Recitation 9 solutions - CSCI 2200 (FOCS)

I. Suppose that A and B are independent.

(a) Which of the following pairs of events are independent?

First, remind that $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$ for independent probabilities.

(i) $A, \overline{B}$

$\mathbb{P}[A \cap \overline{B}] = \mathbb{P}[A] - \mathbb{P}[A \cap B]$
$= \mathbb{P}[A] - \mathbb{P}[A]\mathbb{P}[B]] = \mathbb{P}[A](1 - \mathbb{P}[B])$
$= \mathbb{P}[A]\mathbb{P}[\overline{B}]$
Thus independent.

(ii) $\overline{A}, \overline{B}$

$\mathbb{P}[\overline{A} \cap \overline{B}] = 1 - \mathbb{P}[A] - \mathbb{P}[B] + \mathbb{P}[A \cap B]$
$= (1 - \mathbb{P}[A])(1 - \mathbb{P}[B])$
$= \mathbb{P}[\overline{A}]\mathbb{P}[\overline{B}]$
Thus independent.

(iii) $\overline{A}, B$

Use the same step used in (i). Independent.

(iv) $A, \overline{A} \cap B$

if independent, $\mathbb{P}[A \cap (\overline{A} \cap B)] = \mathbb{P}[\emptyset] = 0$
this requires either $\mathbb{P}[A]$ or $\mathbb{P}[\overline{A} \cap B] = \mathbb{P}[\overline{A}]\mathbb{P}[B]$ is 0,
which is not generally true.

We also can say $\overline{A}$ only happens when A does not happen, which implies the dependency.

(v) $A, \overline{A} \cup B$

you can say the same thing here, or,
$\mathbb{P}[A \cap (\overline{A} \cup B)] = \mathbb{P}[(A \cap \overline{A}) \cup (A \cap B)] = \mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$
Thus not independent.

(b) Show that:

(i) $\mathbb{P}[A|\overline{B}] = \mathbb{P}[A]$

$$\mathbb{P}[A|\overline{B}] = \frac{\mathbb{P}[A \cap \overline{B}]}{\mathbb{P}[\overline{B}]}$$
$$= \frac{\mathbb{P}[A]\mathbb{P}[\overline{B}]}{\mathbb{P}[\overline{B}]}$$
$$= \mathbb{P}[A]$$

(ii) $\mathbb{P}[\overline{A}|B] = \mathbb{P}[\overline{A}]$

$$\mathbb{P}[\overline{A}|B] = \frac{\mathbb{P}[\overline{A} \cap B]}{\mathbb{P}[B]}$$
$$= \frac{\mathbb{P}[\overline{A}]\mathbb{P}[B]}{\mathbb{P}[B]}$$
$$= \mathbb{P}[\overline{A}]$$

(c) If A and B have strictly positive probabilities, can they be disjoint events?

Disjoint events means that:

$$\mathbb{P}[A]\mathbb{P}[B] = 0$$

...which only happens when either one is 0, counters that they should have strictly positive probabilites.

Thus, No.

(d) Can $\mathbb{P}[A] = 0$?

Yes, if A is independent.

(e) Show that $\mathbb{P}[\overline{A \ \cup \ B}] = \mathbb{P}[\overline{A}]\mathbb{P}[\overline{B}]$

$\mathbb{P}[\overline{A \cup B}] = 1 - \mathbb{P}[A \cup B]$
$= 1 - \mathbb{P}[A] - \mathbb{P}[B] + \mathbb{P}[A \cap B]$
By independency,
$= 1 - \mathbb{P}[A] - \mathbb{P}[B] + \mathbb{P}[A]\mathbb{P}[B]$
$= (1 - \mathbb{P}[A])(1 - \mathbb{P}[B])$
$= \mathbb{P}[\overline{A}]\mathbb{P}[\overline{B}]$

II. A 90-sided die with faces $1, \ldots, 90$ is rolled 6 times, Compute the probability that all rolls are different.

Total possibility will be: 90^6
For counting all different numbers (for each dice):

