COL202: Discrete Math **Tutorial 7 Solutions**

Instructor: Venkata Koppula

1 Tutorial Submission Problem (⋆)

The following problem has a few parts. You only need to submit the ones that are (⋆) marked.

In our lectures so far, we have computed the expected value of several random variables. Why do we care about the expected value of the random variable? By definition, this is the *average value* of the random variable. Can we conclude that if we take a random sample w from the distribution, then $X(w)$ will be close to $\mathbb{E}[X]$? Not always. For instance, if you sample a number uniformly at random from $\{1, 2, \ldots, n\}$, then the expected value of the sample is $(n+1)/2$. However, the sampled number can be anything from 1 to n (with equal probability).
However, in many cases that arise in CS, we have some more information about the distribution, and in such cases, we can prove that the sampled value is concentrated near the expected value! These are called *concentration inequalities* and are extensively used in the analysis of randomized algorithms and processes.
For this, we will start with a simple inequality named Markov's inequality (which you may have seen in high school/other courses).

1.1. Markov Inequality

For a *non-negative* random variable X with $\mathbb{E}[X] = \mu$ and $\epsilon > 0$ prove that

$$\Pr[X \geq \epsilon] \leq \frac{\mu}{\epsilon}$$

Solution: We are proving for the discrete case. The continuous case is analogous

$$\mathbb{E}[X] = \sum_x x \Pr[X = x] = \sum_{x<\epsilon} x \Pr[X = x] + \sum_{x \geq \epsilon} x \Pr[X = x]$$

Since X is non-negative, $\sum_{x<\epsilon} x \Pr[X = x] \geq 0$, so that

$$\mathbb{E}[X] \geq \sum_{x \geq \epsilon} x \Pr[X = x] \geq \sum_{x \geq \epsilon} \epsilon \Pr[X = x] = \epsilon \Pr[X \geq \epsilon]$$

form which we obtain the result

Note that Markov's inequality can only be applied for non-negative random variables. Moreover, the bound is meaningless for $\epsilon \leq \mu$. However, it is useful for proving other concentration inequalities where we have some more information about the random variable.

Definition 1. *Let (Ω, p) be a discrete probability distribution, and $X : \Omega \to \mathbb{R}$ any random variable such that $\mathbb{E}[X] = \mu$. The variance of X, $\mathsf{Var}[X]$, is defined as $\mathbb{E}[(X-\mu)^2]$.*

Note that $\mathsf{Var}[X] = \mathbb{E}\left[X^2\right] - 2\mu\mathbb{E}[X] + \mu^2 = \mathbb{E}\left[X^2\right] - \mu^2$.

1.2. Chebyshev Inequality

For any random variable X with $\mathbb{E}[X] = \mu$ and $\mathsf{Var}[X] = \sigma^2$ and $\epsilon > 0$ prove that

$$\Pr[|X-\mu| \geq \epsilon] \leq \frac{\sigma^2}{\epsilon^2}$$

Solution: Take $Y = (X-\mu)^2$, then $\mathbb{E}[Y^2] = \mathbb{E}[(X-\mu)^2] = \sigma^2$. By Markov's inequality,

$$\Pr[Y \geq \epsilon^2] \leq \frac{\mathbb{E}[Y^2]}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}$$

Which gives the result.

But sometimes, even the guarantees given by Chebyshev are not enough. Note that in order to apply Chebyshev's inequality, we only need a bound on the variance. What if you had more information about the random variable? Recall, in many of the examples that we saw in class, the random variable can be decomposed a sum of simpler random variables. If these random variables are *independent*, then we can prove much stronger concentration bounds. First, let us define independence of random variables.

Definition 2. *Let X, Y be random variables corresponding to discrete probability distribution (Ω, p). We say that X and Y are independent random variables if for all $a \in \mathsf{Supp}(X)$ and $b \in \mathsf{Supp}(Y)$,*

$$\Pr[(X = a) \wedge (Y = b)] = \Pr[X = a] \cdot \Pr[Y = b].$$

For independent random variables X and Y, we can express the expectation of the product of random variables as the product of the expectations.

Theorem 1. *For any random variables X and Y, $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.*

The proof of this is left as an exercise. Note that, unlike linearity of expectation, the above holds only for random variables that are independent.

We are now ready to state, and prove, strong concentration bounds.

