

Contents

1	Probability and Counting	1
1.1	Counting	1
1.2	Story proofs	5
1.3	Naive Definition of Probability	6
1.4	Inclusion-Exclusion	9
1.5	Independence	10
1.6	Thinking Conditionally	11
2	Conditional Probability	13
2.1	Conditioning on evidence	13
2.2	Independence and conditional independence	14

1 Probability and Counting

1.1 Counting

- How many ways are there to permute the letters in the word MISSISSIPPI?

- The answer is obtained by finding the total number of permutation and dividing by the number of permutation of the individual letters. The total number of permutations is $11!$ because there are 11 characters in that word. For example the number of permutations of the four characters S is $4!$, as such the number of permutations is

$$\frac{11!}{1!4!4!2!} = 34650.$$

- (a) How many 7-digit phone numbers are possible, assuming that the first digit can't be a 0 or a 1?

- Assuming all other digits are possible there are $10 - 2 = 8$ possible digits for the first position and 10 for the remaining $7 - 1 = 6$ positions as such the number of phone numbers is

$$8 * 10^6 = 8000000.$$

- (b) Re-solve (a), except now assume also that the phone number is not allowed to start with 911 (since this is reserved for emergency use, and it would not be desirable for the system to wait to see whether more digits were going to be dialed after someone has dialed 911).

- The total count of numbers starting with 911 is 10^4 as such removing those yields

$$8000000 - 10^4 = 7990000.$$

3. Fred is planning to go out to dinner each night of a certain week, Monday through Friday, with each dinner being at one of his ten favorite restaurants.

- How many possibilities are there for Fred's schedule of dinners for that Monday through Friday, if Fred is not willing to eat at the same restaurant more than once?
 - There are 5 days and 10 restaurants which are picked without repetition. As such he has 10 choices on the first day, 9 choices on the second day and so on, as such in total

$$\frac{10!}{5!} = 30240.$$

- How many possibilities are there for Fred's schedule of dinners for that Monday through Friday, if Fred is willing to eat at the same restaurant more than once, but is not willing to eat at the same place twice in a row (or more)?
 - There are 10 possibilities on the first day. On the second day he can choose any restaurant which is different from the first one. Same situation on the third day and so on. As such in total

$$10 * 9^4 = 65610.$$

4. A round-robin tournament is being held with n tennis players; this means that every player will play against every other player exactly once.

- (a) How many possible outcomes are there for the tournament (the outcome lists out who won and who lost for each game)?

- Every player will have exactly $n - 1$ opponents, summing up over all players gives $n(n - 1)$, but of course we doublecounted everything because every opponent is also a player. As such the total number of matches is $\frac{n(n-1)}{2}$. Because every match is either a win or a loss the total number of outcomes is

$$2^{\frac{n(n-1)}{2}}.$$

- (b) How many games are played in total?

- As mentioned in the first part, it's $\frac{n(n-1)}{2}$.

5. A knock-out tournament is being held with $2n$ tennis players. This means that for each round, the winners move on to the next round and the losers are eliminated, until only one person remains. For example, if initially there are $24 = 16$ players, then there are 8 games in the first round, then the 8 winners move on to round 2, then the 4 winners move on to round 3, then the 2 winners move on to round 4, the winner of which is declared the winner of the tournament. (There are various systems for determining who plays whom within a round, but these do not matter for this problem.)

- (a) How many rounds are there?
- We half the amount of players remaining in every step so there are $\log_2(2^n) = n$ rounds.
- (b) Count how many games in total are played, by adding up the numbers of games played in each round.
- If there are 2^k players in a round then 2^{k-1} games have been played, as such the total number of games is

$$\sum_{k=1}^n 2^{k-1} = 2^n - 1.$$

- (c) Count how many games in total are played, this time by directly thinking about it without doing almost any calculation.
- We filtered everybody but one person, every game filters exactly one person (the loser of the 1v1), as such there have to be 2^{n-1} matches.
6. There are 20 people at a chess club on a certain day. They each find opponents and start playing. How many possibilities are there for how they are matched up, assuming that in each game it does matter who has the white pieces (in a chess game, one player has the white pieces and the other player has the black pieces)?
- Imagine we have 20 cells and the $2k$ th person plays with the $2k + 1$ st person. Then the problem becomes how many unique assignments do we have. The important thing is not in what cell the person is, but with whom they are in a double-cell with and if they are in the odd or even position (corresponding to white / black pieces). There are $20!$ ways to assign the people, there are $10!$ arrangement of the cells, as such the answer is