1) 90 total possibilities
2) 89 total possibilities
3) 88 total possibilities
4) 87 total possibilities
5) 86 total possibilities
6) 85 total possibilities

Thus, it will be:

$$\frac{90 \times 89 \times 88 \times 87 \times 86 \times 85}{90^6}$$

III

Joan and Tariq try to access a database at time steps 1, 2, 3, ... If both try to access the database, both get locked out for that time step. Joan and Tariq implement a randomized algorithm. Each independently attempts to access the database with probability p (independently at every time step). Let $J(i) = P[\text{Joan gains access to the database at time step } i]$. Similarly define $T(i)$ for Tariq. Let $A(i)$ be the probability that one of them gains access at time step i.

(i) Compute $J(i)$, $T(i)$, and $A(i)$. Set p to the value that maximizes $J(i)$.

For Joan to gain access at time step i, we need:

- Joan must attempt to access the database (probability p)
- Tariq must not attempt to access the database (probability $1 - p$)

Since Joan and Tariq make their decisions independently, we have:

$$\begin{aligned} J(i) &= P[\text{Joan attempts access}] \cdot P[\text{Tariq does not attempt access}] \\ &= p \cdot (1 - p) \end{aligned}$$

Tariq's probability is the same:

$$\begin{aligned} T(i) &= P[\text{Tariq attempts access}] \cdot P[\text{Joan does not attempt access}] \\ &= p \cdot (1 - p) \end{aligned}$$

The probability that someone gains access at time step i is:

$$\begin{aligned} A(i) &= P[\text{Exactly one person attempts access}] \\ &= P[\text{Joan attempts access and Tariq doesn't}] + P[\text{Tariq attempts access and Joan doesn't}] \\ &= p(1 - p) + p(1 - p) \\ &= 2p(1 - p) \end{aligned}$$

To find the value of p that maximizes $J(i) = p(1 - p)$, we take the derivative and set it equal to zero:

$$\begin{aligned} \frac{d}{dp}[p(1 - p)] &= \frac{d}{dp}[p - p^2] \\ &= 1 - 2p \\ 1 - 2p &= 0 \\ p &= \frac{1}{2} \end{aligned}$$

Therefore, $p = \frac{1}{2}$ maximizes $J(i)$. For this p, we have:

$$\begin{aligned} J(i) = T(i) &= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \\ A(i) &= 2 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2} \end{aligned}$$

(ii) Show that $P[\text{Joan waits } k \text{ steps for access}] = (\frac{3}{4})^{k-1}\frac{1}{4}$.

This means:

- Joan does not get access for the first $k - 1$ steps
- Joan gets access at step k

We already established that the probability of Joan getting access at any step is $J(i) = \frac{1}{4}$.
For Joan to not get access at a step, either:

- Joan doesn't attempt access (probability $1 - p = \frac{1}{2}$), OR
- Joan attempts access but Tariq also attempts access (probability $p \cdot p = \frac{1}{4}$)

Therefore:

$$\begin{aligned} P[\text{Joan doesn't get access at a step}] &= (1-p) + p \cdot p \\ &= \frac{1}{2} + \frac{1}{4} \\ &= \frac{3}{4} \end{aligned}$$

Since each step is independent, we have:

$$\begin{aligned} P[\text{Joan waits } k \text{ steps for access}] &= P[\text{No access for } k-1 \text{ steps}] \cdot P[\text{Access at step } k] \\ &= \left(\frac{3}{4}\right)^{k-1} \cdot \frac{1}{4} \end{aligned}$$

(iii) Show that $P[\textbf{First successful access is after } k \textbf{ steps}] = (\frac{1}{2})^k$.

For the first successful access to be after step k, there must be no successful access by either Joan or Tariq in the first k steps.

At each step, the probability that neither person gets access is:

$$\begin{aligned} P[\text{No access at step } i] &= 1 - A(i) \\ &= 1 - 2p(1-p) \\ &= 1 - 2 \cdot \frac{1}{2} \cdot \frac{1}{2} \\ &= 1 - \frac{1}{2} \\ &= \frac{1}{2} \end{aligned}$$

Since the steps are independent, we have:

$$\begin{aligned} P[\text{First successful access is after } k \text{ steps}] &= P[\text{No access in first } k \text{ steps}] \\ &= \left(\frac{1}{2}\right)^k \end{aligned}$$

IV

The surface of a sphere is arbitrarily painted red and blue with 90% of the surface painted red. Prove that it is possible to inscribe a cube whose vertices are all red.

