CS 70 Discrete Mathematics and Probability Theory
Summer 2016 Dinh, Psomas, and Ye HW 4

Due Monday July 18 at 1:59PM

1. **Sample Space and Events (1/1/1/1/1/1/2/2 points)**
Consider the sample space Ω of all outcomes from flipping a coin 3 times.

   (a) List all the outcomes in Ω. How many are there?
   (b) Let A be the event that the first flip is a heads. List all the outcomes in A. How many are there?
   (c) Let B be the event that the third flip is a heads. List all the outcomes in B. How many are there?
   (d) Let C be the event that the first and third flip are heads. List all outcomes in C. How many are there?
   (e) Let D be the event that the first or the third flip is heads. List all outcomes in D. How many are there?
   (f) Are the events A and B disjoint? Express C in terms of A and B. Express D in terms of A and B.
   (g) Suppose now the coin is flipped $n \geq 3$ times instead of 3 flips. Compute $|\Omega|, |A|, |B|, |C|, |D|$.
   (h) Your gambling buddy found a website online where he could buy trick coins that are heads or tails on both sides. He puts three coins into a bag: one coin that is heads on both sides, one coin that is tails on both sides, and one that is heads on one side and tails on the other side. You shake the bag, draw out a coin at random, put it on the table without looking at it, then look at the side that is showing. Suppose you notice that the side that is showing is heads. What is the probability that the other side is heads? Show your work. (Hint: the answer is NOT 1/2).

Answer:

   (a) Each flip results in either heads (H) or tails (T). So in total the total number of outcomes is 8, which we represent by length 3 strings of H's and T's. We have
   $$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$
   (b) These are the strings that start with H. We have $A = \{HHH, HHT, HTH, HTT\}$. There are 4 such outcomes.
   (c) These are the strings that end with an H. We have $B = \{HHH, HTH, THH, TTH\}$. There are 4 such outcomes.
   (d) These are the strings that start and end with an H. We have $C = \{HHH, HTH\}$. There are 2 such outcomes.
   (e) We have $D = \{HHH, HHT, HTH, HTT, THH, TTH\}$. There are 6 such outcomes.
   (f) No, A and B are not disjoint. For example HHH belongs to both of them.
   The event C is the intersection of A and B, because in C we require exactly both A (the first coin being heads) and B (the third coin being heads) to happen. So $C = A \cap B$.
   The event D is the union of A and B, because in D we require at least one of A (the first coin being heads) or B (the second coin being heads) to happen. So $D = A \cup B$.

(g) First, obviously $|\Omega| = 2^n$.

Note that for each outcome in the three-coin case, there are now 2^{n-3} outcomes, each corresponding to a possible configuration of the 4th flip and beyond. Since A, B, C, and D do not care about the outcomes of the 4th flip and beyond, this means that the size of each set is simply multiplied by 2^{n-3}. Therefore we have $|A| = 4 \times 2^{n-3} = 2^{n-1}$, $|B| = 4 \times 2^{n-3} = 2^{n-1}$, $|C| = 2 \times 2^{n-3} = 2^{n-2}$, and $|D| = 6 \times 2^{n-3} = 3 \times 2^{n-2}$.

(h) There are 6 possible outcomes which are all equally likely. We have 3 choices for the coin that we draw (which we represent by HH, HT and TT). Then for each coin we have two choices, we either see the first side or the second side (which we represent by 1 and 2). So the outcomes are $\Omega = \{(HH,1),(HH,2),(HT,1),(HT,2),(TT,1),(TT,2)\}$. Now given that we saw a heads, we can get rid of 3 of the outcomes, and the possible remaining outcomes are $\{(HH,1),(HH,2),(HT,1)\}$ which are all equally likely. In this space, the event that the coin has two heads is $\{(HH,1),(HH,2)\}$ which consists of two equally likely outcomes. So the probability is $2/3$.

2. **To Be Fair (5 points)**

Suppose you have a biased coin with $P(heads) \neq 0.5$. How could you use this coin to simulate a fair coin? (Hint: Think about pairs of tosses.)

Answer: Let H be the event that the coin flip comes up heads and T be the event that the coin flip comes up tails. We do not know the probability of getting heads since we've been told the coin is biased, but let's denote it by the variable p (i.e., $p = P(H)$).

We conduct a mini experiment in which we flip the same coin twice. Let HT be the event that we get heads and then tails in two consecutive flips. Similarly, let TH be the event that we get tails and then heads in two consecutive flips. Also, we'll let HH be the event that we get two heads in a row, and TT be the event that we get two tails in a row. Our sample space for the experiment is then $\Omega = \{HH,HT,TH,TT\}$.

Since we know $P(H) = p$, we can write the probabilities of the events in our sample space as follows:

$$P(HH) = P(\text{First toss H})P(\text{Second toss H}|\text{First toss H})$$

Because the nature of coins tells us that the second toss doesn't remember the first toss so

$$P(HH) = P(H)P(H) = p^2.$$

Similarly,

$$\begin{aligned}
P(HT) &= P(\text{First toss H})P(\text{Second toss T}|\text{First toss H}) = P(H)(1-P(H)) = p(1-p),\\
P(TH) &= P(\text{First toss T})P(\text{Second toss H}|\text{First toss T}) = (1-P(H))P(H) = p(1-p),\\
P(TT) &= P(\text{First toss T})P(\text{Second toss T}|\text{First toss T}) = (1-P(H))(1-P(H)) = (1-p)^2.
\end{aligned}$$

We notice that the probability $P(HT)$ and the probability $P(TH)$ are equal, i.e., they are both $p(1-p)$. By symmetry, these two probabilities must be the same since the coin doesn't know what it came up before. Since these two are the same, we can simply condition on the fact that something was returned to get that the resulting simulated coin toss is fair.

Therefore, we can simulate a fair coin using the following process. We toss the coin twice. If the outcome turns out to be heads both times or tails both times, we throw away the result and repeat the whole process again. Otherwise, if the outcome is HT, we return "heads", and if the outcome is TH, we return "tails".

How do we know this procedure will return a result at all? How do we know that this can't go on forever? We are not yet in a position to prove this rigorously because we haven't built the tools yet. But intuitively, this would require getting an infinite sequence of HH's or TT's. Why would the coin always agree with itself? This seems like it defies the nature of coin tossing: that the coin doesn't know what it did before.