$$\frac{20!}{10!} = 670442572800.$$

7. Two chess players, A and B, are going to play 7 games. Each game has three possible outcomes: a win for A (which is a loss for B), a draw (tie), and a loss for A (which is a win for B). A win is worth 1 point, a draw is worth 0.5 points, and a loss is worth 0 points.
- (a) How many possible outcomes for the individual games are there, such that overall player A ends up with 3 wins, 2 draws, and 2 losses?
- There are $7!$ total permutations, and we have to remove the $3!$ for the wins, and $2!$ for the draws and losses individually, as such

$$\frac{7!}{3!2!2!} = 210.$$

- (b) How many possible outcomes for the individual games are there, such that A ends up with 4 points and B ends up with 3 points?

- Note that a game always increases the total points by 1, just varies up the distribution. Because $4 + 3 = 7$ are the total points there have been 7 matches played. The number of ties has to be even because the players have even points, meaning we only have to consider the cases where $\# \text{ties} \in \{0, 2, 4, 6\}$. If $2k, 0 \leq k \leq 3$ is the number of ties then $(3 - k, 2k, 4 - k)$ is a valid outcome, and those are the only ones. In total this means the number of outcomes is

$$|\{0, 2, 4, 6\}| = 4.$$

- (c) Now assume that they are playing a best-of-7 match, where the match will end when either player has 4 points or when 7 games have been played, whichever is first. For example, if after 6 games the score is 4 to 2 in favor of A, then A wins the match and they don't play a 7th game. How many possible outcomes for the individual games are there, such that the match lasts for 7 games and A wins by a score of 4 to 3?

- The only possible arrangement for this outcome is if the 6th round was a 3 to 3 and A wins the last round. As such we are interested in the number of ways to get to 3 to 3. We can draw out a binary tree for this and simply count the number of leafs. Note that once one of the players reaches 3 the tree can be cut there because there is only one possible solution. With this method we get the result of 17.

8. (a) How many ways are there to split a dozen people into 3 teams, where one team has 2 people, and the other two teams have 5 people each?
- (b) How many ways are there to split a dozen people into 3 teams, where each team has 4 people?
9. (a) How many paths are there from the point $(0, 0)$ to the point $(110, 111)$ in the plane such that each step either consists of going one unit up or one unit to the right?

- Denote by $C((a, b), (x, y))$ where $(a, b) \leq (x, y)$ (elementwise) the number of such paths starting from (a, b) and ending on (x, y) . Note that $C((a, b), (x, y)) = C((a+1, b), (x, y)) + C((a, b+1), (x, y))$ and $C((a, b), (x, y)) = C((b, a), (y, x))$. Furthermore for $(a, b) \leq (x', y') \leq (x, y)$ we have

$$C((a, b), (x, y)) = C((a, b), (x', y')) + C((x', y'), (a, b)).$$

Note that to get from (a, b) to (x, y) we'll need exactly $y - b$ steps up and $x - a$ steps to the right. The only difference between different paths is the order of the steps. The total number of sets

then is $n := (y - b) + (x - a)$. The number of total permutations of this is $n!$, but all of the right and up steps are the same, so we need to factor out those internal permutations, which are $(x - a)!$ and $(y - b)!$ respectively. As such

$$C((a, b), (x, y)) = \frac{((y - b) + (x - a))!}{(y - b)!(x - a)!}$$

As such

$$C((0, 0), (110, 111)) = \frac{221!}{110!111!}.$$

Note that this is exactly

$$\binom{n}{n - (y - b)}.$$

- (b) How many paths are there from $(0, 0)$ to $(210, 211)$, where each step consists of going one unit up or one unit to the right, and the path has to go through $(110, 111)$?