(a) Consider a randomly inscribed cube with vertices $v_1, \ldots, v_8$. Show that $P[v_i \textbf{ is blue}] = 0.1$.

When we randomly inscribe a cube inside a sphere, each vertex of the cube touches the surface of the sphere. Since 90% of the sphere's surface is painted red and 10% is painted blue, and we are choosing the position of the vertices randomly, the probability that any specific vertex v_i is on a blue part of the sphere is equal to the proportion of the sphere that is blue:

$$P[v_i \text{ is blue}] = 0.1$$

This is because a random placement of the cube will place each vertex uniformly at random on the sphere's surface.

(b) Show that $P[v_1$ OR v_2 OR $\cdots$ OR v_8 is blue$] \leq 0.8$.

To find the probability that at least one of the eight vertices is blue, we can use the union bound. The probability of a union of events is at most the sum of the probabilities of the individual events.

Let B_i be the event that vertex v_i is blue. Then:

$$P[B_1 \cup B_2 \cup \cdots \cup B_8] \leq \sum_{i=1}^{8} P[B_i]$$

We know from part (a) that $P[B_i] = 0.1$ for each i. Therefore:

$$\begin{aligned} P[B_1 \cup B_2 \cup \cdots \cup B_8] &\leq \sum_{i=1}^{8} P[B_i] \\ &= \sum_{i=1}^{8} 0.1 \\ &= 8 \times 0.1 \\ &= 0.8 \end{aligned}$$

Thus, the probability that at least one vertex is blue is at most 0.8.

(c) Hence show that $P[$all eight vertices are red$] \geq 0.2$.

The event "all eight vertices are red" is the complement of the event "at least one vertex is blue". Therefore:

$$\begin{aligned} P[\text{all eight vertices are red}] &= 1 - P[\text{at least one vertex is blue}] \\ &= 1 - P[B_1 \cup B_2 \cup \cdots \cup B_8] \end{aligned}$$

From part (b), we determined that $P[B_1 \cup B_2 \cup \cdots \cup B_8] \leq 0.8$. Therefore:

$$\begin{aligned} P[\text{all eight vertices are red}] &= 1 - P[B_1 \cup B_2 \cup \cdots \cup B_8] \\ &\geq 1 - 0.8 \\ &= 0.2 \end{aligned}$$

Thus, the probability that all eight vertices are red is at least 0.2.

(d) What does it mean if a probability is positive?

If a probability is positive, it means that the event can occur. Specifically, if $P(E) > 0$ for some event E, then E is possible with a non-zero chance.In this question, since the probability of having all eight vertices be red is at least 0.2 (which is positive), this means it is definitely possible to inscribe a cube whose vertices are all red.

Problem V

Let X be the number of successes in n independent Bernoulli trials, each having success probability p. Thus, $X \sim \text{Binomial}(n, p)$, so

$$P(X = k) \;=\; \binom{n}{k} p^k \, (1-p)^{\,n-k} \quad \text{for } k = 0, 1, \ldots, n.$$

(a) *Probability of at least one success and at least one failure.*

$$P[X \geq 1] \;=\; 1 - P[X = 0] \;=\; 1 - (1-p)^n.$$
$$P[\text{at least one failure}] \;=\; 1 - P[\text{no failures}] \;=\; 1 - p^n.$$

(b) *Probability of an even number of successes and an even number of failures.*

$$P[X \text{ is even}] \;=\; \sum_{\substack{k=0 \\ k \text{ even}}}^{n} \binom{n}{k} p^k \, (1-p)^{n-k}.$$

We can use the fact that:

$$(p + (1-p))^n = 1 \quad \text{and} \quad ((1-p) - p)^n = (1-2p)^n,$$

which implies

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1, \quad \sum_{k=0}^{n} (-1)^k \binom{n}{k} p^k (1-p)^{n-k} = (1-2p)^n.$$