3. **Picking CS Classes (2/3/5 points)**

The EECS (Elegant Etiquette Charm School) department has d different classes being offered in Summer 2016. These include classes such as dressing etiquette, dining etiquette, and social etiquette, etc. Let's assume that all the classes are equally popular and each class has essentially unlimited seating! Suppose that c students are enrolled this semester and the registration system, EleBEARS (Elegant Bears), requires each student to choose a class s/he plans to attend.

(a) What is the probability that a given student chooses the first class, dressing etiquette?

Answer: There are d different choices for each student. The probability of choosing "Dressing etiquette" class is $\frac{1}{d}$.

(b) What is the probability that a given class is chosen by no student?

Answer: The probability a given student of choosing a class is $\frac{1}{d}$ and hence the probability of a given student not choosing the class is $\frac{d-1}{d}$. Each student chooses the class independently and there are c students. Hence the probability of no student choosing the class is $\left(\frac{d-1}{d}\right)^c$.

(c) If there are $d = 20$ classes, what should c be in order for the probability to be at least one half that (at least) two students enroll in the same class?

Answer: From the notes, we know that for d bins, the probability of no collision is less than $\frac{1}{2}$ for approximately $\lfloor 1.177\sqrt{d} \rfloor$ balls. The problem of at least two students enrolling in the same class can be viewed as at least one collision ocurring for $d = 20$ bins. Thus, in order for this probability to be at least $\frac{1}{2}$, the number of students needed is approximately $c = \lceil 1.177\sqrt{20} \rceil$, which gives us $c = 5$ students.
If we list out the exact values of probabilities of no collision for varying values c, we get the following results:
For $c = 1$, obviously $P(\text{no collision}) = 1$
For $c = 2$, the second student has $20 - 1 = 19$ choices, so $P(\text{no collision}) = \frac{19}{20}$
For $c = 3$, the second student has 19 choices and the third student has 18 choices, $P(\text{no collision}) = \frac{19 \times 18}{20} = \frac{342}{400}$
Similarly,
For $c = 4$ $P(\text{no collision}) = \frac{19 \times 18 \times 17}{20^3} = \frac{5814}{8000}$
For $c = 5$, $P(\text{no collision}) = \frac{19 \times 18 \times 17 \times 16}{20^4} = \frac{93024}{160000}$
For $c = 6$, $P(\text{no collision}) = \frac{19 \times 18 \times 17 \times 16 \times 15}{20^5} = \frac{139536}{320000}$, which is finally less than $\frac{1}{2}$.
Hence it takes at least $c = 6$ students for the probability of collision to be at least half, very close to our approximation of $c = 5$ students.

4. **Poisoned pancakes (2/3/5 points)**

You have been hired as an actuary by IHOP corporate headquarters, and have been handed a report from Corporate Intelligence that indicates that a covert team of ninjas hired by Denny's will sneak into some IHOP, and will have time to poison ten of the pancakes being prepared (they can't stay any longer to avoid being discovered by Pancake Security). Given that an IHOP kitchen has 100 pancakes being prepared, and there are twenty patrons, each ordering five pancakes (which are chosen uniformly at random from the pancakes in the kitchen), calculate the probabilities that a particular patron:

(a) will not receive any poisoned pancakes;

Answer: Our sample space Ω consists of all possible sets of 5 pancakes that the patron can receive, so $|\Omega| = \binom{100}{5}$ and each sample point has the same probability. Let A be the event that the patron receives no poisoned pancakes. Since there are 90 unpoisoned pancakes, $|A| = \#$ of ways to choose 5 pancakes out of 90 $= \binom{90}{5}$. So

$$Pr[\text{patron receives no poisoned pancakes}] = Pr[A] = \frac{|A|}{|\Omega|} = \frac{\binom{90}{5}}{\binom{100}{5}} = \frac{90 \times 89 \times 88 \times 87 \times 86}{100 \times 99 \times 98 \times 97 \times 96} \approx 0.584$$

Alternatively, we can argue as follows. For $1 \le i \le 5$, let A_i denote the event that the ith pancake received is not poisoned. Then we have

$$Pr\left[\bigcap_{i=1}^{5} A_i\right] = Pr[A_1] \times Pr[A_2|A_1] \times \cdots \times Pr[A_5|A_1 \cap A_2 \cap A_3 \cap A_4] = \frac{90}{100} \times \frac{89}{99} \times \frac{88}{98} \times \frac{87}{97} \times \frac{86}{96}$$

which is of course the same answer as above, as in the alternative argument above.

(b) will receive exactly one poisoned pancake;

Answer: We work with the same sample space as in (a). Let B denote the event that the patron receives precisely one poisoned pancake. There are $\binom{90}{4}$ ways to choose the 4 unpoisoned pancakes, and $\binom{10}{1}$ ways to choose the 1 poisoned pancake. So $|B| = \binom{90}{4}\binom{10}{1}$. Then

$$Pr[B] = \frac{|B|}{|\Omega|} = \frac{\binom{90}{4}\binom{10}{1}}{\binom{100}{5}} = \frac{90 \times 89 \times 88 \times 87 \times 10 \times 5}{100 \times 99 \times 98 \times 97 \times 96} \approx 0.339$$

(c) will receive at least one poisoned pancake.

Answer: Let C denote the event that the patron receives at least one poisoned pancake. Then $\bar{C}$ is the event that the patron receives no poisoned pancakes, which we calculated in part (a). Thus we get

$$Pr[C] = 1 - Pr[\bar{C}] = 1 - \frac{\binom{90}{5}}{\binom{100}{5}} \approx 0.416$$

5. **Cliques in random graphs (3/5/7 points)**

Consider a graph $G(V,E)$ on n vertices which is generated by the following random process: for each pair of vertices u and v, we flip a fair coin and place an (undirected) edge between u and v if and only if the coin comes up heads. So for example if $n = 2$, then with probability 1/2, $G(V,E)$ is the graph consisting of two vertices connected by an edge, and with probability 1/2 it is the graph consisting of two unconnected vertices.

(a) What is the size of the sample space?