•

$$\begin{aligned} & C((0, 0), (110, 111)) \cdot C((110, 111), (210, 211)) \\ &= C((0, 0), (110, 111)) \cdot C((0, 0), (100, 100)) \\ &= \frac{221!}{110!111!} \cdot \frac{200!}{100!100!} \end{aligned}$$

1.2 Story proofs

1. Give a story proof that

$$\sum_{k=0}^n \binom{n}{k} = 2^n.$$

2. Show that for all positive integers n and k with $n \geq k$,

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

doing this in two ways: (a) algebraically and (b) with a story giving an interpretation for why both sides count the same thing.

Hint for the story proof: Imagine an organization consisting of $n+1$ people, with one of them pre-designated as the president of the organization.

3. Give a story proof that

$$\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}$$

for all positive integers n .

- When we want to consider the number of ways to pick n elements of $\{1, 2, \dots, n, n+1, \dots, 2n\}$ then for every permutation there exists a $0 \leq k \leq n$ such that it picks k elements from $\{1, 2, \dots, n\}$ and $n-k$ elements from $\{n+1, \dots, 2n\}$. For fixed k there are

$$\binom{n}{k} \binom{n}{n-k}$$

such possibilities. This proves that $\binom{2n}{n} = \sum_{k=0}^n \binom{n}{k} \binom{n}{n-k}$. Note that picking $n-k$ elements from k is equivalent to picking the other k elements, as such

$$\binom{n}{k} = \binom{n}{n-k}.$$

In total we have shown that

$$\binom{2n}{n} = \sum_{k=0}^n \binom{n}{k} \binom{n}{n-k} = \sum_{k=0}^n \binom{n}{k}^2.$$

4. Give a story proof that

$$\frac{(2n)!}{2^n \cdot n!} = (2n-1)(2n-3) \cdots 3 \cdot 1.$$

- Suppose we are given $2n$ people and we want to count the number of pairs.

To see that the right hand side describes this note that the first person has $2n-1$ people to choose, the second person has $2n-3$ people to choose and so on, until we are left with two people and then there is only one choice.

On the other hand we could also just line up the people, there are $(2n)!$ ways to do this and pair the people such that $2k$ and $2k+1$ are paired up. For every pair we can swap the arrangement of people so we overcount by 2 in each pair, for a total of 2^n . On the other hand the order of the pairs in the line also doesn't matter. There are $n!$ ways to arrange those pairs, as such the total number of pairings is

$$\frac{(2n)!}{2^n \cdot n!}.$$

1.3 Naive Definition of Probability

1. A certain family has 6 children, consisting of 3 boys and 3 girls. Assuming that all birth orders are equally likely, what is the probability that the 3 eldest children are the 3 girls?
 - There are $6!$ ways to arrange the three daughters (or equivalently the three sons) in the birth order. We don't care about the order of the

boys or girls within their gender so the total number of arrangements is

$$\frac{6!}{3!3!} = 20.$$

Of those permutations every single arrangement is equally likely and there is only one we are interested in, so the probability is

$$1/20 = 0.05\%.$$

2. A college has 10 (non-overlapping) time slots for its courses, and blithely assigns courses to time slots randomly and independently. A student randomly chooses 3 of the courses to enroll in (for the PTP, to avoid getting fined). What is the probability that there is a conflict in the student's schedule?

- The first course doesn't cause overlap, the second course has a 10% to cause an overlap and the second course has a 20% chance to cause an overlap. So the chance that there is overlap is $1 - (1 - 0.1)(1 - 0.2) = 0.28 = 28\%$.

3. A city with 6 districts has 6 robberies in a particular week. Assume the robberies are located randomly, with all possibilities for which robbery occurred where equally likely. What is the probability that some district had more than 1 robbery?

- Let the first robber choose a district randomly. The second robber has a $5/6$ chance to hit a new district, the third one has a chance of $4/6$ and so on. In total the chance of no overlap is

$$\frac{5}{6} \frac{4}{6} \frac{3}{6} \frac{2}{6} \frac{1}{6} = \frac{5!}{6^5} = \frac{6!}{6^6}.$$

As such the probability of having an overlap is the complement of this

$$1 - \frac{6!}{6^6} \cong 0.9846.$$

4. Elk dwell in a certain forest. There are N elk, of which a simple random sample of size n are captured and tagged ("simple random sample" means that all $\binom{N}{n}$ sets of n elk are equally likely). The captured elk are returned to the population, and then a new sample is drawn, this time with size m . This is an important method that is widely-used in ecology, known as capture-recapture.

