

Math for CS 2015/2019 solutions to “In-Class Problems Week 13, Fri. (Session 33)”

<https://github.com/spamegg1>

November 21, 2022

Contents

1	Problem 1	2
1.1	(a)	2
1.2	(b)	2
1.3	(c)	3
2	Problem 2	3
2.1	(a)	3
2.2	(b)	3
2.3	(c)	3
2.4	(d)	4
3	Problem 3	4
3.1	(a)	4
3.2	(b)	4
3.3	(c)	4
3.4	(d)	5
3.5	(e)	5
3.6	(f)	5
4	Problem 4 (Supplementary Problem)	6
4.1	(a)	6
4.2	(b)	6
4.3	(c)	6
4.4	(d)	7
4.5	(e)	8
4.6	(f)	8
4.7	(g)	8
5	Problem 5 (Supplementary Problem)	9
5.1	(a)	9

5.2	(b)	9
5.3	(c)	9

1 Problem 1

A herd of cows is stricken by an outbreak of cold cow disease. The disease lowers a cow's body temperature from normal levels, and a cow will die if its temperature goes below 90 degrees F. The disease epidemic is so intense that it lowered the average temperature of the herd to 85 degrees. Body temperatures as low as 70 degrees, but no lower, were actually found in the herd.

1.1 (a)

Use Markov's Bound to prove that at most $3/4$ of the cows could survive.

Proof. Let T be the temperature of a random cow. Then the fraction of cows that survive is the probability that $T \geq 90$, and $E[T]$ is the average temperature of the herd.

Applying Markov's Bound to T :

$$Pr[T \geq 90] \leq \frac{E[T]}{90} = \frac{85}{90} = \frac{17}{18}$$

But $17/18 > 3/4$, so this bound is not good enough.

Instead, we apply Markov's Bound to $T - 70$:

$$Pr[T \geq 90] = Pr[T - 70 \geq 20] \leq \frac{E[T - 70]}{20} = \frac{85 - 70}{20} = 3/4$$

□

1.2 (b)

Suppose there are 400 cows in the herd. Show that the bound from part (a) is the best possible by giving an example set of temperatures for the cows so that the average herd temperature is 85 and $3/4$ of the cows will have a high enough temperature to survive.

Proof. Let 100 cows have temperature 70 degrees and 300 have 90 degrees. So the probability that a random cow has a high enough temperature to survive is exactly $3/4$. Also, the mean temperature is $(1/4)70 + (3/4)90 = 85$. So this distribution of temperatures satisfies the conditions under which the Markov bound implies that the probability of having a high enough temperature to survive cannot be larger than $3/4$. □

1.3 (c)

Notice that the results of part (b) are purely arithmetic facts about averages, not about probabilities. But you verified the claim in part (a) by applying Markov's bound on the deviation of a random variable. Justify this approach by regarding the temperature, T , of a cow as a random variable. Carefully specify the probability space on which T is defined: what are the outcomes? what are their probabilities? Explain the precise connection between properties of T and average herd temperature that justifies the application of Markov's Bound.

Proof. ???

□

2 Problem 2

A gambler plays 120 hands of draw poker, 60 hands of black jack, and 20 hands of stud poker per day. He wins a hand of draw poker with probability $1/6$, a hand of black jack with probability $1/2$, and a hand of stud poker with probability $1/5$.

2.1 (a)

What is the expected number of hands the gambler wins in a day?

Proof. $120(1/6) + 60(1/2) + 20(1/5) = 54$.

□

2.2 (b)

What would the Markov bound be on the probability that the gambler will win at least 108 hands on a given day?

Proof. The expected number of games won is 54, so by Markov, $\Pr[R \geq 108] \leq 54/108 = 1/2$.

□

2.3 (c)

Assume that the outcomes of the card games are pairwise, but possibly not mutually, independent. What is the variance of the number of hands won per day? You may answer with a numerical expression that is not completely evaluated.

Proof. The variance can also be calculated using linearity of variance. For an individual hand the variance is $p(1-p)$ where p is the probability of winning. Therefore the variance is $120(1/6)(5/6) + 60(1/2)(1/2) + 20(1/5)(4/5) = 523/15$.