Adding these two sums picks out the even terms, giving:

$$P[X \text{ is even}] \;=\; \frac{1 + (1-2p)^n}{2}.$$

Since the number of failures is $n - X$, its parity depends on n:

$$P[\text{even number of failures}] \;=\; \begin{cases} P[X \text{ is even}] = \dfrac{1 + (1-2p)^n}{2}, & \text{if } n \text{ is even,} \\ 1 - P[X \text{ is even}] = \dfrac{1 - (1-2p)^n}{2}, & \text{if } n \text{ is odd.} \end{cases}$$

Solution to Problem VI

You and a friend each try independently at each time step to access a wireless channel, each with probability p. An *access* occurs at a step if exactly one of you tries (i.e. not both). Define:

$$X = (\text{number of steps until } \textit{your} \text{ first access}), \quad Y = (\text{number of steps until your friend's first a$$

We want the PDFs of X, Y, and $Z = X - Y$.

(a) *PDF of* X.

At each step, the probability that *you alone* try is $p(1-p)$. Thus X follows a (discrete) geometric distribution with parameter $r = p(1-p)$. Under the convention $X = 1$ means success on the first trial, its PDF is:

$$P(X = k) \;=\; [\,1 - p(1-p)\,]^{k-1}\, p(1-p), \quad k \in \mathbb{N}$$

(b) *PDF of* Y.

By symmetry, your friend's waiting time until they *alone* access follows the same distribution:

$$P(Y = k) \;=\; [\,1 - p(1-p)\,]^{k-1}\, p(1-p), \quad k \in \mathbb{N}$$

(c) *PDF of* $Z = X - Y$.

Since X and Y are i.i.d. geometric($p(1-p)$) random variables, the distribution of Z (the difference) is well-known to be:

$$P(Z = z) \;=\; \frac{p(1-p)}{2 - p(1-p)} \left[1 - p(1-p)\right]^{|z|}, \quad z \in \mathbb{Z}.$$

VII. Let the success probability in a trial be p. Let $\mathbf{X}$ be the waiting time for r successes. Derive the PDF of $\mathbf{X}$, i.e. compute $\mathbb{P}[\mathbf{X}=t]$.

(a) At which step is the rth success? In how many ways can you arrange the first $r-1$ successes?

Let's suppose that the rth success occurs at time t. Therefore, we know that from times 1 to $t-1$, there must be $r-1$ successes. Since any subset of size $r-1$ of the $t-1$ times can be success set, we have $\binom{t-1}{r-1}$ ways to arrange the first $r-1$ successes.

(b) Show that $P_{\mathbf{X}}(t) = \binom{t-1}{r-1}p^r(1-p)^{t-r}$.

Using our result from the previous part, the probability of waiting t time steps for r success is equal to the probability of r successes times the probability of $t-r$ failures times the number of ways the first $r-1$ successes can be arranged. This gives us

$$P_{\mathbf{X}}(t) = \binom{t-1}{r-1}p^r(1-p)^{t-r}$$

(c) Show that the above formula matches the PDF of the waiting time for one success when $r=1$.

When $r=1$, we have that

$$\begin{aligned} P_{\mathbf{X}}(t) &= \binom{t-1}{1-1}p^r(1-p)^{t-1} \\ &= p^r(1-p)^{t-1} \end{aligned}$$

which matches the PDF of a geometric random variable with success probability p.

VIII. The number of fish in a lake is $\mathbf{X} \in [200, 400]$, with each value of $\mathbf{X}$ being equally likely. In a study, a biologist randomly caught 30 fish (without replacement), marked them, and replaced them in the lake. The next day, the biologist caught 30 random fish and found 20 to be marked.

(a) What was the PDF of $\mathbf{X}$ before the biologist did anything?