What is the probability that exactly k of the m elk in the new sample were previously tagged? (Assume that an elk that was captured before doesn't become more or less likely to be captured again.)

- Let X be the set of elks that have been captured before, $|X| = n$ and Y be the set of elks captured now, $|Y| = m$. The set of all elks is denoted by Z and $|Z| = N$. Note that $Z = X \cup X^c$. We are interested in the probability that $|X \cap Y| = k$, or equivalently that $|Y \setminus X| = m - k$.

Let's pick a random elk $e \in Z$ then the probability that $e \in X$ is equal to $\frac{n}{N}$. For there to be exactly k elks from X we have to pick k elks from X and $m - k$ elks from X^c , which has probability

$$\left(\frac{n}{N}\right)^k \left(\frac{N-n}{N}\right)^{m-k}$$

There are a total of $\binom{N}{k}$ choices like this so the probability is

$$\binom{N}{k} \left(\frac{n}{N}\right)^k \left(\frac{N-n}{N}\right)^{m-k}.$$

Note that this is simply the hypergeometric distribution.

5. A jar contains r red balls and g green balls, where r and g are fixed positive integers. A ball is drawn from the jar randomly (with all possibilities equally likely), and then a second ball is drawn randomly.
 - (a) Explain intuitively why the probability of the second ball being green is the same as the probability of the first ball being green.
 - If we are returning the ball after picking it then the state after and before picking it up is the same, so the probabilities don't change.
 - (b) Define notation for the sample space of the problem, and use this to compute the probabilities from (a) and show that they are the same.
 - $P(\text{green}) = \frac{g}{g+r}$, after picking the ball up and putting it back the probabilities don't change.
 - (c) Suppose that there are 16 balls in total, and that the probability that the two balls are the same color is the same as the probability that they are different colors. What are r and g (list all possibilities)?
6. A norepeatword is a sequence of at least one (and possibly all) of the usual 26 letters a,b,c,...,z, with repetitions not allowed. For example, "course" is a norepeatword, but "statistics" is not. Order matters, e.g., "course" is not the same as "source". A norepeatword is chosen randomly, with all norepeatwords equally likely. Show that the probability that it uses all 26 letters is very close to $1/e$.
 - Let's fix $1 \leq k \leq 26$. To compute the number of nonrepeat words of length k , which we denote by $n_k \in \mathbb{N}$. We first have to select k out of 26 letters, there are $\binom{26}{k}$ possibilities for this. Then Every single

permutation of those letters is a new valid norepeatword. There are $k!$ permutations like that, as such

$$n_k = \binom{26}{k} k! = \frac{26!}{(26-k)!k!} k! = \frac{26!}{(26-k)!}.$$

We are interested in the number

$$p = \frac{n_{26}}{\sum_{k=1}^{25} n_k} = \frac{26!}{\sum_{k=1}^{25} \frac{26!}{(26-k)!}} = \frac{1}{\sum_{k=1}^{25} \frac{1}{(26-k)!}}.$$

Recall that

$$e = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!}$$

so that

$$e \sim \sum_{k=0}^{25} \frac{1}{k!} = \sum_{k=1}^{25} \frac{1}{(26-k)!}$$

Note that we are missing $n_0 = 1$.

1.4 Inclusion-Exclusion

1. For a group of 7 people, find the probability that all 4 seasons (winter, spring, summer, fall) occur at least once each among their birthdays, assuming that all seasons are equally likely.
 - Let X_i denote the number of times the i th season appears. We are interested in $P(X_1 > 0, X_2 > 0, X_3 > 0, X_4 > 0)$ note that this is the same as

$$1 - P(X_1 = 0 \cup X_2 = 0 \cup X_3 = 0 \cup X_4 = 0).$$

Using the principle of inclusion exclusion and noting that because $\sum_{i=1}^4 X_i = 7$ we can't have $X_i = 0$ for all $1 \leq i \leq 4$, we can calculate

$$\begin{aligned} & \sum_{i=1}^4 P(X_i = 0) - \sum_{i \neq j} P(X_i = 0, X_j = 0) \\ & + \sum_{1 \leq i, j, k \leq 4, \text{ p.w. disjoint}} P(X_i = 0, X_j = 0, X_k = 0), \end{aligned}$$

noting that the order of the elements doesn't matter we have

$$\binom{4}{1} P(X_1 = 0) - \binom{4}{2} P(X_1 = 0, X_2 = 0) + \binom{4}{3} P(X_1 = 0, X_2 = 0, X_3 = 0).$$

