CS 70 — Discrete Mathematics and Probability Theory
Fall 2020
HW 12

Due: Friday 11/20 2020 at 10:00 PM
Grace period until Friday 11/20 2020 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Graph

Consider a random graph (undirected, no multi-edges, no self-loops) on n nodes, where each possible edge exists independently with probability p. Let X be the number of isolated nodes (nodes with degree 0).

(a) What is $\text{Var}(X)$?

2 Whitening

Let X and Y be two random variables, with $\text{Var}(X) > 0, \text{Var}(Y) > 0$. Show that it is possible to construct $\tilde{X} = aX + bY$ and $\tilde{Y} = cX + dY$, where a, b, c, d are scalars to be chosen subject to the constraint $ad - bc \neq 0$, such that $\text{cov}(\tilde{X}, \tilde{Y}) = 0$.

You may find it unnecessary to transform Y, that is, you only need to solve for a, b to get $\text{cov}(\tilde{X}, Y) = 0$.

3 Probabilistic Bounds

A random variable X has variance $\text{Var}(X) = 9$ and expectation $\mathbb{E}[X] = 2$. Furthermore, the value of X is never greater than 10. Given this information, provide either a proof or a counterexample for the following statements.

(a) $\mathbb{E}[X^2] = 13$.

(b) $\mathbb{P}[X = 2] > 0$.

(c) $\mathbb{P}[X \geq 2] = \mathbb{P}[X \leq 2]$.

(d) $\mathbb{P}[X \le 1] \le 8/9$.

(e) $\mathbb{P}[X \ge 6] \le 9/16$.

4 Subset Card Game

Jonathan and Yiming are playing a card game. Jonathan has $k > 2$ cards, and each card has a real number written on it. Jonathan tells Yiming (truthfully), that the sum of the card values is 0, and that the sum of squares of the values on the cards is 1. Specifically, if the card values are $c_1, c_2, \ldots, c_k$, then we have $\sum_{i=1}^{k} c_i = 0$ and $\sum_{i=1}^{k} c_i^2 = 1$. Jonathan and Yiming also agree on a positive target value of α.

The cards are then going to be dealt randomly in the following fashion: for each card in the deck, a fair coin is flipped. If the coin lands heads, then the card goes to Yiming, and if the coin lands tails, the card goes to Jonathan. Note that it is possible for either player to end up with no cards/all the cards.

A player wins the game if the sum of the card values in their hand is at least α, otherwise it is a tie.

(a) Prove that the probability that Yiming wins is at most $\frac{1}{8\alpha^2}$.

(b) Find a deck of k cards and target value α where the probability that Yiming wins is exactly $\frac{1}{8\alpha^2}$.

5 Just One Tail, Please

Let X be some random variable with finite mean and variance which is not necessarily non-negative. The *extended* version of Markov's Inequality states that for a non-negative function $\phi(x)$ which is monotonically increasing for $x > 0$ and some constant $\alpha > 0$,

$$\mathbb{P}(X \ge \alpha) \le \frac{\mathbb{E}[\phi(X)]}{\phi(\alpha)}$$

Suppose $\mathbb{E}[X] = 0$, $\text{Var}(X) = \sigma^2 < \infty$, and $\alpha > 0$.

(a) Use the extended version of Markov's Inequality stated above with $\phi(x) = (x+c)^2$, where c is some positive constant, to show that:

$$\mathbb{P}(X \ge \alpha) \le \frac{\sigma^2 + c^2}{(\alpha + c)^2}$$

(b) Note that the above bound applies for all positive c, so we can choose a value of c to minimize the expression, yielding the best possible bound. Find the value for c which will minimize the RHS expression (you may assume that the expression has a unique minimum). Plug in the minimizing value of c to prove the following bound:

$$\mathbb{P}(X \ge \alpha) \le \frac{\sigma^2}{\alpha^2 + \sigma^2}.$$

(c) Recall that Chebyshev's inequality provides a two-sided bound. That is, it provides a bound on $\mathbb{P}(|X - \mathbb{E}[X]| \geq \alpha) = \mathbb{P}(X \geq \mathbb{E}[X] + \alpha) + \mathbb{P}(X \leq \mathbb{E}[X] - \alpha)$. If we only wanted to bound the probability of one of the tails, e.g. if we wanted to bound $\mathbb{P}(X \geq \mathbb{E}[X] + \alpha)$, it is tempting to just divide the bound we get from Chebyshev's by two. Why is this not always correct in general? Provide an example of a random variable X (does not have to be zero-mean) and a constant α such that using this method (dividing by two to bound one tail) is not correct, that is, $\mathbb{P}(X \geq \mathbb{E}[X] + \alpha) > \frac{\text{Var}(X)}{2\alpha^2}$ or $\mathbb{P}(X \leq \mathbb{E}[X] - \alpha) > \frac{\text{Var}(X)}{2\alpha^2}$.

Now we see the use of the bound proven in part (b) - it allows us to bound just one tail while still taking variance into account, and does not require us to assume any property of the random variable. Note that the bound is also always guaranteed to be less than 1 (and therefore at least somewhat useful), unlike Markov's and Chebyshev's inequality!

(d) Let's try out our new bound on a simple example. Suppose X is a positively-valued random variable with $\mathbb{E}[X] = 3$ and $\text{Var}(X) = 2$. What bound would Markov's inequality give for $\mathbb{P}[X \geq 5]$? What bound would Chebyshev's inequality give for $\mathbb{P}[X \geq 5]$? What about for the bound we proved in part (b)? (*Note*: Recall that the bound from part (b) only applies for zero-mean random variables.)

6 Sum of Poisson Variables

Assume that you were given two independent Poisson random variables X_1, X_2. Assume that the first has mean λ_1 and the second has mean λ_2. Prove that $X_1 + X_2$ is a Poisson random variable with mean $\lambda_1 + \lambda_2$.

Hint: Recall the binomial theorem.

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$