1.3. ($\star$) Chernoff Bounds

Let $X_1, X_2 \ldots X_n$ be independent random variables such that $\Pr[X_i = 1] = p_i$ and $\Pr[X_i = 0] = 1 - p_i$, where $0 < p_i < 1$ for all $i \in [n]$. Let $X = \sum_{i=1}^{n} X_i$, with

$\mu = \mathbb{E}[X] = \sum_{i=1}^{n} p_i$. Then prove that

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu} \qquad \text{for } \delta > 0$$

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu} \qquad \text{for } 0 < \delta < 1$$

Proving this requires some tools from high-school mathematics, together with Markov's inequality and Theorem 1. First, let us consider $\Pr[X > (1+\delta)\mu]$. Note that for any $t > 0$,

$$\Pr[X > (1+\delta)\mu] = \Pr\left[e^{tX} > e^{t(1+\delta)\mu}\right].$$

Next, you should simplify this (using Markov's inequality, Theorem 1 and the fact that $1 + x < e^{x}$). Finally, you will get an upper bound on $\Pr[X > (1+\delta)\mu]$ in terms of t and the p_is. This bound holds for all positive values of t, and therefore, the best upper bound would be the one obtained by finding a positive t that minimizes this expression. Find this minimal value for t (in terms of δ).

Solution: Using Markov's inequality:

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}] \leq e^{-t(1+\delta)\mu}\mathbb{E}[e^{tX}]$$

Now, due to independence of the random variables

$$\mathbb{E}[e^{tX}] = \prod_{i=1}^{n}\left(\mathbb{E}[e^{tX_i}]\right)$$

and

$$\mathbb{E}[e^{tX_i}] = e^{t}p_i + (1-p_i) = 1 + p_i(e^{t}-1) \leq e^{p_i(e^{t}-1)}$$

Then

$$\mathbb{E}[e^{tX}] = e^{\sum_i p_i(e^{t}-1)} = e^{\mu(e^{t}-1)}$$

Then

$$\Pr[X \geq (1+\delta)\mu] \leq \inf_{t} e^{-t(1+\delta)\mu+\mu(e^{t}-1)}$$

On differentiating wrt t, we see that the minima is obtained at $t = \ln(1+\delta)$ which gives the result. The other inequality can be obtained similarly.

The above form is often too cumbersome to work with. So we use some looser bounds that can be derived from the above inequalities using $\frac{2\delta}{2+\delta} \leq \ln(1+\delta)$. (You don't need to submit these for the tutorial, though you are encouraged to derive them on your

own)

$$\Pr[X \geq (1+\delta)\mu] \leq e^{-\delta^2\mu/(2+\delta)}$$
$$\Pr[X \leq (1-\delta)\mu] \leq e^{-\delta^2\mu/2}$$
$$\Pr[|X-\mu| \geq \delta\mu] \leq 2e^{-\delta^2\mu/3}$$

2 Problems - General Probability

2.1. **Random permutations** Let σ be a uniformly random permutation over n elements. For any $1 \leq k \leq n$, what is the expected number of k-cycles in a uniformly random permutation?

Solution:
Hint:

Calculate the probability that a specific k-tuple forms a k-cycle in a permutation and use linearity of expectation

Answer:

The expected number of k-cycles is $\frac{1}{k}$.

2.2. **Balls in Bins:** Suppose we have n identical balls and n identical bins. We throw each ball to one of the bins uniformly at random. Let X_i denote the number of balls in i^{th} bin, and let $X = \max_i X_i$. Prove that $\mathbb{E}[X] = \Theta\left(\frac{\log n}{\log\log n}\right)$.

Solution: We will take the balls and bins to be distinct in the analysis. Observe that the expected maximum load will remain same in either case.

Upper Bound: First, we will show that there exists a constant $c > 0$ such that $\mathbb{E}[X] \leq \frac{c\log n}{\log\log n}$.
Let us first approximate the probability that a bin has $\geq k$ balls. Using a union bound over all subsets of $[n]$ of size k, we get

$$\Pr[X_i \geq k] \leq \binom{n}{k}\frac{1}{n^k}$$

Further, since $\binom{n}{k} \leq \frac{n^k}{k!}$,

$$\Pr[X_i \geq k] \leq \frac{1}{k!}$$

Taking $k = c\frac{\log n}{\log\log n}$, observe that using Stirling's approximation $k! \geq n^2$. Thus using a union bound again:

$$\Pr[\exists i : X_i \geq k] \leq \frac{n}{k!} \leq \frac{1}{n}$$

$$\implies \Pr[X \geq k] \leq \frac{1}{n}$$

Finally

$$\mathbb{E}[X] \leq \Pr[X < k](k-1) + \Pr[X \geq k]n \leq 1(k-1) + \frac{1}{n}n = k$$

which gives the required upper bound.

