CS–174 Combinatorics and Discrete Probability, Spring 2000

Solution to Midterm Exam 1

Solution to 1: Suppose two fair sided dice are tossed. Let X be the number on one of the dice, and let Y be the sum of the two dice.

a. (4 pts) What is $\mathbf{E}[X+Y]$?
Let $Y = X + Z$ where Z is is the number on the second dice. Then X and Z are identically distributed independent random variables. $\mathbf{E}[X] = \mathbf{E}[Z] = 7/2$, therefore $\mathbf{E}[X+Y] = 3\mathbf{E}[X] = 21/2$.

b. (7 pts) What is $\mathbf{Var}[X]$?
$\mathbf{Var}[X] = \mathbf{E}[X^2] \Leftrightarrow \mathbf{E}[X]^2 = \frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2) \Leftrightarrow (21/2)^2 = 35/12$.

c. (9 pts) What is $Cov(X,Y)$?

$$
\begin{aligned}
Cov(X,Y) &= \mathbf{E}[XY] \Leftrightarrow \mathbf{E}[X]\mathbf{E}[Y] = \mathbf{E}[X(X+Z)] \Leftrightarrow \mathbf{E}[X]\mathbf{E}[X+Z] \\
&= \mathbf{E}\left[X^2\right] + \mathbf{E}[XZ] \Leftrightarrow \mathbf{E}\left[X^2\right] \Leftrightarrow \mathbf{E}[X]\mathbf{E}[Z] \\
&= \mathbf{Var}[X] + Cov(X,Z)
\end{aligned}
$$

But since X and Z are independent, $Cov(X,Y) = 0$ so we get $Cov(X,Y) = \mathbf{Var}[X] = 35/12$.

Solution to 2: Suppose we begin with a graph G with n vertices and no edges, and then add edges i,j one at a time by choosing i,j independently and uniformly at random from $\{1,\ldots,n\}$.

a. (5 pts) What is the expected number of edges m added before some vertex is hit twice, i.e. before some vertex has two edges touching it?
This is just like throwing balls into n bins where each edge is equivalent to throwing two balls. The birthday paradox tells us that the expected number of balls before some bin contains two balls is $\sqrt{2n}$. To get the expected number of edges, we must divide by 2:

$$\mathbf{E}[m] = \sqrt{n/2}$$

b. (5 pts) What is the expected number of edges m before every vertex is hit at least once?
This is equivalent to coupon collecting where each edge corresponds to acquiring two independent coupons. We know that the expected number of coupons we must acquire before having all n of them is asymptotically $n \ln n$, so the expected number of edges must be half of that: $\frac{n \ln n}{2}$.

c. (10 pts) Suppose the graph G has m edges, n vertices, and exactly two components. Let p be the probability that the next random edge connects the two components. Give upper and lower bounds for p.
Suppose one component contains k vertices and the second one contains $n \Leftrightarrow k$ vertices. There are n^2 possibilities for the choice of the next edge (n choices for i and n choices for j). The number of choices which will connect the two components is $k(n \Leftrightarrow k) + (n \Leftrightarrow k)k = 2k(n \Leftrightarrow k)$. Thus, $p = 2k(n \Leftrightarrow k)/n^2$. The lower bound for p is when one component is just a single vertex and the upper bound is when both components contain exactly $n/2$ vertices:

$$2(n \Leftrightarrow 1)/n^2 \le p \le 1/2$$

Solution to 3: Give the name of the probability distributions for the random variable X, and compute the mean and variance in each case:

a. (6 pts) Suppose 5 cards are drawn from a deck of 52 cards, one-at-a-time, with replacement. Let X be the number of aces.

Since each card is chosen independently and uniformly from the whole deck, this is the binomial distribution with $n = 5$ and $p = 4/52 = 1/13$ since there are only 4 aces in a deck.

$\mathbf{E}[X] = np = 5/13$ and $\mathbf{Var}[X] = np(1 \Leftrightarrow p) = \frac{5}{13}\frac{12}{13} = 60/169$.

b. (8 pts) Suppose n people seat themselves randomly around a meeting table with n seats before lunch. Then after lunch, they sit down again in new random positions. Let X be the number of people who sit in the same seat. Assume n is large.

Since two people cannot sit in the same seat, the second seating is a random permutation of the first one, and X just counts the number of fixed points in that permutation. We saw that when n is large, this distribution is Poisson with parameter 1. Hence $\mathbf{E}[X] = \mathbf{Var}[X] = 1$.

c. (6 pts) We draw cards again from a deck of 52 cards, this time in turns. On each turn, we pick two cards without replacement. Then we replace them, reshuffle and repeat. Let X be the number of turns up to and including the first time we draw a pair.

Each turn is independent of the other ones and has the same probability of succeeding, so this is just the geometric distribution. The probability that we pull a pair is $3/51 = 1/17$ since whatever the first card is, there are only 3 other cards in the deck with the same value and only 51 cards left in the deck. This means $\mathbf{E}[X] = 1/p = 17$ and $\mathbf{Var}[X] = (1 \Leftrightarrow p)/p^2 = (16/17)/(1/17)^2 = 272$.

Solution to 4: (20 pts) Suppose we have a random variable X such that $\mathbf{E}[X] = 2n$ and $\mathbf{Var}[X] = 4n$. Find a function $f(n)$ such that

$$\mathbf{Pr}(X > \mathbf{E}[X] + f(n)) < \frac{1}{\sqrt{n}}$$

Without more information we cannot use Chernoff bounds, but we can use Chebyshev bounds since we know the variance:

$$\mathbf{Pr}(X > \mathbf{E}[X] + f(n)) \leq \mathbf{Pr}(|X \Leftrightarrow \mathbf{E}[X]| > f(n)) < \mathbf{Var}[X] / f(n)^2$$

This means we need to have:

$$\frac{\mathbf{Var}[X]}{f(n)^2} < \frac{1}{\sqrt{n}} \Leftrightarrow f(n) > 2n^{\frac{3}{4}}$$