□

2.4 (d)

What would the Chebyshev bound be on the probability that the gambler will win at least 108 hands on a given day? You may answer with a numerical expression that is not completely evaluated.

Proof.

$$\Pr[R - 54 \geq 54] \leq \Pr[|R - 54| \geq 54] \leq \frac{\text{Var}[R]}{54^2} = \frac{523}{(15)(54)^2}$$

□

3 Problem 3

The hat-check staff has had a long day serving at a party, and at the end of the party they simply return the n checked hats in a random way such that the probability that any particular person gets their own hat back is $1/n$.

Let X_i be the indicator variable for the i th person getting their own hat back. Let S_n be the total number of people who get their own hat back.

3.1 (a)

What is the expected number of people who get their own hat back?

Proof. $\Pr[X_i = 1] = 1/n$, so by the linearity of expectation

$$\text{Ex}[S_n] = \text{Ex}[X_1 + \dots + X_n] = \text{Ex}[X_1] + \dots + \text{Ex}[X_n] = n \cdot \frac{1}{n} = 1$$

□

3.2 (b)

Write a simple formula for $\text{Ex}[X_i \cdot X_j]$ for $i \neq j$.

Hint: What is $\Pr[X_j = 1 \mid X_i = 1]$?

Proof. The probability that X_i and X_j are both 1 is $\frac{1}{n} \cdot \frac{1}{n-1}$. Thus $X_i \cdot X_j = 1$ with probability $\frac{1}{n(n-1)}$ and 0 otherwise. So $\text{Ex}[X_i \cdot X_j] = \frac{1}{n(n-1)}$. □

3.3 (c)

Explain why you cannot use the variance of sums formula to calculate $\text{Var}[S_n]$.

Proof. The indicator random variables, X_i , are not even pairwise independent. This can be seen by comparing the marginal and conditional probability of a particular person, Alice, getting her hat back. The marginal probability, unconditioned on any

other events, is $1/n$ as we've computed before. However, if compute this probability conditioned on the event that a second person, Bob, got his hat back, we find that the probability of Alice getting her hat back is $1/(n-1)$. \square

3.4 (d)

Show that $\text{Ex}[(S_n)^2] = 2$. Hint: $(X_i)^2 = X_i$.

Proof.

$$\begin{aligned} \text{Ex}[(S_n)^2] &= \sum_{i=1}^n \text{Ex}[(X_i)^2] + \sum_{i=1}^n \sum_{j \neq i}^n \text{Ex}[X_i \cdot X_j] \\ &= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} \\ &= 2 \end{aligned}$$

\square

3.5 (e)

What is the variance of S_n ?

Proof.

$$\begin{aligned} \text{Var}[S_n] &= \text{Ex}[(S_n)^2] - (\text{Ex}[S_n])^2 \\ &= 2 - (n \cdot (1/n))^2 \\ &= 1 \end{aligned}$$

\square

3.6 (f)

Show that there is at most a 1% chance that more than 10 people get their own hat back.

Proof.

$$\begin{aligned} \text{Pr}[S_n \geq 11] &= \text{Pr}[S_n - \text{Ex}[S_n] \geq 11 - \text{Ex}[S_n]] \\ &= \text{Pr}[S_n - \text{Ex}[S_n] \geq 10] \\ &= \frac{\text{Var}[S_n]}{10^2} = 0.01 \end{aligned}$$

Note that the X_i 's are Bernoulli variables but are not independent, so S_n does not have a binomial distribution and the binomial estimates from Lecture Notes do not apply. \square

4 Problem 4 (Supplementary Problem)

Let K_n be the complete graph with n vertices. Each of the edges of the graph will be randomly assigned one of the colors red, green, or blue. The assignments of colors to edges are mutually independent, and the probability of an edge being assigned red is r , blue is b , and green is g (so $r + b + g = 1$).

A set of three vertices in the graph is called a triangle. A triangle is monochromatic if the three edges connecting the vertices are all the same color.

4.1 (a)

Let m be the probability that any given triangle, T , is monochromatic. Write a simple formula for m in terms of r, b , and g .

