}{2}}$.

(b) Let R,B,Y be the events that the bag we pick is red, blue, and purple yellow. Moreover, let G and S denote the events that the bag we pick contains a gold or a silver coin, respectively. Since every bag is exactly one of the stated colors, and since every bag contains either a gold or a silver coin, we have that $\mathbb{P}[R]+\mathbb{P}[B]+\mathbb{P}[Y] = \mathbb{P}[G]+\mathbb{P}[S] = 1$. Now, we are given in the problem that $\mathbb{P}[R] = \frac{1}{2}$, $\mathbb{P}[B] = \frac{1}{3}$, and since two-thirds of the red bags and one-fourth of the blue bags contain gold coins, we have that $\mathbb{P}[G \mid R] = \frac{2}{3}$ and $\mathbb{P}[G \mid B] = \frac{1}{4}$. Lastly, we are given $\mathbb{P}[S] = \frac{1}{2}$.

Now, we can see from $\mathbb{P}[R]+\mathbb{P}[B]+\mathbb{P}[Y] = \mathbb{P}[G]+\mathbb{P}[S] = 1$ that $\mathbb{P}[Y] = \frac{1}{6}$ and $\mathbb{P}[G] = \frac{1}{2}$. Using the Law of Total Probability and Bayes' Theorem, we can expand

$$\begin{aligned}
\mathbb{P}[G] &= \mathbb{P}[G\cap R]+\mathbb{P}[G\cap B]+\mathbb{P}[G\cap Y] \\
&= \mathbb{P}[R]\cdot\mathbb{P}[G \mid R]+\mathbb{P}[B]\cdot\mathbb{P}[G \mid B]+\mathbb{P}[Y]\cdot\mathbb{P}[G \mid Y] \\
&= \frac{1}{2}\cdot\frac{2}{3}+\frac{1}{3}\cdot\frac{1}{4}+\frac{1}{6}\cdot\mathbb{P}[G \mid Y].
\end{aligned}$$

Now, since $\mathbb{P}[G] = \frac{1}{2}$, we can solve the above to get that $\mathbb{P}[G \mid Y] = \frac{1}{2}$. Thus, $\mathbb{P}[S \mid Y] = 1-\mathbb{P}[G \mid Y] = \frac{1}{2}$. Finally, we can use Bayes' Theorem again to get that

$$\mathbb{P}[Y \mid S] = \frac{\mathbb{P}[S \mid Y]\cdot\mathbb{P}[Y]}{\mathbb{P}[S]} = \frac{\frac{1}{2}\cdot\frac{1}{6}}{\frac{1}{2}} = \boxed{\frac{1}{6}}.$$

4 Monty Hall's Revenge

Due to a quirk of the television studio's recruitment process, Monty Hall has ended up drawing all the contestants for his game show from among the ranks of former CS70 students. Unfortunately for Monty, the former students' amazing probability skills have made his cars-and-goats gimmick unprofitable for the studio. Monty decides to up the stakes by asking his contestants to generalise to three new situations with a variable number of doors, goats, and cars:

(a) There are n doors for some $n > 2$. One has a car behind it, and the remaining $n-1$ have goats. As in the ordinary Monty Hall problem, Monty will reveal one door with a goat behind it after you make your first selection. How would switching affect the odds that you select the car?

(Hint: Think about the size of the sample space for the experiment where you *always* switch. How many of those outcomes are favorable?)

(b) Again there are $n > 2$ doors, one with a car and $n-1$ with goats, but this time Monty will reveal $n-2$ doors with goats behind them instead of just one. How does switching affect the odds of winning in this modified scenario?

(c) Finally, imagine there are $k < n-1$ cars and $n-k$ goats behind the $n > 2$ doors. After you make your first pick, Monty will reveal $j < n-k$ doors with goats. What values of j,k maximize the relative improvement in your odds of winning if you choose to switch? (i.e. what j,k maximizes the ratio between your odds of winning when you switch, and your odds of winning when you do not switch?)

Solution:

Throughout the solution, we will refer to W as the event that the contestant wins, and $\mathbb{P}_S[W]$ and $\mathbb{P}_N[W]$ as the probabilities of this event happening if the contestant is (S)witching or (N)ot switching, respectively.