(b) A k-clique in graph is a set of k vertices which are pairwise adjacent (every pair of vertices is connected by an edge). For example a 3-clique is a triangle. What is the probability that a particular set of k vertices forms a k-clique?

(c) Prove that the probability that the graph contains a k-clique for $k = 4\lceil \log n \rceil + 1$ is at most $1/n$.

Answer:

(a) There are two choices for each of the $\binom{n}{2}$ pairs of vertices, so the size of the sample space is $2^{\binom{n}{2}}$.

(b) For a set of k vertices to be a k-clique, all of the $\binom{k}{2}$ pairs of those vertices have to be connected by an edge. The probability of this event is $1/2^{\binom{k}{2}}$.

(c) Let A_S denote the event that S is a k-clique, where $S \subseteq V$ is of size k. Then, the event that the graph contains a k-clique can be described as the union of A_S's over all $S \subseteq V$ is of size k. Using the union bound,

$$Pr\left[\bigcup_{S\subseteq V,|S|=k} A_S\right] \le \sum_{S\subseteq V,|S|=k} Pr[A_S] = \sum_{S\subseteq V,|S|=k} \frac{1}{2^{\binom{k}{2}}}$$

Now since there $\binom{n}{k}$ ways of choosing a subset $S \subseteq V$ is of size k, the right-hand side of the above equality is

$$\frac{\binom{n}{k}}{2^{\binom{k}{2}}} = \frac{\binom{n}{k}}{2^{\frac{k(k-1)}{2}}} \le \frac{n^k}{\left(2^{\frac{(k-1)}{2}}\right)^k} \le \frac{n^k}{\left(2^{\frac{(4\log n+1-1)}{2}}\right)^k} = \frac{n^k}{(2^{2\log n})^k} = \frac{n^k}{n^{2k}} = \frac{1}{n^k} \le \frac{1}{n}$$

6. **Drunk man (1/2/2/2/3 points)**

Imagine that you have a drunk man moving along the horizontal axis (that stretches from $x = -\infty$ to $x = +\infty$). At time $t = 0$, his position on this axis is $x = 0$. At each time point $t = 1, t = 2$, etc., the man moves forward (that is, $x(t+1) = x(t) + 1$) with probability 0.5, backward (that is, $x(t+1) = x(t) - 1$) with probability 0.3, and stays exactly where he is (that is, $x(t+1) = x(t)$) with probability 0.2.

(a) What are all his possible positions at time t, $t \ge 0$?

Answer: Clearly, by time t, the man could have moved at most t positions to the right, and at most t positions to the left. Furthermore, within this range $[-t, t]$, the man could be occupying any integer position. Therefore, the possible values for the position $x(t)$ of the man at time t are exactly the integers in the closed range $[-t, t]$.

(b) Calculate the probability of each possible position at $t = 1$.

Answer: Clearly, at time $t = 1$, the man could be either in position -1, or in position 0, or in position 1. We know the man starts at position 0 at $t = 0$, and at time $t = 1$, he has taken at most 1 step; if this step were taken backward (w.p. 0.3), he would be in position -1, and if this step were forward (w.p. 0.5), he would be in position $+1$. And if he had chosen to remain wherever he was (w.p. 0.2), he would be in position 0. There is no other way he could have been in any of these positions. So, his possible positions are $[-1,\ 0,\ 1]$, with probabilities $[0.3,\ 0.2,\ 0.5]$ respectively.

(c) Calculate the probability of each possible position at $t = 2$.

Answer: From the discussion above, at time $t = 2$, the man can be in any one of the 5 positions $[-2,\ -1,\ 0,\ 1,\ 2]$. The probability associated with each of these positions can be calculated from the probabilities that we just computed above (for the man's position at time $t = 1$).

For example, what is the probability that the man is in position -2 at time 2? Clearly, this can happen under only one circumstance: the man should have been in position -1 at time 1, and moved backwards at time 2. Thus we have:

$$\begin{aligned}
P(x(2) = -2) &= P(x(1) = -1 \cap \text{man moves backward at } t = 2) \\
&= P(\text{man moves backward at } t = 2 \mid x(1) = -1) \times P(x(1) = -1) \\
&= 0.3 \times 0.3 \\
&= 0.09.
\end{aligned}$$

In general, using the law of total probability, we can write the probability that the man is in position i at time $t+1$ as

$$\begin{aligned}
&P(x(t+1) = i) \\
=&P(x(t+1) = i|x(t) = i-1) \times P(x(t) = i-1) + P(x(t+1) = i|x(t) = i+1) \times P(x(t) = i+1) \\
&+ P(x(t+1) = i|x(t) = i) \times P(x(t) = i) \\
=&0.5 \times P(x(t) = i-1) + 0.3 \times P(x(t) = i+1) + 0.2 \times P(x(t) = i).
\end{aligned} \tag{1}$$

Now, using (1),

$$\begin{aligned}
&P(x(2) = -1) \\
=&P(x(2) = -1|x(1) = -2) \times P(x(1) = -2) + P(x(2) = -1|x(1) = 0) \times P(x(1) = 0) \\
&+ P(x(2) = -1|x(1) = -1) \times P(x(1) = -1) \\
=&0 + 0.3 \times 0.2 + 0.2 \times 0.3 \\
=&0.12.
\end{aligned}$$

Doing this for the rest of the values, we get

$$
\begin{aligned}
&P(x(2)=0)\\
=&P(x(2)=0|x(1)=-1)\times P(x(1)=-1)+P(x(2)=0|x(1)=1)\times P(x(1)=1)\\
&+P(x(2)=0|x(1)=0)\times P(x(1)=0)\\
=&0.5\times 0.3+0.3\times 0.5+0.2\times 0.2\\
=&0.34.
\end{aligned}
$$

$$
\begin{aligned}
&P(x(2)=1)\\
=&P(x(2)=1|x(1)=0)\times P(x(1)=0)+P(x(2)=1|x(1)=2)\times P(x(1)=2)\\
&+P(x(2)=1|x(1)=1)\times P(x(1)=1)\\
=&0.5\times 0.2+0+0.2\times 0.5\\
=&0.2.
\end{aligned}
$$

$$
\begin{aligned}
&P(x(2)=2)\\
=&P(x(2)=2|x(1)=1)\times P(x(1)=1)+P(x(2)=2|x(1)=2)\times P(x(1)=2)\\
=&0.5\times 0.5+0\\
=&0.25.
\end{aligned}
$$

Notice that the 5 probabilities above add up to 1, as we would expect.