Note that $P(X_1 = 0) = \left(\frac{3}{4}\right)^4$, $P(X_1 = 0, X_2 = 0) = \left(\frac{1}{2}\right)^4$ and $P(X_1 = 0, X_2 = 0, X_3 = 0) = \left(\frac{1}{4}\right)^4$. In total this means the desired probability is

$$1 - \left(4 \left(\frac{3}{4}\right)^4 - 6 \left(\frac{1}{2}\right)^4 + 4 \left(\frac{1}{4}\right)^4 \right) \cong 0.513.$$

2. Alice attends a small college in which each class meets only once a week. She is deciding between 30 non-overlapping classes. There are 6 classes to choose from for each day of the week, Monday through Friday. Trusting in the benevolence of randomness, Alice decides to register for 7 randomly selected classes out of the 30, with all choices equally likely. What is the probability that she will have classes every day, Monday through Friday? (This problem can be done either directly using the naive definition of probability, or using inclusion-exclusion.)

- Let X_i denote the number of courses in the i th day chosen. We know $\sum_{i=1}^5 X_i = 30$. Similarly to before

$$P(X_i = 0) = \frac{24}{30} \frac{23}{29} \frac{22}{28} \frac{21}{27} = \frac{26!}{30!} \frac{24}{20} = \frac{\binom{24}{7}}{\binom{30}{7}}$$

3. A card player is dealt a 13-card hand from a well-shuffled, standard deck of cards. What is the probability that the hand is void in at least one suit (“void in a suit” means having no cards of that suit)?

- Denote by X_i the number of cards of the i th suit then

$$\begin{aligned} P(X_i = 0) &= \frac{39}{52} \frac{38}{51} \cdots \frac{26}{39} \\ &= \frac{38!}{52!} \frac{39!}{25!} \\ &= \frac{\frac{39!}{25!}}{\frac{52!}{38!}} \\ &= \frac{\frac{39!}{25!} 14!}{\frac{52!}{38!} 14!} \\ &= \frac{\binom{39}{14}}{\binom{52}{14}}. \end{aligned}$$

Now we can apply inclusion exclusion to calculate the desired probability

$$P(X_1 = 0 \cup X_2 = 0 \cup X_3 = 0 \cup X_4 = 0) = \binom{4}{1} P(X_1 = 0) + \binom{4}{2} P(X_1 = 0, X_2 = 0) + \binom{4}{3} P(X_1 = 0, X_2 = 0, X_3 = 0) + \binom{4}{4} P(X_1 = 0, X_2 = 0, X_3 = 0, X_4 = 0)$$

$P(X_1 = 0)$ we have already calculated above,

$$P(X_1 = 0, X_2 = 0) = ,$$

and

$$P(X_1 = 0, X_2 = 0, X_3 = 0)$$

can be calculated similarly.

1.5 Independence

1. Is it always true that if A and B are independent events, then A^c and B^c are independent events? Show that it is, or give a counterexample.
 - Suppose A, B are independent, then

$$P(A, B) = P(A)P(B).$$

Note that $P(A^c)P(B^c) = (1 - P(A))(1 - P(B)) = 1 - P(A) - P(B) + P(A)P(B) = 1 - P(A) - P(B) + P(A, B)$. By DeMorgan $P(A^c, B^c) = P((A \cup B)^c) = 1 - P(A \cup B) = 1 - P(A) - P(B) + P(A, B)$.

2. Give an example of 3 events A, B, C which are pairwise independent but not independent. Hint: find an example where whether C occurs is completely determined if we know whether A occurred and whether B occurred, but completely undetermined if we know only one of these things.
3. Give an example of 3 events A, B, C which are not independent, yet satisfy $P(A \cap B \cap C) = P(A)P(B)P(C)$. Hint: consider simple and extreme cases.
 - First note that $P(X) \neq 0$ for $X \in \{A, B, C\}$, otherwise the remaining variables would have to be independent. Let $A = C$ then

$$P(A, B) = P(A, B, A) = P(A)^2 P(B)$$

as such $P(A, B) \neq P(A)P(B)$. Of course A and $C = A$ are not independent. And also $P(C, B) = P(A, B)$ is not independent.