Lower Bound: We will now prove that there exists a constant $d > 0$ such that $\mathbb{E}[X] \geq \frac{d \log n}{\log \log n}$. Note that it suffices to show that $\Pr\left[X \leq \frac{\log n}{100 \log \log n}\right] \leq 0.5$, because then $\mathbb{E}[X] \geq \frac{\log n}{200 \log \log n}$.

Let Y denote the number of bins with at least $z = \frac{\log n}{100 \log \log n}$ balls. Suppose $\mathbb{E}[Y] = \mu$ and $\mathsf{Var}[Y] = \sigma^2$. Using Chebyshev's inequality,

$$\Pr[Y = 0] \leq \Pr\left[|Y - \mu| \geq \mu\right] \leq \frac{\sigma^2}{\mu^2}$$

Hence it suffices to provide an upper bound on σ^2 and a lower bound on μ.

2.3. (♦) Let $N \in \mathbb{N}$, and let $2 = p_1 < p_2 < \ldots < p_t \leq N$ be the set of prime numbers in $[N]$. Suppose n is sampled uniformly at random from $[N]$. Let X denote the number of *distinct* prime factors of n. Compute $\mathbb{E}[X]$ (in terms of $p_1, p_2, \ldots, p_t$).

Solution:
Hint:

Use indicator random variables for each prime and apply linearity of expectation.

Answer:

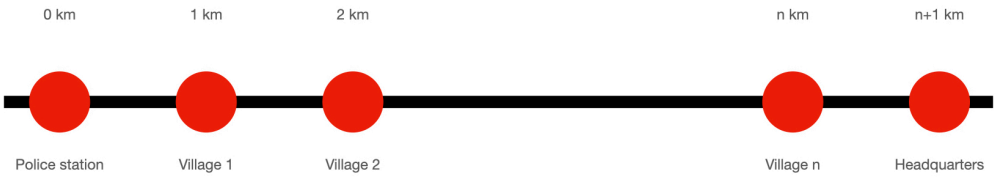
$$\mathbb{E}[X] = \sum_{i=1}^{t} \frac{\lfloor \frac{N}{p_i} \rfloor}{N}$$

2.4. Recall the Coupon Collector Problem discussed in class. Show that for $c > 0$

$$\Pr[X > n \log n + cn] \leq e^{-c}$$

where X is a random variable corresponding to the number of coupons collected, n is the total number of distinct coupons.

2.5. (♦) There is a police officer who needs to visit n villages for inspection. Fortunately, all the n villages are located on the highway, and the i^{th} village is located exactly i km from the police station. Finally, after visiting all the villages, the officer needs to submit these reports at the headquarters, also located on the same highway, at distance $(n+1)$ km from the police station.



The inspector picks a uniformly random permutation σ over $[n]$. Starting at the police station, he first visits the village $\sigma(1)$, then visits village $\sigma(2)$, and so on. Finally, after visiting village $\sigma(n)$, the officer goes to the headquarters.

Let X denote the distance travelled by the police officer. Compute $\mathbb{E}[X]$.

Solution:
Hint:
Use indicator random variables for the distance between each pair of consecutive villages and apply linearity of expectation.

Answer:

$$\mathbb{E}[X] = \frac{(n+1)(n+2)}{3}$$

2.6. (♦) **Ramsey Numbers:** In an graph $G = (V, E)$, a **clique** is a subset of vertices such that for every pair of vertices in that set, there exists an edge between them. Further, an **independent set** is a set of vertices such that no two vertices have an edge between them. The Ramsey number $R(k, l)$ is defined as the minimum number of vertices such that any graph with at least $R(k, l)$ vertices either has a clique of size k or an independent set of size l. Prove that for $k \geq 3$ $R(k, k) > 2^{k/2-1}$.

Hint: Consider a random graph on n vertices, where each edge is chosen w.p. $1/2$.