(d) Calculate the probability of each possible position at $t=3$.

Answer: From the discussion above, at time $t=3$, the man can be in any one of the 7 positions -3, -2, -1, 0, 1, 2, or 3. The probability associated with each of these positions can be calculated from the probabilities that we just computed above (for the man's position at time $t=2$).

The calculations are carried out in exactly the same way as in the previous part, by considering all possible ways in which the man can occupy position x at time 3, for each x satisfying $-3\le x\le 3$.

$$
\begin{aligned}
&P(x(3)=-3)\\
=&P(x(3)=-3|x(2)=-2)\times P(x(2)=-2)+P(x(3)=-3|x(2)=-3)\times P(x(2)=-3)\\
=&0.3\times 0.09+0\\
=&0.027.
\end{aligned}
$$

$$
\begin{aligned}
&P(x(3)=-2)\\
=&P(x(3)=-2|x(2)=-3)\times P(x(2)=-3)+P(x(3)=-2|x(2)=-1)\times P(x(2)=-1)\\
&+P(x(3)=-2|x(2)=-2)\times P(x(2)=-2)\\
=&0.3\times 0.12+0.2\times 0.09\\
=&0.054.
\end{aligned}
$$

Proceeding in a similar fashion, the probabilities for the man to be in positions -3, -2, -1, 0, 1, 2, and 3 are 0.027, 0.054, 0.171, 0.188, 0.285, 0.15, and 0.125 respectively for $t=3$. Again, as expected, these probabilities add up to 1.

(e) If you know the probability of each position at time t, how will you find the probabilities at time $t+1$?

Answer: The solution to the previous part of the problem suggests a nice algorithm for computing the probability of each position the man can take at time $t+1$, provided these probabilities are known for time t.

Let X_t be the list of all possible positions that the man can be in at time t. From the arguments above, we know that:

$$X_t = [-t,\ -(t-1),\ \ldots,\ -1,\ 0,\ 1,\ \ldots,\ (t-1),\ t]\ .$$

Let Pt denote a list of probabilities corresponding to the positions X_t. Our goal is to find a way to calculate P_{t+1} (the *next probabilities*) given P_t (the *current probabilities*).

```
#! /usr/bin/env python2

import sys

def next_pvec (pvec, pf, pb, pc):

    qvec = []
    for idx in range(len(pvec)+2):
        q = pvec[idx]*pb if 0 <= idx < len(pvec) else 0
        q += pvec[idx-1]*pc if 0 <= idx-1 < len(pvec) else 0
        q += pvec[idx-2]*pf if 0 <= idx-2 < len(pvec) else 0
        qvec.append(q)

    return qvec


if __name__ == '__main__':

    (pf, pc, pb) = (0.5, 0.2, 0.3)
    tf = int(sys.argv[1])

    (t, pvec) = (0, [1])
    while t < tf:
        pvec = next_pvec (pvec, pf, pb, pc)
        t += 1

    print(pvec)
```

The figure above shows Python code for calculating the above next probabilities. The function `next_pvec` takes as input the current list of probabilities `pvec` (at time t), and values for `pf`, `pb`, and `pc` (the forward, backward, and "stay put" probabilities), and it produces as output a list of the next probabilities (at time $t+1$).

First of all, observe that the length of the list X_{t+1}, and hence P_{t+1} is two more than the length of X_t (hence P_t). This is because, at time $t+1$ the man can be in two additional possible positions that he could not have been in at time t.

Also, for each possible position at time $t+1$, there are at most 3 possible positions the man could have been in at time t. Therefore, the rules described above for multiplying the relevant probabilities and adding up these products generalize quite readily.

Thus, given the positional probabilities at time t, the man's positional probabilities at time $t+1$ can be readily calculated. And the man's initial position is known to be $x(0) = 0$. Therefore, starting from this initial condition, the positional probabilities can be calculated at any desired future time. Indeed, the `main` part of the above program does exactly this; it accepts a future time `tf` from the user and prints out a list of probabilities corresponding to every possible position the man can be in at time `tf`.

Note: Those of you who are familiar with Linear Algebra will readily recognize that the "next probabilities" list is simply a linear combination of the "current probabilities" list, which corresponds to pre-multiplying the current probabilities list by a (tall and thin) rectangular matrix. Indeed, this idea can be used to considerably speed-up the probability calculations above.

Note: The following parts are optional.

The Drunk Man has regained some control over his movement, and no longer stays in the same spot; he only moves forwards or backwards. More formally, let the Drunk Man's initial position be $x(0) = 0$. Every second, he either moves forward one pace or backwards one pace, *i.e.*, his position at time $t+1$ will be one of $x(t+1) = x(t)+1$ or $x(t+1) = x(t)-1$.

We want to compute the number of paths in which the Drunk Man returns to 0 at time t and it is his first return, *i.e.*, $x(t) = 0$ and $x(s) \neq 0$ for all s where $0 < s < t$. Note, we **no longer** care about probabilities. We are just counting paths here.

(a) How many paths can the Drunk Man take if he returns to 0 at $t = 6$ and it is his first return?

Answer: We use an "F" to represent that the Drunk Man moves forward one pace and a "B" to represent that the Drunk Man moves backward one pace.

$\boxed{4}$ possible paths: FFFBBB, FFBFBB, BBBFFF, and BBFBFF. The last two paths can also be obtained by exchanging F's and B's in the first two paths.

(b) How many paths can the Drunk Man take if he returns to 0 at $t = 7$ and it is his first return?

Answer: $\boxed{0}$ possible path because it needs the same number of forward paces and backward paces.

(c) How many paths can the Drunk Man take if he returns to 0 at $t = 8$ and it is his first return?

Answer: $\boxed{10}$ possible paths: FFFFBBBB, FFFBFBBB, FFFBBFBB, FFBFFBBB, FFBFBFBB, and the other five by exchanging F's and B's.

(d) How many paths can the Drunk Man take if he returns to 0 at $t = 2n+1$ for $n \in \mathbb{N}$ and it is his first return?

Answer: $\boxed{0}$ possible path because it needs the same number of forward paces and backward paces.

