Homework 2

CSCI 4030 / CSCI 6220: Randomized Algorithms — Fall 2026

Rensselaer Polytechnic Institute — 100 points

Justify every claim. A numerical answer without an argument earns no credit. Acknowledge any collaborators and any external sources you consult, including AI tools, and describe clearly what role they played in your submission.

Problems 1–5 are required of everyone and total 100 points. Problem 6 is required for CSCI 6220 and is worth 20 points of extra credit for CSCI 4030.

Throughout, you may use without proof the following results from lecture. For a random variable X with mean μ and variance σ^2 and any $t > 0$, Chebyshev's inequality gives $\mathbb{P}[|X - \mu| \geq t] \leq \sigma^2/t^2$. The Laplace–Chernoff inequality gives, for any random variable X, any $x \in \mathbb{R}$, and any $\theta > 0$,

$$\mathbb{P}[X \geq x] \leq e^{-\theta x}\mathbb{E}[e^{\theta X}].$$

For $S = \sum_{i=1}^{n} X_i$ a sum of *independent* Bernoulli variables with $\mu = \mathbb{E}[S]$, the multiplicative upper- and lower-tail Chernoff bounds give, for $0 \leq \delta \leq 1$,

$$\mathbb{P}[S \geq (1+\delta)\mu] \leq e^{-\delta^2\mu/3}, \qquad \mathbb{P}[S \leq (1-\delta)\mu] \leq e^{-\delta^2\mu/2}.$$

For $S = \sum_{i=1}^{n} X_i$ a sum of independent variables with $X_i \in [a_i, b_i]$ and $R^2 = \sum_i (b_i - a_i)^2$, Hoeffding's inequality gives

$$\mathbb{P}[|S - \mathbb{E}[S]| \geq x] \leq 2e^{-2x^2/R^2}.$$

[10 points] Problem 1 (The moment generating function of a Poisson random variable).

A discrete Poisson random variable X with parameter $\mu > 0$ takes values on the nonnegative integers with

$$\mathbb{P}[X = j] = \frac{e^{-\mu}\mu^j}{j!}, \qquad j = 0, 1, 2, \ldots$$

(a) (4 points) Compute the moment generating function $M_X(t) = \mathbb{E}[e^{tX}]$. State for which real t your expression is valid.

(b) (3 points) Using M_X, compute $\mathbb{E}[X]$ and $\mathrm{Var}[X]$. State the general facts about derivatives of M_X that you are using. Then compute $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$ directly from the definition of the Poisson distribution, and verify that the values agree with those obtained from the MGF.

(c) (3 points) Look back at the proof of the multiplicative Chernoff upper-tail bound for a sum of independent Bernoulli random variables. Explain why the same upper-tail bound also holds for a Poisson random variable with mean μ.

[20 points] Problem 2 (Pairwise independence is enough, until it isn't).

Parts (a)–(e) adapted from Mitzenmacher & Upfal, Exercise 3.23.

Flip a fair coin $n \geq 3$ times to obtain independent uniform random bits $b_1, \ldots, b_n \in \{0,1\}$. Fix an enumeration of all $m = \binom{n}{2}$ unordered pairs of indices. For the i-th pair $\{j,k\}$, let

$$Y_i = b_j \oplus b_k$$

be the exclusive-or of the two bits, and set $Y = \sum_{i=1}^{m} Y_i$.

(a) (2 points) Show that $\mathbb{P}[Y_i = 0] = \mathbb{P}[Y_i = 1] = 1/2$ for every i, and compute $\mathbb{E}[Y]$.

(b) (3 points) Show that $Y_1, \ldots, Y_m$ are *not* mutually independent. Give an explicit triple of indices witnessing the failure.

(c) (4 points) Show that $\mathbb{E}[Y_i Y_j] = \mathbb{E}[Y_i]\,\mathbb{E}[Y_j]$ for all $i \neq j$. Treat both cases: the two pairs share a bit, and the two pairs are disjoint.

(d) (4 points) Using (c) and the variance-of-a-sum identity, compute $\mathrm{Var}[Y]$ exactly. State precisely which hypothesis of the identity you are invoking, and why mutual independence is not needed.

(e) (3 points) Use Chebyshev's inequality to bound $\mathbb{P}\big[|Y - \mathbb{E}[Y]| \geq n\big]$. Comment briefly on the strength of the bound: does it give a constant-probability guarantee or a high-probability guarantee as n grows?

(f) (4 points) A student proposes to replace part (e) by the Chernoff bound, arguing that Y is a sum of m Bernoulli(1/2) variables. In the standard Chernoff derivation one uses

$$\mathbb{E}\Big[e^{\theta \sum_i Y_i}\Big] \stackrel{?}{=} \prod_i \mathbb{E}[e^{\theta Y_i}].$$

Explain exactly why this equality is not justified here. Why does the calculation in part (c), which only concerns products of two distinct variables at a time, not justify this factorization?

[15 points] Problem 3 (A geometric tail bound in two ways).