Since our random variable $\mathbf{X}$ follows a uniform distribution and there are 201 equally likely possibilities for what the value of $\mathbf{X}$ can be, our PDF of $\mathbf{X}$ is given as follows:

$$P_{\mathbf{X}}(x) = \begin{cases} \frac{1}{201} & \text{if } x \in [200, 400] \\ 0 & \text{otherwise} \end{cases}$$

(b) What is the updated PDF of $\mathbf{X}$ after finding that 20 fish in the new sample are marked? (give a plot)

What we are now looking for is the probability that $\mathbf{X} = x$ given that 20 fish in the new sample are marked. By Bayes' Theorem,

$$\mathbb{P}[\mathbf{X} = x \mid 20 \text{ fish marked}] = \frac{\mathbb{P}[20 \text{ fish marked} \mid \mathbf{X} = x]\,\mathbb{P}[\mathbf{X} = x]}{\mathbb{P}[20 \text{ fish marked}]}$$

To find $\mathbb{P}[20 \text{ fish marked} \mid \mathbf{X} = x]$, we can divide the number of favorable outcomes by the number of total outcomes. Given that we have x fish in the lake and that 30 of them are marked, the number of ways we can choose 20 marked fish and 10 unmarked fish is given by $\binom{30}{20}\binom{x-30}{10}$. The total number of ways to choose 30 fish from our x total fish is given by $\binom{x}{30}$. Therefore

$$\mathbb{P}[20 \text{ fish marked} \mid \mathbf{X} = x] = \frac{\binom{30}{20}\binom{x-30}{10}}{\binom{x}{30}}$$

We already found that $\mathbb{P}[\mathbf{X} = x] = \frac{1}{201}$ if $x \in [200, 400]$ and since we are only considering x values in this range, we can just say that $\mathbb{P}[\mathbf{X} = x] = \frac{1}{201}$.

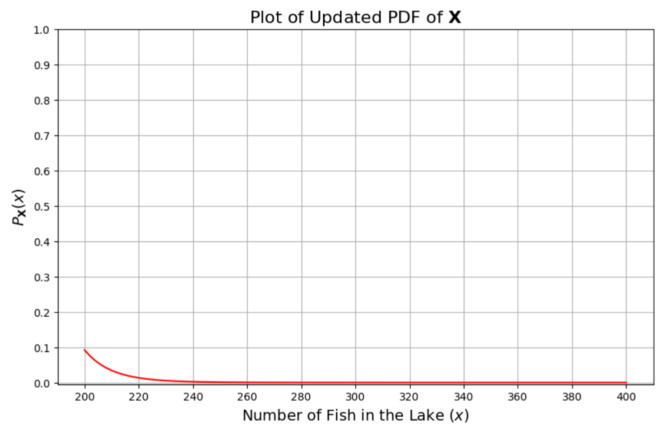
To find $\mathbb{P}[20 \text{ fish marked}]$, we can use the law of total probabilities which states that

$$\begin{aligned}\mathbb{P}[20 \text{ fish marked}] &= \sum_{k=200}^{400} \mathbb{P}[20 \text{ fish marked} \mid \mathbf{X} = k]\,\mathbb{P}[\mathbf{X} = k] \\ &= \frac{1}{201}\sum_{k=200}^{400} \frac{\binom{30}{20}\binom{k-30}{10}}{\binom{k}{30}}\end{aligned}$$

Therefore our final PDF for the updated value of $\mathbf{X}$ is

$$\begin{aligned}\mathbb{P}[\mathbf{X} = x \mid 20 \text{ fish marked}] &= \frac{\frac{1}{201} \cdot \frac{\binom{30}{20}\binom{x-30}{10}}{\binom{x}{30}}}{\frac{1}{201}\sum_{k=200}^{400} \frac{\binom{30}{20}\binom{k-30}{10}}{\binom{k}{30}}} \\ &= \frac{\binom{x-30}{10}}{\binom{x}{30}\sum_{k=200}^{400} \frac{\binom{k-30}{10}}{\binom{k}{30}}}\end{aligned}$$

Here is a plot of this updated PDF



(c) What is the most likely number of fish in the lake?

The most likely number of fish in the lake is simply the value that maximizes our new PDF for $\mathbf{X}$. Given our graph, it is clear that this value is 200.