(e) How many paths can the Drunk Man take if he returns to 0 at $t = 2n+2$ for $n \in \mathbb{N}$ and it is his first return? (Hint: read http://en.wikipedia.org/wiki/Catalan_number and use any result there if you need.)

Answer: From Wikipedia, "a Catalan number C_n is the number of monotonic paths along the edges of a grid with $n \times n$ square cells, which **do not pass above** the diagonal. A monotonic path is one which starts in the lower left corner, finishes in the upper right corner, and consists entirely of edges pointing rightwards or upwards." We can regard an F as an edge pointing rightwards, a B as an edge pointing upwards, and a possible path as a monotonic path of a grid with $(n+1) \times (n+1)$ square cells, which **does not touch** the diagonal. If the first pace is an F, then the $(2n+2)$-th path (the last pace) must be B; otherwise, the Drunk Man must have returned 0 before $t = 2n+2$. Therefore, in this case, we can focus on the second pace to the $(2n+1)$-th pace, and the number of possible paths is C_n because it is the number of monotonic paths, which do not pass above the diagonal of $n \times n$ square cells, *i.e.*, do not touch the diagonal of $(n+1) \times (n+1)$ square cells if the first pace is an F and the last pace is a B. Considering the other case that the first pace is a B and the last pace is an F, the total number of possible paths is $2C_n = \boxed{\frac{2}{n+1}\binom{2n}{n}.}$

7. **Independence (2/4/4 points)**

(a) **Independence (due to H.W. Lenstra)**

Suppose we pick a random card from a standard deck of 52 playing cards. Let A represent the event that the card is an queen, B the event that the card is a spade, and C the event that a red card (a heart or a diamond) is drawn.

i. Which two of A, B, and C are independent? Justify your answer carefully. (In other words: For each pair of events (AB, AC, and BC), state and prove whether they are independent or not.)

Answer: A and B are independent, since $Pr[A \cap B] = 1/52 = 1/13 \times 1/4 = Pr[A]Pr[B]$.
A and C are independent, since $Pr[A \cap C] = 2/52 = 1/13 \times 2/4 = Pr[A]Pr[C]$.
B and C are not independent, since $Pr[B \cap C] = 0 \neq 1/4 \times 2/4 = Pr[B]Pr[C]$.

ii. What if a joker is added to the deck? Justify your answer carefully.

Answer: Let A', B', C' denote the corresponding events when a joker is added to the deck. We assume that the joker has neither suit, rank, nor color, so that the joker is neither a queen, a spade, nor a red card. Then:
A' and B' are not independent, since $Pr[A' \cap B'] = 1/53 \neq 4/53 \times 13/53 = Pr[A']Pr[B']$.
A' and C' are not independent, since $Pr[A' \cap C'] = 2/53 \neq 4/53 \times 26/53 = Pr[A']Pr[C']$.
B' and C' are not independent, since $Pr[B' \cap C'] = 0 \neq 13/53 \times 26/53 = Pr[B']Pr[C']$.

(b) **Independence (due to H.W. Lenstra)**

Let Ω be a sample space, and let $A, B \subseteq \Omega$ be two independent events. Let $\overline{A} = \Omega - A$ and $\overline{B} = \Omega - B$ (sometimes written $\neg A$ and $\neg B$) denote the complementary events. For the purposes of this question, you may use the following definition of independence: Two events A, B are *independent* if $Pr[A \cap B] = Pr[A]Pr[B]$.

i. Prove or disprove: $\overline{A}$ and $\overline{B}$ are necessarily independent.

Answer: True. $\overline{A}$ and $\overline{B}$ must be independent:

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

ii. Prove or disprove: A and $\overline{B}$ are necessarily independent.

Answer: True. A and $\overline{B}$ must be independent:

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

iii. Prove or disprove: A and $\overline{A}$ are necessarily independent.

Answer: False in general. If $0 < Pr[A] < 1$, then $Pr[A \cap \overline{A}] = Pr[\emptyset] = 0$ but $Pr[A]Pr[\overline{A}] > 0$, so $Pr[A \cap \overline{A}] \neq Pr[A]Pr[\overline{A}]$; therefore A and $\overline{A}$ are not independent in this case.

iv. Prove or disprove: It is possible that $A = B$.

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

(c) **Bonferroni's inequalities**

i. For events A, B in the same probability space, prove that

$$Pr[A \cap B] \geq Pr[A] + Pr[B] - 1.$$

Answer: To show this we use the Inclusion-Exclusion theorem. We have that for all events A and B,

$$Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$$

Also, we know that the probability of any event is at most 1. Thus, $Pr[A \cup B] \leq 1$. Using this with the Inclusion-Exclusion theorem, we get

$$Pr[A \cap B] = Pr[A] + Pr[B] - Pr[A \cup B] \geq Pr[A] + Pr[B] - 1$$

ii. Generalize part (a) to prove that, for events $A_1, \ldots, A_n$ in the same probability space (and any n),

$$Pr[A_1 \cap \cdots \cap A_n] \geq Pr[A_1] + \cdots + Pr[A_n] - (n-1).$$

Answer: Now we generalize this result to show that:

$$Pr[A_1 \cap A_2 \cap \ldots A_n] \geq Pr[A_1] + Pr[A_2] + \cdots + Pr[A_n] - (n-1) \tag{1}$$

For this, we use induction on n.

The base case $n = 1$ just says that $Pr[A_1] \geq Pr[A_1] - 0$, which is trivially true. For our inductive hypothesis, we assume that equation (1) holds for some arbitrary n and any n events. The inductive step, therefore, is to show that it holds for $n+1$. In other words we need to show:

$$Pr[A_1 \cap A_2 \cap \cdots \cap A_n \cap A_{n+1}] \geq Pr[A_1] + Pr[A_2] + \cdots + Pr[A_n] + Pr[A_{n_1}] - n \tag{2}$$

Now, let B denote the event $A_n \cap A_{n+1}$. Then we have, by the inductive hypothesis applied to the n events $A_1, A_2, \ldots, A_{n-1}, B$,

$$Pr[A_1 \cap A_2 \cap \cdots \cap A_{n-1} \cap B] \geq Pr[A_1] + Pr[A_2] + \cdots + Pr[A_{n-1}] + Pr[B] - (n-1). \quad (3)$$

However, from our proof for the base case $n = 2$ in part (a) we have:

$$Pr[B] = Pr[A_n \cap A_{n+1}] \geq Pr[A_n] + Pr[A_{n+1}] - 1$$

Substituting this into equation (3) gives us our desired equation (2), which completes the induction proof.