(a) $\mathbb{P}_N[W] = 1/n$ since only one out of n initial choices gets us the car. Under the switching strategy two things can happen: Either the first choice hits the car, and so switching (to any of the remaining $n-2$ doors) will inevitably get us the goat, or our first choice picks a goat, leaving one of the remaining $n-2$ doors with the car. This sequence of choices—first choosing from one of n doors, then switching to one of $n-2$ remaining doors—gives us a sample space of size $n(n-2)$. If we divide the number of favorable outcomes by the total number of outcomes, we get

$$\begin{aligned}\mathbb{P}_S[W] &= \left(\underbrace{(n-1)}_{\text{first choice = goat}} \cdot \underbrace{1}_{\text{second choice = car}}\right) \Big/ \underbrace{n(n-2)}_{\text{total \# of choices}} \\ &= \frac{n-1}{n(n-2)} = \frac{1}{n} \cdot \frac{n-1}{n-2}\end{aligned}$$

which is larger than $\mathbb{P}_N[W] = 1/n$ (ever so slightly so the larger n becomes, which demonstrates the intuitive fact that Monty's help gets decreasingly helpful the more doors there are), so switching doors is the better strategy.

(b) $\mathbb{P}_N[W] = 1/n$ remains unchanged. The same approach as in part (a) yields the same numerator as before. For the denominator, we need to figure out the size of the sample space for the experiment where we first pick a door at random, then switch. Again, there are n ways of making the first choice. Once Monty reveals $n-2$ other doors, though, there is only one remaining option for us to switch to. Thus the denominator is much smaller:

$$\mathbb{P}_S[W] = \left(\underbrace{(n-1)}_{\text{first choice = goat}} \cdot \underbrace{1}_{\text{second choice = car}} \right) / \underbrace{n \cdot 1}_{\text{total \# of choices}}$$
$$= \frac{n-1}{n} = 1 - \frac{1}{n}$$

so switching is again the better strategy.

(c) Now $\mathbb{P}_N[W] = k/n$ since k doors hide a car. Reasoning about sample spaces in the same way we did in part (b) gives us a way to compute the denominator of $\mathbb{P}_S[W]$. However, now the numerator (number of favorable outcomes in the case where we switch) changes too:

$$\mathbb{P}_S[W] = \left(\underbrace{k}_{\text{first choice = car}} \cdot \underbrace{k-1}_{\text{second choice = car}} + \underbrace{(n-k)}_{\text{first choice = goat}} \cdot \underbrace{k}_{\text{second choice = car}} \right) / \underbrace{n(n-j-1)}_{\text{total \# of choices}}$$
$$= \frac{k(n-1)}{n(n-j-1)} = \frac{k}{n} \cdot \frac{n-1}{n-j-1}.$$

From here we see that $\mathbb{P}_S[W]/\mathbb{P}_N[W] = \frac{n-1}{n-j-1}$, which is maximal if $j = n-k-1$. In other words, if Monty reveals all but one goat (which he does in the original show where $n = 3, k = 1$ and $j = 1 = n-k-1$), then the contestant can increase their chances of winning by a factor of $\frac{n-1}{k}$ (which is a factor of 2 in the original show). In particular, the largest relative advantage of switching is achieved when $k = 1$.

5 Random Polynomials

Consider the following scenarios, where we apply probability to polynomials. We generate a new real polynomial Q, by picking 6 numbers in the set $\{0,1,2,3,4,5,6\}$ independently and uniformly at random (with replacement), for a polynomial of the form

$$a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0.$$

The first number we pick is a_5, the second number is a_4, etc.

(a) What is probability we have a polynomial of degree less than 4?

(b) What is the probability that the polynomial has degree at least 4?

(c) Now, consider only degree 5 polynomials that were randomly generated using the scheme described above. What is the probability that the sum of its coefficients is equal to 6?

Solution:

(a) For a polynomial to have degree less than 4, we need $a_5 = 0$ and $a_4 = 0$. This happens with probability $\frac{1}{7} \cdot \frac{1}{7} = \frac{1}{49}$.

(b) This is 1 minus the previous part. So, the probability is $1 - \frac{1}{49} = \frac{48}{49}$

(c) We note that there are $6(7^5)$ different coefficients; there are only 6 options for the first coefficient since it cannot be 0. We now compute all the possible ways to achieve a sum of 6. This is simply stars and bars, where the first bin must have at least one ball (a_5 cannot be zero). We distribute the sum 5 among 6 bins: $\binom{10}{5}$. Thus, our probability is

$$\frac{\binom{10}{5}}{6(7^5)}.$$

6 (Un)conditional (In)equalities

Let us consider a sample space $\Omega = \{\omega_1, \ldots, \omega_N\}$ of size $N > 2$ and two probability functions $\mathbb{P}_1$ and $\mathbb{P}_2$ on it. That is, we have two probability spaces: $(\Omega, \mathbb{P}_1)$ and $(\Omega, \mathbb{P}_2)$.