Adapted from Mitzenmacher & Upfal, Exercise 4.12.

Let $X_1, \ldots, X_n$ be independent geometric random variables with mean 2, so that

$$\mathbb{P}[X_i = k] = 2^{-k}, \qquad k = 1, 2, \ldots,$$

and let $X = \sum_{i=1}^{n} X_i$. Thus X is the number of independent fair-coin tosses required to see n heads. Fix $\delta > 0$ and set

$$L = \lceil 2n(1+\delta) \rceil + 1.$$

(a) (6 points) Show that the event $\{X \geq L\}$ is exactly the event that among the first $L-1$ independent fair-coin tosses there are at most $n-1$ heads. Use the multiplicative lower-tail Chernoff bound stated above to obtain an explicit upper bound on $\mathbb{P}[X \geq L]$.

(b) (6 points) Derive an upper bound on $\mathbb{P}[X \geq L]$ a second time using the moment generating function of a geometric random variable. First compute $M_{X_i}(\theta)$ and determine exactly for which θ it is finite. Then apply Laplace–Chernoff, factorize the MGF of X, and optimize or make an explicit valid choice of θ. Explain how the domain of the MGF restricts your choice.

(c) (3 points) Compare the two arguments. They analyze the same random variable but expose different structure. Which route gives the cleaner bound for this problem, and where does independence enter in each argument?

[25 points] Problem 4 (The median-of-means estimator).

Adapted from Mitzenmacher & Upfal, Exercise 4.9.

We can draw independent samples $X_1, X_2, \ldots$ of a real random variable X with $\mathbb{E}[X] > 0$ and finite variance, and we wish to estimate $\mathbb{E}[X]$. Since X need not be bounded and need not be a 0–1 variable, neither Chernoff nor Hoeffding applies to the samples directly; and we do not assume that the moment generating function of X is finite. All we will assume is knowledge of the ratio

$$r = \frac{\sqrt{\mathrm{Var}[X]}}{\mathbb{E}[X]},$$

the relative standard deviation of X. Fix an accuracy $\varepsilon \in (0, 1)$ and a failure probability $\delta \in (0, 1)$. We want an estimate $\hat{\mu}$ with

$$\mathbb{P}\big[|\hat{\mu} - \mathbb{E}[X]| \le \varepsilon\, \mathbb{E}[X]\big] \ge 1 - \delta.$$

(a) (5 points) Let $\hat{\mu}_t = \frac{1}{t}\sum_{i=1}^{t} X_i$. Using Chebyshev's inequality, show that

$$t \ge \frac{r^2}{\varepsilon^2 \delta}$$

suffices for the desired guarantee.

(b) (4 points) Call $\hat{\mu}_t$ a *weak estimate* if it satisfies the accuracy requirement with probability at least $3/4$ rather than $1 - \delta$. Show that $t = 4r^2/\varepsilon^2$ samples suffice for a weak estimate.

(c) (10 points) Choose an odd integer k and compute k independent weak estimates $Z_1, \ldots, Z_k$, each from a fresh block of samples. Output their median $\hat{\mu} = \mathrm{med}(Z_1, \ldots, Z_k)$. For each block, let B_i be the indicator that Z_i is inaccurate, and let $B = \sum_{i=1}^{k} B_i$. Show that $\mathbb{E}[B] \le k/4$ and that the median can fail only if $B \ge k/2$. Then apply Hoeffding's inequality to the independent variables $B_1, \ldots, B_k$ to prove

$$\mathbb{P}[\hat{\mu} \text{ is inaccurate}] \le 2e^{-k/8}.$$

Conclude that choosing an odd integer

$$k \ge 8 \ln \frac{2}{\delta}$$

suffices to make the median accurate with probability at least $1 - \delta$.

(d) (3 points) Conclude that

$$O\left(\frac{r^2 \log(1/\delta)}{\varepsilon^2}\right)$$

samples suffice in total.

(e) (3 points) Compare the sample counts from (a) and (d) at $\varepsilon = 0.1$, $r = 1$, and $\delta = 10^{-6}$. Explain in one paragraph what structural feature of the median-of-means construction converts the $1/\delta$ dependence from Chebyshev into $\log(1/\delta)$, even though the underlying samples themselves have not become better behaved.

[30 points] Problem 5 (Randomized Quicksort is $O(n \log n)$ with high probability).

Adapted from Mitzenmacher & Upfal, Exercise 4.21.

Randomized Quicksort is run on n distinct elements. At each recursive call, the pivot is chosen uniformly at random from the current subarray, independently of all previous pivot choices. A partition step on a subarray of size s is called *good* if both resulting subarrays have size at most $2s/3$; otherwise it is *bad*.

For a fixed input element x, follow the recursive calls whose subarrays contain x, stopping when x is chosen as a pivot or reaches a singleton. This is the *recursion path of* x. A good partition step on this path shrinks the subarray containing x by a constant factor. The idea of the problem is to show that sufficiently many such good steps occur within $O(\log n)$ partition steps.