8. **Birthdays (2/3/5 points)**
Suppose you record the birthdays of a large group of people, one at a time until you have found a match, i.e., a birthday that has already been recorded. (Assume there are 365 days in a year.)

(a) What is the probability that it takes more than 20 people for this to occur?

Answer: Pr[it takes more than 20 people] $=$ Pr[20 people don't have the same birthday] $= \frac{\frac{365!}{(365-20)!}}{365^{20}} = \boxed{\frac{\frac{365!}{345!}}{365^{20}}} \approx .589$

Another explanation that does not use counting:
Let b_i be the birthday of the i-th person.

$$\begin{aligned}
&\Pr[\text{it takes more than 20 people}] \\
=&\Pr[b_{20} \neq b_i \mid b_i\text{'s are all different } \forall 1 \leq i \leq 19] \times \Pr[b_i\text{'s are all different } \forall 1 \leq i \leq 19] \\
=&\Pr[b_{20} \neq b_i \mid b_i\text{'s are all different } \forall 1 \leq i \leq 19] \times \Pr[b_{19} \neq b_i \mid b_i\text{'s are all different } \forall 1 \leq i \leq 18] \times \\
&\quad \cdots \times \Pr[b_3 \neq b_i \mid b_i\text{'s are all different } \forall 1 \leq i \leq 2] \times \Pr[b_2 \neq b_1] \\
=&\frac{365-19}{365} \times \frac{365-18}{365} \times \cdots \times \frac{363}{365} \times \frac{364}{365} \\
\approx&.589
\end{aligned}$$

(b) What is the probability that it takes exactly 20 people for this to occur?

Answer: Pr[it takes exactly 20 people] $=$

Pr[first 19 have different birthdays **and** 20^{th}person shares a birthday with one of the first 19].
How total ways can the birthdays be chosen for 20 people? 365^{20}. How many ways can the birthdays be chosen so the first 19 have different birthdays and the 20^{th} person shares a birthday with the first 19? Well, the firs person has 365 choices, the second has 364 choices left, and so on until the nineteenth person has $(365-19+1) = 347$ choices left. Then, the 20^{th} person has 19 choices for his birthday. So in total, there are $365 \cdot 364 \cdot \cdots \cdot 348 \cdot 347 \cdot 19 = \frac{365!}{346!} \cdot 19$ ways of getting what we want. So Pr[it takes exactly 20 people] $= \frac{365 \cdot 364 \cdot \cdots \cdot 348 \cdot 347 \cdot 19}{365^{20}} = \boxed{\frac{\frac{365!}{346!} \cdot 19}{365^{20}}} \approx .032$

Another explanation that does not use counting:
Let b_i be the birthday of the i-th person.

$$\begin{aligned}
&\Pr[\text{it takes exactly 20 people}]\\
=&\Pr[b_{20} \text{ is equal to one of the } b_i\text{'s} \mid b_i\text{'s are all different } \forall 1 \le i \le 19] \times\\
&\Pr[b_i\text{'s are all different } \forall 1 \le i \le 19]\\
=&\Pr[b_{20} \text{ is equal to one of the } b_i\text{'s} \mid b_i\text{'s are all different } \forall 1 \le i \le 19] \times\\
&\Pr[b_{19} \ne b_i \mid b_i\text{'s are all different } \forall 1 \le i \le 18] \times \cdots \times\\
&\Pr[b_3 \ne b_i \mid b_i\text{'s are all different } \forall 1 \le i \le 2] \times \Pr[b_2 \ne b_1]\\
=&\frac{19}{365} \times \frac{365-18}{365} \times \cdots \times \frac{363}{365} \times \frac{364}{365}\\
\approx&.032
\end{aligned}$$

(c) Suppose instead that you record the birthdays of a large group of people, one at a time, until you have found a person whose birthday matches your own birthday. What is the probability that it takes exactly 20 people for this to occur?

Answer: Pr[it takes exactly 20 people] =

Pr[first 19 don't have your birthday **and** 20$^{\text{th}}$person has your birthday].

Similar to the last problem, there are 364 choices for the first person's birthday to be different than yours, 364 for the second person, and so on until the nineteenth person has 364 choices. Then, the 20$^{\text{th}}$ person has exactly 1 choice to have your birthday. So the total number of ways to get what we want is $364^{19} \cdot 1$. There are 365^{20} possibilities total. So Pr[it takes exactly 20 people] = $\boxed{\frac{364^{19}}{365^{20}}} \approx .0026$

Another explanation that does not use counting:

$$\begin{aligned}
\Pr[\text{it takes exactly 20 people}] =&\Pr[\text{the 1st person does not have the same birthday as yours}] \times\\
&\Pr[\text{the 2nd person does not have the same birthday as yours}] \times\\
&\cdots \times \Pr[\text{the 19th person does not have the same birthday as yours}] \times\\
&\Pr[\text{the 20th person has the same birthday as yours}]\\
=&\frac{364}{365} \times \frac{364}{365} \times \cdots \frac{364}{365} \times \frac{1}{365}\\
=&\frac{364^{19} \times 1}{365^{20}}\\
\approx&0.0026
\end{aligned}$$

9. **Blood Type (2/3/5 points)**
Consider the three alleles, A, B, and O, for human blood types. As each person inherits one of the 3 alleles from each parent, there are 6 possible genotypes: AA, AB, AO, BB, BO, and OO. Blood groups A and B are dominant to O. Therefore, people with AA or AO have type A blood. Similarly, BB and BO result in type B blood. The AB genotype is called type AB blood, and the OO genotype is called type O blood. Each parent contributes one allele randomly. Now, suppose that the frequencies

of the A, B, and O alleles are 0.4, 0.25, and 0.35, respectively, in Berkeley. Alice and Bob, two residents of Berkeley are married and have a daughter, Mary. Alice has blood type AB.

(a) What is the probability that Bob's genotype is AO?