(a) Suppose that for every subset $A \subseteq \Omega$ of size $|A| = 2$ and for every outcome $\omega \in \Omega$, it is true that $\mathbb{P}_1[\omega \mid A] = \mathbb{P}_2[\omega \mid A]$. Is it necessarily true that $\mathbb{P}_1[\omega] = \mathbb{P}_2[\omega]$ for all $\omega \in \Omega$? That is, if $\mathbb{P}_1$ and $\mathbb{P}_2$ are equal conditional on events of size 2, are they equal unconditionally? (*Hint*: Remember that probabilities must add up to 1.)

(b) Suppose that for every subset $A \subseteq \Omega$ of size $|A| = k$, where k is some fixed element in $\{2, \ldots, N\}$, and for every outcome $\omega \in \Omega$, it is true that $\mathbb{P}_1[\omega \mid A] = \mathbb{P}_2[\omega \mid A]$. Is it necessarily true that $\mathbb{P}_1[\omega] = \mathbb{P}_2[\omega]$ for all $\omega \in \Omega$?

For the following two parts, assume that $\Omega = \left\{(a_1, \ldots, a_k) \mid \sum_{j=1}^{k} a_j = n\right\}$ is the set of configurations of n balls into k labeled bins, and let $\mathbb{P}_1$ be the probabilities assigned to these configurations by throwing the balls independently one after another and they will land into any of the k bins uniformly at random, and let $\mathbb{P}_2$ be the probabilities assigned to these configurations by uniformly sampling one of these configurations.

(c) Let A be the event that all n balls are in exactly one bin.

  (i) What are $\mathbb{P}_1[\omega \mid A]$ and $\mathbb{P}_2[\omega \mid A]$ for any $\omega \in A$?

  (ii) Repeat part (i) for $\omega \in \Omega \setminus A$.

  (iii) Is it true that $\mathbb{P}_1[\omega] = \mathbb{P}_2[\omega]$ for all $\omega \in \Omega$?

(d) For the special case of $n = 9$ and $k = 3$, provide two outcomes B and C, so that $\mathbb{P}_1[B] < \mathbb{P}_2[B]$ and $\mathbb{P}_1[C] > \mathbb{P}_2[C]$. Provide justification.

Solution:

(a) Yes, this is indeed true. To see why, let's take the subset $A = \{\omega_i, \omega_j\}$ for some $i, j \in \{1, \ldots, N\}$ and compute: For any $k \in \{1,2\}$, we have $\mathbb{P}_k[\omega_i \mid A] = \frac{\mathbb{P}_k[\omega_i]}{\mathbb{P}_k[A]}$. Since this expression (by assumption) is the same for $k = 1$ and $k = 2$, we conclude that $\frac{\mathbb{P}_1[\omega_i]}{\mathbb{P}_2[\omega_i]} = \frac{\mathbb{P}_1[A]}{\mathbb{P}_2[A]}$. Repeating the reasoning for ω_j, we similarly find that $\frac{\mathbb{P}_1[\omega_j]}{\mathbb{P}_2[\omega_j]} = \frac{\mathbb{P}_1[A]}{\mathbb{P}_2[A]}$, and whence $\frac{\mathbb{P}_1[\omega_i]}{\mathbb{P}_1[\omega_j]} = \frac{\mathbb{P}_2[\omega_i]}{\mathbb{P}_2[\omega_j]}$. Since this is true for any $i, j \in \{1, \ldots, N\}$, we can sum over i to get

$$\frac{1}{\mathbb{P}_1[\omega_j]} = \sum_{i=1}^{N} \frac{\mathbb{P}_1[\omega_i]}{\mathbb{P}_1[\omega_j]} = \sum_{i=1}^{N} \frac{\mathbb{P}_2[\omega_i]}{\mathbb{P}_2[\omega_j]} = \frac{1}{\mathbb{P}_2[\omega_j]},$$

which shows that $\mathbb{P}_1[\omega_j] = \mathbb{P}_2[\omega_j]$ for all $j \in \{1, \ldots, N\}$.

(b) Yes, it indeed would. There are two ways of verifying this. The first way to convince ourselves that part (b) is true, is to observe that none of the arguments used in part (a) really relied on A having size 2, and so the very same reasoning carries through for A of size k.