Solution: For the sake of explainability, let the edges chosen be indicated as being the edges colored red and the edges not chosen be indicated as being colored blue ($\binom{n}{2}$ - # red edges).
So the problem translates to: $R(k, k)$ is defined as the smallest number of vertices which has either a red clique of size k or a blue colored clique of size k.
Let us now color each edge with red color with $\frac{1}{2}$ probability and blue with $\frac{1}{2}$ probability.
Now let us calculate the expected number of red cliques of size k in the graph. There are $\binom{n}{k}$ cliques of size k in the graph, number them $i = \{1 \dots \binom{n}{k}\}$. Let X_i be an indicator variable indicating whether all edges of clique i are red.

$$E[\text{red cliques}] = E[X_1 + X_2 + X_3 \dots] = \binom{n}{k} E[X_1] = \frac{\binom{n}{k}}{2^{\binom{k}{2}}}.$$

Similarly,

$$E[\text{blue cliques}] = E[X_1 + X_2 + X_3 \dots] = \binom{n}{k} E[X_1] = \frac{\binom{n}{k}}{2^{\binom{k}{2}}}.$$

Thus,

$$E[\text{monochromatic cliques}] = E[\text{red cliques}] + E[\text{blue cliques}] = 2 \cdot \frac{\binom{n}{k}}{2^{\binom{k}{2}}}.$$

Claim: If for $n = n'$, $E[\text{monochromatic cliques}] < 1$, then $R(k,k) > n'$.

Proof: $E[\text{monochromatic cliques}] < 1$ implies that there exists a graph of size n' which has 0 monochromatic cliques, hence by definition, $R(k,k)$ (the smallest number of vertices that guarantees the existence of a monochromatic clique) must be greater than n'.

Claim: $2 \cdot \frac{\binom{n}{k}}{2^{\binom{k}{2}}}$ for $n = 2^{k/2-1}$ is less than 1.

Side note 1: Note the subtle point that $E[\text{monochromatic cliques}] > 1$ does not ensure the existence of a clique, hence this is just a lower bound for Ramsey's number.

Fun PHP problem: Prove that $R(3,3) = 6$.

2.7. **Hypergraph Coloring:** A $k-$uniform hypergraph is a a pair (X, S) where X is the set of vertices and $S = \binom{X}{k}$ is the set of hyperedges (think of them as generalization of graphs, where instead of having an edge (u, v) between any two vertices, we have an hyperedge between $k-$vertices). A hypergraph is c-colorable if its vertices can be colored with $c-$colors so that no hyperedge is monocromatic. Let $m(k)$ denote the smallest number of hyperedges in a k-uniform hypergraph that is not 2-colorable. Prove that for any $k \geq 2$, $m(k) \geq 2^{k-1}$.

Solution: Consider an arbitrary k-uniform hypergraph having $m(k)$ hyperedges. Let the graph be assigned a random coloring, where every vertex is colored red or blue independently and with probability 1/2.
Then for a hyperedge e, we have

$$\mathbf{P}[e \text{ is monochoromatic}] = 1/2^k + 1/2^k = 1/2^{k-1}$$

First $1/2^k$ is for 'all vertex is red' and second $1/2^k$ is for 'all vertex is blue'
By the Union Bound, we get,

$$\mathbf{P}[\exists e, e \text{ is monochromatic }] \leq \sum_e \mathbf{P}[e \text{ is monochromatic}] = m(k)/2^{k-1}$$

with respect to the random coloring

Hence,

$$\begin{aligned}\mathbf{P}[\text{Graph is not 2 colorable}] &= \mathbf{P}[\exists e,\ \text{e is monochromatic}] \\ &\leq m(k)/2^{k-1} < 1 \quad \text{for } m(k) < 2^{k-1}\end{aligned}$$

This means that there is an assignment of colors which makes the graph 2-colorable which makes the graph 2-colorable if $m(k) < 2^{k-1}$
Hence, $m(k) \geq 2^{k-1}$

This is version 2.0 of the tutorial sheet. We added one problem to the tutorial sheet, and updated the set of ♦ marked problems. Let me know if something is unclear. In case of any doubt or for help regarding writing proofs, feel free to contact me or TAs.

Venkata Koppula - kvenkata@iitd.ac.in
Ananya Mathur - cs5200416@iitd.ac.in
Anish Banerjee - cs1210134@cse.iitd.ac.in
Eshan Jain - cs5200424@cse.iitd.ac.in
Mihir Kaskhedikar - cs1210551@iitd.ac.in
Naman Nirwan - Naman.Nirwan.cs521@cse.iitd.ac.in
Pravar Kataria - Pravar.Kataria.cs121@cse.iitd.ac.in
Shashwat Agrawal - csz248012@cse.iitd.ac.in
Subhankar Jana - csz248009@cse.iitd.ac.in