You may use the following extension of the lower-tail Chernoff bound. Consider a sequence of trials that may stop early. Suppose that whenever the i-th trial occurs, its probability of success, conditioned on everything that has happened previously, is at least p. Then for any k and any $0 \le \delta \le 1$,

$$\mathbb{P}\left[\begin{array}{c}\text{at least } k \text{ trials occur, but fewer than} \\ (1-\delta)pk \text{ of the first } k \text{ trials succeed}\end{array}\right] \le \exp\left(-\frac{\delta^2 pk}{2}\right).$$

Thus independence of the trials is not required here; it is enough that each new trial has success probability at least p, no matter what happened before.

(a) (5 points) Show that, no matter what happened at earlier recursive calls, the next partition step is good with probability at least $1/3$. Explain why a good step reduces the size of the subarray on the recursion path of x by a factor of at most $2/3$.

(b) (5 points) Show that after

$$L = \left\lceil \log_{3/2} n \right\rceil$$

good partition steps, the recursion path of x must have ended.

(c) (10 points) Let

$$k = \lceil 18 \log_2 n \rceil .$$

If the recursion path of x contains at least k partition steps, explain why fewer than L of the first k steps can be good. Show that, for all sufficiently large n,

$$L - 1 \le \frac{k}{10}.$$

Apply the extension of the lower-tail Chernoff bound stated above with $p = 1/3$ to show that

$$\mathbb{P}[\text{the recursion path of a fixed } x \text{ has at least } k \text{ partition steps}] \le \frac{1}{n^2}.$$

(d) (5 points) Apply a union bound over the n input elements to show that, with probability at least $1 - 1/n$, every recursion path has length $O(\log n)$.

(e) (5 points) Conclude that Randomized Quicksort performs $O(n \log n)$ comparisons with probability at least $1 - 1/n$. Be explicit: explain why, at any fixed recursion depth, the active subarrays are disjoint and therefore the total number of comparisons performed by all partition operations at that depth is $O(n)$.

[20 points] Problem 6 ([CSCI 6220 required; CSCI 4030 extra credit] A Chernoff bound without independence).

Adapted from Mitzenmacher & Upfal.

In Lecture 4 we observed that independence entered the Chernoff derivation at exactly one point: the factorization $\mathbb{E}[e^{\theta \sum_i X_i}] = \prod_i \mathbb{E}[e^{\theta X_i}]$. We also claimed, without proof, that this factorization is more than we actually need — an *inequality* in the right direction suffices — and that certain dependent families satisfy it even though they are far from independent. Problem 2 exhibited a dependent family for which the Chernoff conclusion genuinely fails. This problem exhibits one for which it survives.

Throw n balls independently and uniformly at random into n bins. Let $X_i = 1$ if the i-th bin is empty and $X_i = 0$ otherwise, and let $X = \sum_{i=1}^{n} X_i$ be the number of empty bins. The variables $X_1, \ldots, X_n$ are *not* independent: if many bins are empty, the remaining bins are more crowded.

As a comparison family, let $Y_1, \ldots, Y_n$ be *independent* Bernoulli random variables, each equal to 1 with probability $p = (1 - 1/n)^n$, and let $Y = \sum_{i=1}^{n} Y_i$.

(a) (2 points) Show that $\mathbb{E}[X_i] = p$ for every i, so that $\mathbb{E}[X] = \mathbb{E}[Y] = np$. Then show that X_1 and X_2 are negatively correlated, i.e. $\mathrm{Cov}(X_1, X_2) < 0$.

(b) (6 points) Show that $\mathbb{E}[X_1 X_2 \cdots X_k] \le \mathbb{E}[Y_1 Y_2 \cdots Y_k]$ for every $k \ge 1$. Interpret the left-hand side as the probability of an explicit event before computing it.

(c) (6 points) Show that $\mathbb{E}[e^{tX}] \le \mathbb{E}[e^{tY}]$ for all $t \ge 0$. (Suggestion: expand e^{tX} as a power series and compare $\mathbb{E}[X^j]$ with $\mathbb{E}[Y^j]$ term by term. When you expand X^j, every monomial reduces to a product over a set of distinct indices because $X_i^a = X_i$ for every integer $a \ge 1$. Then use part (b).)

(d) (4 points) For $0 \le \delta \le 1$, show that the multiplicative upper-tail Chernoff bound stated above also holds for X:

$$\mathbb{P}[X \ge (1+\delta)\mathbb{E}[X]] \le \exp\left(-\frac{\delta^2 \mathbb{E}[X]}{3}\right).$$

Identify the single step in the usual proof for independent Bernoulli variables at which part (c) is used instead.

(e) (2 points) Problem 2 and this problem both concern sums of dependent indicator variables, and the Chernoff conclusion fails in one case and holds in the other. Explain in a short paragraph what distinguishes them. Your answer should make clear why "the variables are correlated" is not by itself the relevant criterion.