COL202: Discrete Math	Tutorial 7
Instructor: Venkata Koppula	Release date: 10th October, 2024

The solutions for the ($\star$) marked problems must be submitted on Gradescope by 11:59 am on 18th October, 2024.

This tutorial sheet requires basic probability, conditional probability and random variables. Here is a summary of the definitions, and theorems/results that we have seen in class (Lectures 16 and 17).

- A probability distribution is defined using a sample space Ω (which is finite or countably infinite) and a function $p : \Omega \to \mathbb{R}^{\geq 0}$ such that $\sum_{x \in \Omega} p(x) = 1$. The probability of any event $\mathcal{E} \subseteq \Omega$ is $\sum_{x \in \mathcal{E}} p(x)$.

- (Union Bound) For any k events $A_1, A_2, \ldots, A_k$,

$$\Pr\left[\bigcup_{i=1}^{k} A_i\right] \leq \sum_{i=1}^{k} \Pr[A_i].$$

- (Birthday bound) Suppose we sample t numbers from $\{1, 2, \ldots, n\}$, independently and uniformly at random. Let p_{coll} denote the probability that at least two of the sampled elements are equal. Then

$$1 - e^{-t(t-1)/2n} \leq p_{\mathsf{coll}} \leq \frac{t(t-1)}{2n}$$

- Given two events A, B such that $\Pr[B] > 0$, we define the conditional probability

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}.$$

- (Law of Total Probability) Let (Ω, p) denote a probability distribution over sample space Ω. Let $(\Omega_1, \ldots, \Omega_k)$ be any partitioning of Ω. Then for any event A,

$$\Pr[A] = \sum_{i=1}^{k} \Pr[A \cap \Omega_i] = \sum_{i=1}^{k} \Pr[A \mid \Omega_i] \cdot \Pr[\Omega_i].$$

- (Random variables and Expectation) Let (Ω, p) be a probability distribution, and $X : \Omega \to \mathbb{R}$ be any random variable. Let $Z = \{X(w) : w \in \Omega\}$, and for any $z \in Z$, let $X^{-1}(z) = \{w \in \Omega \ : \ X(w) = z\}$. The expectation of this random variable is

$$\mathbb{E}[X] = \sum_{w \in \Omega} X(w) \cdot p(w) = \sum_{z \in Z} w \cdot \Pr\left[X^{-1}(z)\right].$$

- (Linearity of Expectation) For any random variables $X_1, X_2, \ldots, X_k$ over the same probability distribution (Ω, p),

$$\mathbb{E}[X_1 + X_2 + \ldots + X_k] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \ldots + \mathbb{E}[X_k].$$

In Lecture 18, we introduced the *probabilistic method*, which can be used to prove the existence of combinatorial objects with certain property $\mathcal{P}$. The high-level idea is the following: define an appropriate probability distribution, and show that if the combinatorial object is sampled from this distribution, then it will have the property $\mathcal{P}$ with positive probability. We saw two examples for this:

- for any 3CNF formula ϕ over n variables with m clauses, there exists an assignment to the variables such that at least $7m/8$ clauses are satisfied. Consider a uniformly random assignment, and let X denote the number of clauses satisfied by this random assignment. $\mathbb{E}[X] = 7m/8$, and therefore there exists an assignment that satisfies at least $7m/8$ clauses.

- graphs with large bipartite subgraphs: Let $G = (V, E)$ be an undirected graph over $|V| = n$ vertices with $|E| = m$ edges. There exists a subgraph $H = (V, F)$ such that $|F| \geq m/2$ and H is bipartite. To prove this, consider a uniformly random subset W of V. Let H be the subgraph of G with edges with one endpoint in W and another in $V \setminus W$, and let X denote the number of edges in H. Since $\mathbb{E}[X] = m/2$, there exists a subgraph H such that H has $m/2$ edges, and H is bipartite.

In Lecture 19, we studied some properties of uniformly random permutations over n elements.

- A uniformly random permutation has a cycle of length $\lceil n/2 \rceil$ with probability $\approx \ln 2$.

- Let σ be a uniformly random permutation, and X denote the number of points i such that $\sigma(i) = i$. Then $\mathbb{E}[X] = 1$.

- Let σ be a uniformly random permutation. For any $k \in [n]$, $\Pr[\text{1 is in a cycle of length k}] = 1/n$.

- Let σ be a uniformly random permutation, and X denote the number of cycles in σ. Then $\mathbb{E}[X] \approx \ln n$.

We concluded the lecture with the Coupon Collector Problem. In this problem, there are n distinct coupons, and the objective is to collect all n coupons. Every time you sample, you receive one of the n coupons, uniformly at random. We showed that the expected time to sample all n coupons is approximately $n \ln n$. The coupon collector problem is a useful abstraction that has several applications in the analysis of randomized processes/algorithms.

1 Tutorial Submission Problem ($\star$)

The following problem has a few parts. You only need to submit the ones that are ($\star$) 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

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

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

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

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).

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 $\frac{2\delta}{2+\delta} \le \ln(1+\delta)$. (You don't need to submit these for the tutorial, though you are encouraged to derive them on your own)

$$\Pr[X \ge (1+\delta)\mu] \le e^{-\delta^2\mu/(2+\delta)}$$

$$\Pr[X \le (1-\delta)\mu] \le e^{-\delta^2\mu/2}$$

$$\Pr[|X - \mu| \ge \delta\mu] \le 2e^{-\delta^2\mu/3}$$

2 Problems - General Probability

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

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 $\mathbb{E}[X] = \Theta\left(\frac{\log n}{\log\log n}\right)$.

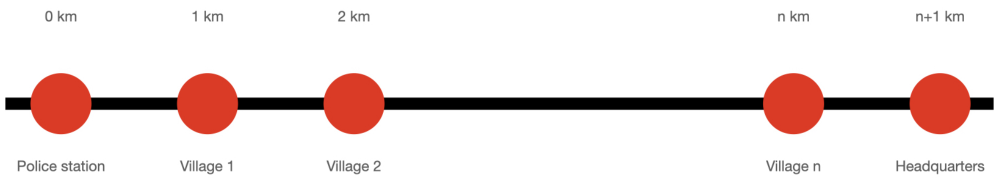
2.3. (♦) Let $N \in \mathbb{N}$, and let $2 = p_1 < p_2 < \ldots < p_t \le N$ be the set of prime numbers in $[N]$. Suppose n is sampled uniformly at random from $[N]$. Let X denote the number of *distinct* prime factors of n. Compute $\mathbb{E}[X]$ (in terms of $p_1, p_2, \ldots, p_t$).

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

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

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 $\mathbb{E}[X]$.

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 $k \ge 3$ $R(k,k) > 2^{k/2-1}$.

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

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 $k \ge 2$, $m(k) \ge 2^{k-1}$.

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@iitd.ac.in