Answer: Let B_{1A}, B_{1B} and B_{1O} be the events that Bob's first allele is A, B, and O, respectively. Let B_{2A}, B_{2B} and B_{2O} be the events that Bob's second allele is A, B, and O respectively. Bob's blood type can be AA, AB, AO, BB, or BO. Let B_{AA} be the event that Bob has type AA blood, B_{AB} be the event that Bob has type AB blood, B_{AO} be the event that Bob has type AO blood, B_{BB} be the event that Bob has type BB blood, and B_{BO} be the event that Bob has type BO blood. The sample space is $\Omega = \{B_{AA}, B_{AB}, B_{AO}, B_{BB}, B_{BO}\}$. Note that we have

$$\begin{aligned}
B_{AA} &= B_{1A} \cap B_{2A} \\
B_{AB} &= (B_{1A} \cap B_{2B}) \cup (B_{1B} \cap B_{2A}) \\
B_{AO} &= (B_{1A} \cap B_{2O}) \cup (B_{1O} \cap B_{2A}) \\
B_{BB} &= B_{1B} \cap B_{2B} \\
B_{BO} &= (B_{1B} \cap B_{2O}) \cup (B_{1O} \cap B_{2B}).
\end{aligned}$$

Since the first allele and second allele don't know about each other, the occurrence of the first allele will not affect the second, and vice versa. Therefore, using the rules that $P(A \cap B) = P(A|B)P(B)$ and $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,

$$\begin{aligned}
P(B_{AA}) &= P(B_{1A}|B_{2A})P(B_{2A}) = P(B_{1A})P(B_{2A}) = (.4)(.4) = .16 \\
P(B_{AB}) &= P(B_{1A}|B_{2B})P(B_{2B}) + P(B_{1B}|B_{2A})P(B_{2A}) - P((B_{1A} \cap B_{2B}) \cap (B_{1B} \cap B_{2A})) \\
&= P(B_{1A})P(B_{2B}) + P(B_{1B})P(B_{2A}) = 2(.4)(.25) = .2 \\
P(B_{AO}) &= P(B_{1A}|B_{2O})P(B_{2O}) + P(B_{1O}|B_{2A})P(B_{2A}) - P((B_{1A} \cap B_{2O}) \cap (B_{1O} \cap B_{2A})) \\
&= P(B_{1A})P(B_{2O}) + P(B_{1O})P(B_{2A}) = 2(.4)(.35) = .28 \\
P(B_{BB}) &= P(B_{1B}|B_{2B})P(B_{2B}) = P(B_{1B})P(B_{2B}) = (.25)(.25) = .0625 \\
P(B_{BO}) &= P(B_{1B}|B_{2O})P(B_{2O}) + P(B_{1O}|B_{2B})P(B_{2B}) - P((B_{1B} \cap B_{2O}) \cap (B_{1O} \cap B_{2B})) \\
&= P(B_{1B})P(B_{2O}) + P(B_{1O})P(B_{2B}) = 2(.25)(.35) = .175.
\end{aligned}$$

Therefore $P(B_{AO}) = .28$.

(b) Assume that Bob's genotype is AO. What is the probability that Mary's blood type is AB?

Answer: Since Alice has type AB and Bob has type AO, the sample space of possible genotypes for Mary is $\{AA, AO, AB, BO\}$. Since there is uniform probability of inheriting either allele from a given parent, there is a $1/4$ chance that Mary will have type AB blood.

(c) Assume Mary's blood type is AB. What is the probability that Bob's genotype is AA?

Answer: Bob's blood type can be AA, AB, AO, BB, or BO. As in part (a), let B_{AA} be the event that Bob has type AA blood, B_{AB} be the event that Bob has type AB blood, B_{AO} be the event that Bob has type AO blood, B_{BB} be the event that Bob has type BB blood, and B_{BO} be the event that Bob has type BO blood. We already computed the probability of Bob having these blood types in part (a):

$$\begin{aligned}
Pr(B_{AA}) &= .16 \\
Pr(B_{AB}) &= .2 \\
Pr(B_{AO}) &= .28 \\
Pr(B_{BB}) &= .0625 \\
Pr(B_{BO}) &= .175.
\end{aligned}$$

Now, let the event that Mary has blood type AB be M_{AB}. The problem asks us to find $Pr(B_{AA}|M_{AB})$. We can compute this using Bayes' formula, which says that

$$Pr(B_{AA}|M_{AB}) = \frac{Pr(M_{AB}|B_{AA}) \cdot Pr(B_{AA})}{Pr(M_{AB})}.$$

To find $Pr(M_{AB})$, we can use the Law of Total Probability, which says that

$$\begin{aligned}
Pr(M_{AB}) = Pr(M_{AB}|B_{AA}) \cdot Pr(B_{AA}) + Pr(M_{AB}|B_{AB}) \cdot Pr(B_{AB}) + Pr(M_{AB}|B_{AO}) \cdot Pr(B_{AO}) \\
+ Pr(M_{AB}|B_{BB}) \cdot Pr(B_{BB}) + Pr(M_{AB}|B_{BO}) \cdot Pr(B_{BO}).
\end{aligned}$$

To calculate this, we must find the conditional probabilities that Mary has AB blood given Bob's blood type. Recall that Alice has type AB blood.

- If Bob has AA blood, the possible combinations of their alleles are AA, AA, AB, and AB, so $Pr(M_{AB}|B_{AA}) = 1/2$.
- If Bob has AB blood, the possible combinations of their alleles are AA, AB, AB, and BB, so $Pr(M_{AB}|B_{AB}) = 1/2$.
- If Bob has AO blood, the possible combinations of their alleles are AA, AO, AB, and BO, so $Pr(M_{AB}|B_{AO}) = 1/4$.
- If Bob has BB blood, the possible combinations of their alleles are AB, AB, BB, and BB, so $Pr(M_{AB}|B_{BB}) = 1/2$.
- If Bob has BO blood, the possible combinations of their alleles are AB, AO, BB, and BO, so $Pr(M_{AB}|B_{BO}) = 1/4$.