The second (more rigorous) one is to observe that if $A' \subset A$ and $\omega \in A'$, then $\mathbb{P}_1[\omega \mid A'] = \mathbb{P}_1[\omega \mid A' \cap A] = \frac{\mathbb{P}_1[\omega \mid A]}{\mathbb{P}_1[A' \mid A]} = \frac{\mathbb{P}_2[\omega \mid A]}{\mathbb{P}_2[A' \mid A]} = \mathbb{P}_2[\omega \mid A']$, where the second equality follows from the product rule (Theorem 13.1): $\mathbb{P}_1[A] \cdot \mathbb{P}_1[A' \mid A] \cdot \mathbb{P}_1[\omega \mid A \cap A'] = \mathbb{P}_1[\{\omega\} \cap A \cap A'] = \mathbb{P}_1[\omega] = \mathbb{P}_1[A]\mathbb{P}_1[\omega \mid A]$. That is, if $\mathbb{P}_1$ and $\mathbb{P}_2$ coincide conditional on some event A, they also coincide conditional on any smaller event A'. In particular, if they coincide on all events of size k, they also coincide on all events of size 2, which we have already dealt with in part (a).

(c) There are exactly k outcomes in A (namely, $(n, 0, \ldots, 0), (0, n, 0, \ldots), \ldots, (0, \ldots, 0, n)$; i.e. each bin could be the full one), and all of them are equally likely under either $\mathbb{P}_1$ or $\mathbb{P}_2$. That is, if $\omega \in A$, then $\mathbb{P}_1[\omega] = \left(\frac{1}{k}\right)^n$, and $\mathbb{P}_2[\omega] = \left[\binom{n+k-1}{k-1}\right]^{-1}$. Consequently, for $\omega \in A$,

$$\mathbb{P}_1[\omega \mid A] = \frac{k^{-n}}{k \cdot k^{-n}} = \frac{1}{k} \qquad\qquad \mathbb{P}_2[\omega \mid A] = \frac{\binom{n+k-1}{k-1}^{-1}}{k \cdot \binom{n+k-1}{k-1}^{-1}} = \frac{1}{k}.$$

If $\omega \notin A$, then $\mathbb{P}_1[\omega \mid A] = \mathbb{P}_2[\omega \mid A] = 0$, and so $\mathbb{P}_1[\omega \mid A]$ and $\mathbb{P}_2[\omega \mid A]$ coincide for all $\omega \in \Omega$. This, however, does *not* imply that $\mathbb{P}_1$ and $\mathbb{P}_2$ are the same! Indeed, when computing the probability of $\omega \in A$ above, we saw that $\mathbb{P}_1[\omega] \neq \mathbb{P}_2[\omega]$ (remember that the assumption of part (b) was that the conditional probabilities coincide for *all* events of size k, here we have only shown equality conditional on *one* such event).

(d) Intuitively, throwing balls independently one after another makes it much less likely that all balls stack up in one bin as opposed to spreading out more evenly. This suggests taking, e.g., $A = \{\text{all balls land in bin } 1\}$, whose probability we already computed in part (c). That is, to show that $\mathbb{P}_1[A] < \mathbb{P}_2[A]$, we need to show that $k^{-n} < \binom{n+k-1}{k-1}^{-1}$. Plugging in $k = 3$ and $n = 9$, we have

$$k^{-n} = 3^{-9} = 3^{-2} \cdot 3^{-3} \cdot 3^{-4} = \frac{1}{9 \cdot 27 \cdot 3^4} < \frac{1}{5 \cdot 11} = \frac{2}{11 \cdot 10} = \binom{11}{2}^{-1} = \binom{n+k-1}{k-1}^{-1},$$

as desired.

Conversely, the same reasoning suggests that evenly distributed balls are much more likely under $\mathbb{P}_1$ than under $\mathbb{P}_2$. And indeed, letting $B = \{\text{each bin has exactly three balls}\}$, we have

$$\mathbb{P}_1[B] = \binom{9}{3}\binom{6}{3}\binom{3}{3}\left(\frac{1}{3}\right)^9 = \frac{9!}{(3!)^3 \cdot 3^9} = \frac{7!}{3^{10}} > \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3}{3^6} \cdot \frac{2}{11 \cdot 10} = \frac{7}{3} \cdot \frac{6 \cdot 5}{3^3} \cdot \frac{4}{3}\mathbb{P}_2[B],$$

and since the last factor is bigger than 1, we have $\mathbb{P}_1[B] > \mathbb{P}_2[B]$ as promised (of course, we could have also just plugged all these powers and factorials into a calculator to get the same result).