Proof. $m = r^3 + b^3 + g^3$ □

4.2 (b)

Let I_T be the indicator variable for whether T is monochromatic. Write simple formulas in terms of m, r, b , and g for $\text{Ex}[I_T]$ and $\text{Var}[I_T]$.

Proof. $\text{Ex}[I_T] = \Pr[T \text{ is monochromatic}] = m$

$\text{Var}[I_T] = \text{Ex}[(I_T)^2] - (\text{Ex}[I_T])^2 = \text{Ex}[I_T] - (\text{Ex}[I_T])^2 = m - m^2$ □

Let T and U be distinct triangles.

4.3 (c)

What is the probability that T and U are both monochromatic if they do not share an edge? . . . if they do share an edge?

Proof. If T and U do not share an edge, T being monochromatic is independent of U being monochromatic. Therefore

$$\Pr[T \text{ and } U \text{ are both monoch.}] = \Pr[T \text{ is monoch.}] \cdot \Pr[U \text{ is monoch.}] = m^2$$

If they share an edge, there are 3 possibilities. Either they share 1 edge, or 2 edges, or 3 edges (which is impossible because T and U are distinct triangles).

Assume they share one edge e .

If e has color red, then the probability that the other two edges of T are also red is r^2 ; and the probability that the other two edges of U are also red is r^2 . Notice these two events are independent (the other two edges are not common to T and U).

$$\Pr[T \text{ and } U \text{ are both monoch.} \mid e \text{ has color red}] = r^4.$$

Similar arguments for the other two color cases for e . Therefore

$$\begin{aligned}
& \Pr[T \text{ and } U \text{ are both monoch} \mid T \text{ and } U \text{ have only one edge } e \text{ in common}] = \\
& \Pr[e \text{ has color red}] \cdot \Pr[T \text{ and } U \text{ are both monoch.} \mid e \text{ has color red}] \\
& + \Pr[e \text{ has color blue}] \cdot \Pr[T \text{ and } U \text{ are both monoch.} \mid e \text{ has color blue}] \\
& + \Pr[e \text{ has color green}] \cdot \Pr[T \text{ and } U \text{ are both monoch.} \mid e \text{ has color green}] \\
& = r \cdot r^4 + b \cdot b^4 + g \cdot g^4 = r^5 + b^5 + g^5.
\end{aligned}$$

We can give a similar analysis when T and U have two edges in common, say e, f . Now there are 9 possibilities for the colors of e, f .

$\Pr[e \text{ and } f \text{ are both red}] = r^2$, and the probability that the remaining third edge of T is also red is r , and the probability that the remaining third edge of U is also red is r (and these two events are independent).

And so on. With this analysis we see that

$$\Pr[T \text{ and } U \text{ are both monoch.} \mid T \text{ and } U \text{ have 2 edges in common}] = r^4 + b^4 + g^4. \quad \square$$

Now assume $r = b = g = 1/3$.

4.4 (d)

Show that I_T and I_U are independent random variables.

Proof. Need to show that the events:

1. $[I_T = 0]$ and $[I_U = 0]$ are independent,
2. $[I_T = 0]$ and $[I_U = 1]$ are independent,
3. $[I_T = 1]$ and $[I_U = 0]$ are independent,
4. $[I_T = 1]$ and $[I_U = 1]$ are independent.

We will only show (2), the others are similar.

For (2) we need to show that $\Pr[(I_T = 0) \cap (I_U = 1)] = \Pr[I_T = 0] \cdot \Pr[I_U = 1]$.

Consider the event $(I_T = 0) \cap (I_U = 1)$. This is the event $[T \text{ is not monoch. AND } U \text{ is monoch.}]$ There are 4 cases for this event: T and U have 0,1,2,3 edges in common (the last one is impossible since T and U are distinct triangles).

If T and U have 0 edges in common, then the event $[T \text{ is not monoch.}]$ and the event $[U \text{ is monoch}]$ are independent. So we have $\Pr[(I_T = 0) \cap (I_U = 1)] = \Pr[I_T = 0] \cdot \Pr[I_U = 1]$ as needed.