We now have all the information we need to plug in and solve. By the Law of Total Probability above, we have

$$\begin{aligned}
Pr(M_{AB}) &= Pr(M_{AB}|B_{AA}) \cdot Pr(B_{AA}) + Pr(M_{AB}|B_{AB}) \cdot Pr(B_{AB}) + Pr(M_{AB}|B_{AO}) \cdot Pr(B_{AO}) \\
&\quad + Pr(M_{AB}|B_{BB}) \cdot Pr(B_{BB}) + Pr(M_{AB}|B_{BO}) \cdot Pr(B_{BO}) \\
&= (.5)(.16) + (.5)(.2) + (.25)(.28) + (.5)(.0625) + (.25)(.175) \\
&= .325,
\end{aligned}$$

and plugging in to Bayes' formula, we find that

$$Pr(B_{AA}|M_{AB}) = \frac{Pr(M_{AB}|B_{AA}) \cdot Pr(B_{AA})}{Pr(M_{AB})} = \frac{(.5)(.16)}{.325} = .246.$$

10. **Expressions (2/4/1/2/1 points)**

For each problem, just write down a mathematical expression.There is no need to justify/explain/derive the answer.

(a) **Bayes Rule - Man Speaks Truth**

i. A man speaks the truth 3 out of 4 times. He flips a biased coin that comes up Heads $\frac{1}{3}$ the time and reports it's Heads. What is the probability it is Heads?

ii. A man speaks the truth 3 out of 4 times. He rolls a fair 6-sided dice and reports it comes up 6. What is the probability it is really 6?

Answer:

i. Let E denote the event the man reports heads, S_1 be the event that the coin comes up heads, and S_2 be the event that the coin comes up tails.
We have: $P(E|S_1) = \frac{3}{4}$,$P(E|S_2) = \frac{1}{4}$,$P(S_2) = \frac{2}{3}$.
We want to compute $P(S_1|E)$, and let's do so by applying Bayes' Rule.

$$P(S_1|E) = \frac{P(S_1E)}{P(E)} = \frac{P(E|S_1)P(S_1)}{P(E|S_1)P(S_1)+P(E|S_2)P(S_2)} = \frac{3/4\times 1/3}{3/4\times 1/3+1/4\times 2/3} = \frac{3}{5}$$

ii. Let D be the event that the dice rolls a 6. Let M be the event that the man says 6.

$$P(D|M) = \frac{P(D\wedge M)}{P(M)} = \frac{P(M|D)P(D)}{P(M|D)P(D)+P(M|\neg D)P(\neg D)} = \frac{3/4\times 1/6}{3/4\times 1/6+1/4\times 1/5\times 5/6} = \frac{3}{4}$$

(b) **Unlikely events**

i. Toss a fair coin x times. What is the probability that you never get heads?
Answer: 0.5^x

ii. Roll a fair die x times. What is the probability that you never roll a six?
Answer: $(1-\frac{1}{6})^x$

iii. Suppose your weekly local lottery has a winning chance of $1/10^6$. You buy lottery from them for x weeks in a row. What is the probability that you never win?
Answer: $(1-\frac{1}{10^6})^x$

iv. How large must x be so that you get a head with probability at least 0.9? Roll a 6 with probability at least 0.9? Win the lottery with probability at least 0.9?
Answer: For coin, want: $0.5^x \leq 0.1$ so $x \geq \frac{\log 0.1}{\log 0.5} \approx 3.32$

For die, want: $(5/6)^x \leq 0.1$ so $x \geq \frac{\log 0.1}{\log 5/6} \approx 12.6$

For ticket, want: $(1-1/10^6)^x \leq 0.1$ so $x \geq \frac{\log 0.1}{\log(1-1/10^6)} \approx 2\times 10^6$.

(c) **Roll Dice**
You roll three fair six-sided dice. What is the probability of rolling a triple (all three dice agree)? What is the probability of rolling a double (two of the dice agree with each other)?

Answer: The sample space Ω consists of all possible outcomes of rolling 3 dies. Therefore, the size of it is: $|\Omega| = 6^3$. Let A be the event of rolling a triple, B be the event of rolling a double. The size of A is 6 since A consists of three ones', three twos', etc.

$$P[\text{rolling a triple}] = \frac{|A|}{|\Omega|} = \frac{6}{6^3}$$

The size of B is $6\cdot 5\cdot \frac{3!}{2!1!}$ because you have 6 ways to choose a number that appears twice in the roll, 5 ways to choose a number that is different the previous number and appears once in the roll. And you have $\frac{3!}{2!1!}$ possible different arrangement for two identical number and one distinct number. So we have

$$P[\text{rolling a double}] = \frac{|B|}{|\Omega|} = \frac{6 \cdot 5 \cdot 3}{6^3} = \frac{5}{12}$$

(d) **Lie Detector**
A lie detector is known to be 80% reliable when the person is guilty and 95% reliable when the person is innocent. If a suspect is chosen from a group of suspects of which only 1% have ever committed a crime, and the test indicates that the person is guilty, what is the probability that he is innocent?

Answer: Let A denote the event that the test indicates that the person is guilty, and B the event that the person is innocent. Note that

$$Pr[B] = 0.99, Pr[\overline{B}] = 0.01, Pr[A|B] = 0.05, Pr[A|\overline{B}] = 0.8$$

Using the Bayesian Inference Rule, we can compute the desired probability as follows:

$$Pr[B|A] = \frac{Pr[B]Pr[A|B]}{Pr[B]Pr[A|B] + Pr[\overline{B}]Pr[A|\overline{B}]} = \frac{0.99 \cdot 0.05}{0.99 \cdot 0.05 + 0.01 \cdot 0.8} \approx 0.86$$

(e) **Chess Squares**
Two squares are chosen at random on 8×8 chessboard. What is the probability that they share a side?

Answer: In 64 squares, there are:
(1) 4 at-corner squares, each has ONLY 2 squares each having a side in common with.
(2) $6 \cdot 4 = 24$ side squares, each has ONLY 3 squares such that each has a side in common with.
(3) $6 \cdot 6 = 36$ inner squares, each has 4 squares such that each has a side in common with.
Notice that the three cases are mutually exclusive. So we just sum up the probabilities.

$$\frac{4}{64} \cdot \frac{2}{63} + \frac{24}{64} \cdot \frac{3}{63} + \frac{36}{64} \cdot \frac{4}{63} = \frac{1}{18}$$