Alternatively we could have $P(A) = 0$ then $P(A, B, C) = P(A)P(B)P(C) = 0$. For any B, C independent such that $P(B) \neq 0$ we then have the desired form.

1.6 Thinking Conditionally

1. A bag contains one marble which is either green or blue, with equal probabilities. A green marble is put in the bag (so there are 2 marbles now), and then a random marble is taken out. The marble taken out is green. What is the probability that the remaining marble is also green?
2. Let G be the event that a certain individual is guilty of a certain robbery. In gathering evidence, it is learned that an event E_1 occurred, and a little later it is also learned that another event E_2 also occurred.
 - (a) Is it possible that individually, these pieces of evidence increase the chance of guilt (so $P(G|E_1) > P(G)$ and $P(G|E_2) > P(G)$), but together they decrease the chance of guilt (so $P(G|E_1, E_2) < P(G)$)?

- (b) Show that the probability of guilt given the evidence is the same regardless of whether we update our probabilities all at once, or in two steps (after getting the first piece of evidence, and again after getting the second piece of evidence). That is, we can either update all at once (computing $P(G|E_1, E_2)$ in one step), or we can first update based on E_1 , so that our new probability function is $P_{\text{new}}(A) = P(A|E_1)$, and then update based on E_2 by computing $P_{\text{new}}(G|E_2)$.

- This problem asks if $P((A|B)|C) = P(A|B, C)$.

$$\begin{aligned}
 P_{\text{new}}(G|E_2) &= \frac{P_{\text{new}}(G, E_2)}{P_{\text{new}}(E_2)} \\
 &= \frac{P(G, E_2|E_1)}{P(E_2|E_1)} \\
 &= \frac{\frac{P(G, E_1, E_2)}{P(E_1)}}{\frac{P(E_1, E_2)}{P(E_1)}} \\
 &= \frac{P(G, E_1, E_2)}{P(E_1, E_2)} = P(G|E_1, E_2).
 \end{aligned}$$

3. A crime is committed by one of two suspects, A and B. Initially, there is equal evidence against both of them. In further investigation at the crime scene, it is found that the guilty party had a blood type found in 10% of the population. Suspect A does match this blood type, whereas the blood type of Suspect B is unknown.

- (a) Given this new information, what is the probability that A is the guilty party?

- Denote by A that A is guilty and by β that A has the given blood type, then

$$\begin{aligned}
 P(A|\beta) &= P(\beta|A) \frac{P(A)}{P(\beta)} \\
 &= \frac{P(\beta|A)P(A)}{P(\beta|A)P(A) + P(\beta|B)P(B)}.
 \end{aligned}$$

Note that $P(\beta|A) = 1$ because the guilty party needs to have that blood type, also $P(A) = 0.5$. Now note that $P(\beta|B) = 0.1$ because if B is the guilty party then the probability of A having that blood type is just the average. $P(B) = 1 - P(A) = 0.5$. In total this shows that

$$P(A|\beta) = \frac{10}{11}.$$

- (b) Given this new information, what is the probability that B's blood type matches that found at the crime scene?

- Note that $P(\beta'|A^c) = 1$ and $P(\beta'|A) = 0.1$. As such we are interested in

$$P(\beta'|\beta) = P(A|\beta)P(\beta'|A) + P(A^c|\beta)P(\beta'|A^c) = \frac{10}{11} \frac{1}{10} + \frac{1}{11} \cdot 1 = \frac{2}{11}.$$

4. You are going to play 2 games of chess with an opponent whom you have never played against before (for the sake of this problem). Your opponent is equally likely to be a beginner, intermediate, or a master. Depending on which, your chances of winning an individual game are 90%, 50%, or 30%, respectively.

(a) What is your probability of winning the first game?

- Let $X = 1, 2, 3$ be the skill of my opponent, then $P(X = i) = \frac{1}{3}, i = 1, 2, 3$ and let W denote whether I win or not $P(W|X_1) = 0.9, P(W|X_2) = 0.5$ and $P(W|X_3) = 0.3$. The total win chance is then simply

$$P(W) = \sum_{i=1}^3 P(W|X_i)P(X_i) = \frac{1}{3} (0.9 + 0.5 + 0.3) = \frac{1.7}{3} = \frac{17}{30}.$$

(b) Congratulations: you won the first game! Given this information, what is the probability that you will also win the second game (assume that, given the skill level of your opponent, the outcomes of the games are independent)?