If T and U have 1 edge in common, say e , consider 3 cases for e : red, blue, green. Consider the red case. $\Pr[\text{the other two edges of } T \text{ are not both red}]$ is $1 - r^2$. $\Pr[\text{the other two edges of } U \text{ are both red}]$ is r^2 . These two events (about the other two edges) are independent, so in this case $\Pr[(I_T = 0) \cap (I_U = 1)] = (1 - r^2)r^2$. Similar for the other two colors. So

$$\Pr[(I_T = 0) \cap (I_U = 1)] = r \cdot (1 - r^2)r^2 + b \cdot (1 - b^2)b^2 + g \cdot (1 - g^2)g^2.$$

Plugging-in $r = b = g = 1/3$ we get $\Pr[(I_T = 0) \cap (I_U = 1)] = \frac{24}{243} = \frac{8}{81}$.

Now $\Pr[I_T = 0] = r \cdot (1 - r^2) + b \cdot (1 - b^2) + g \cdot (1 - g^2) = \frac{24}{27} = \frac{8}{9}$ and $\Pr[I_U = 1] = r \cdot r^2 + b \cdot b^2 + g \cdot g^2 = \frac{3}{27} = \frac{1}{9}$ so $\Pr[I_T = 0] \cdot \Pr[I_U = 1] = \frac{8}{81}$.

So $\Pr[(I_T = 0) \cap (I_U = 1)] = \Pr[I_T = 0] \cdot \Pr[I_U = 1]$ as needed.

Finally the case where T and U have 2 edges e, f in common follows from a similar argument.

So $[I_T = 0]$ and $[I_U = 1]$ are independent, which proves (2). \square

4.5 (e)

Let M be the number of monochromatic triangles. Write simple formulas in terms of n and m for $\text{Ex}[M]$ and $\text{Var}[M]$.

Proof. First of all, how many triangles are there in K_n ? Since it's a complete graph, every set of three vertices form a triangle. So there are $\binom{n}{3}$ triangles.

The probability that any given triangle T is monochromatic is m by part (a).

Therefore $\text{Ex}[M] = \binom{n}{3}m$. (Similarly $\text{Ex}[M^2] = \binom{n}{3}^2 m$.)

$$\text{Var}[M] = \text{Ex}[M^2] - (\text{Ex}[M])^2 = \binom{n}{3}^2 m - \left(\binom{n}{3}m\right)^2 = \binom{n}{3}^2 (m - m^2). \quad \square$$

4.6 (f)

Let $\mu := \text{Ex}[M]$. Use Chebyshev's Bound to prove that

$$\Pr[|M - \mu| > \sqrt{\mu \log(\mu)}] \leq \frac{1}{\log(\mu)}$$

Proof. By Chebyshev's Bound

$$\Pr[|M - \mu| > \sqrt{\mu \log(\mu)}] \leq \frac{\text{Var}[M]}{\mu \log(\mu)}$$

So we need to show $\text{Var}[M] \leq \mu$, in other words $\text{Var}[M] \leq \text{Ex}[M]$.

???

\square

4.7 (g)

Conclude that

$$\lim_{n \rightarrow \infty} \Pr[|M - \mu| > \sqrt{\mu \log(\mu)}] = 0$$

Proof. $\log(\mu) = \log\left(\binom{n}{3}m\right) = \log(n(n-1)(n-2)m/6)$, so

$$\lim_{n \rightarrow \infty} \Pr[|M - \mu| > \sqrt{\mu \log(\mu)}] \leq \lim_{n \rightarrow \infty} \frac{1}{\log(n(n-1)(n-2)m/6)} = 0$$

therefore $\lim_{n \rightarrow \infty} \Pr[|M - \mu| > \sqrt{\mu \log(\mu)}] = 0$. □

5 Problem 5 (Supplementary Problem)

Let R be a positive integer valued random variable.

5.1 (a)

If $\text{Ex}[R] = 2$, how large can $\text{Var}[R]$ be?

Proof. ??? □

5.2 (b)

How large can $\text{Ex}[1/R]$ be?

Proof. ??? □

5.3 (c)

If $R \leq 2$, that is, the only values of R are 1 and 2, how large can $\text{Var}[R]$ be?

Proof. ??? □