- Recall that $P(X_i|W) = P(W|X_i) \frac{P(X_i)}{P(W)}$ holds, and note that $\frac{P(X_i)}{P(W)} = \frac{\frac{1}{3}}{\frac{1.7}{3}} = \frac{10}{17}$. With that we can then calculate

$$P(X_1|W) = \frac{9}{17}, \quad P(X_2|W) = \frac{5}{17}, \quad P(X_3|W) = \frac{3}{17}.$$

With that the probability of winning again is then

$$P(W_2) = \frac{9}{17} \cdot 0.9 + \frac{5}{17} \cdot 0.5 + \frac{3}{17} \cdot 0.3 = \frac{23}{34}.$$

- (c) Explain the distinction between assuming that the outcomes of the games are independent and assuming that they are conditionally independent given the opponent's skill level. Which of these assumptions seems more reasonable, and why?

2 Conditional Probability

2.1 Conditioning on evidence

1. A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails,

the phrase “free money” is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention “free money”. What is the probability that it is spam?

- Let S denote if the email is spam, and F is the phrase “free money” is contained in the email. We are interested in finding $P(S|F)$, to determine this note

$$\begin{aligned} P(S|F) &= \frac{P(S, F)}{P(F)} \\ &= \frac{1}{P(F)} P(F|S)P(S) \\ &= \frac{P(S)}{P(F)} P(F|S). \end{aligned}$$

We know that $P(S) = 0.8$, $P(F|S) = 0.1$ and $P(F|S^c) = 0.01$. This allows us to calculate

$$\begin{aligned} P(F) &= P(F, S) + P(F, S^c) \\ &= P(F|S)P(S) + P(F|S^c)P(S^c) \\ &= 0.1 \cdot 0.8 + 0.01 \cdot (1 - 0.8) \\ &= 0.1 \cdot 0.8 + 0.01 \cdot 0.2 = 0.082 = 8.2\%, \end{aligned}$$

so in total

$$P(S|F) = \frac{P(S)}{P(F)} P(F|S) = \frac{0.8}{0.082} 0.1 \cong 0.9756 = 97.56\%.$$

2. According to the CDC (Centers for Disease Control and Prevention), men who smoke are 23 times more likely to develop lung cancer than men who don't smoke. Also according to the CDC, 21.6% of men in the U.S. smoke. What is the probability that a man in the U.S. is a smoker, given that he develops lung cancer?

- We write $p := P(L|S)$ and calculate

$$\begin{aligned} P(S|L) &= P(L|S) \frac{P(S)}{P(L)} \\ &= P(L|S) \frac{P(S)}{P(L|S)P(S) + P(L|S^c)P(S^c)} \\ &= P(L|S) \frac{0.216}{0.216p + \frac{p}{23}(1 - 0.216)} \\ &= p \frac{0.216}{0.216p + \frac{p}{23}(0.784)} \\ &\cong 0.8637 = 86.37\%. \end{aligned}$$

2.2 Independence and conditional independence

1. A family has 3 children, creatively named A , B , and C .
 - (a) Discuss intuitively (but clearly) whether the event “ A is older than B ” is independent of the event “ A is older than C ”.
 - Note that there are $3! = 6$ different possible birth orders, if A is older than B then the possible birth orders are $\{BAC, BCA, CBA\}$. Out of those A being older than C occurs in $\frac{2}{3}$ of the cases, which is higher than it would be without that information (which of course has to be $1/2$ because the children are interchangeable).
 - (b) Find the probability that A is older than B , given that A is older than C .
 - $P(A > B | A > C) = \frac{2}{3}$ as discussed before.
2. Is it possible that an event is independent of itself? If so, when?
 - Suppose A is independent to itself, i.e., $P(A, A) = P(A)P(A)$. Of course $P(A, A) = P(A)$ so this condition becomes

$$P(A)^2 = P(A)$$

which is the case if and only if $P(A) \in \{0, 1\}$.