CS 174: Combinatorics and Discrete Probability (Spring 2022), UC Berkeley

Practice Midterm Problems

Q1-5 are Midterm problems from Spring 2004, while Q6 is a practice Midterm problem from Spring 2007.

1. A random variable X takes non-negative values and has $\mathbb{E}[X] = 3$.

   (a) Give an upper bound on $\mathbb{P}(X \geq 6)$, and specify a random variable X for which this upper bound is equal to the probability.

   (b) Suppose that, in addition, you know that $\mathrm{Var}[X] = 4$. Give an upper bound on $\mathbb{P}(|X - \mathbb{E}[X]| \geq 3)$, and specify a random variable X for which this upper bound is equal to the probability.

2. Suppose that we have a biased coin, and we wish to use it to generate random numbers from the set $\{1, 2, 3\}$. Consider the following algorithm:

   - Toss the coin three times.
   - If precisely one of the three tosses came up heads, output the number of the toss on which the head appeared. Suppose that the probability that the coin comes up heads is $0 < p < 1$.

   (a) Verify that, if the algorithm outputs a number, it is chosen uniformly from $\{1, 2, 3\}$.

   (b) Suppose we repeatedly call the algorithm until it outputs a number. What is the expected number of coin tosses required? What is the variance of the number of coin tosses required?

3. Suppose that $k = 2^n$ sporting teams compete in a tournament, designed as follows. In the first round, the k teams are randomly grouped into $k/2$ pairs, numbered $1, ..., k/2$, and each pair plays a match. If the outcome is a loss for one team, that team is eliminated from the tournament. If the outcome is a draw, that pair plays an additional first round match, and so on, until one loses and is eliminated. In subsequent rounds, the winners from the previous round are paired and play matches in the same way. This continues until only one team remains. Suppose that the match outcomes are independent, the probability that a match ends in a draw is $1/3$, and all teams are equally matched.

   (a) What is the expected number of matches in the tournament?

   (b) What is the expected number of matches played by the winning team?

4. Consider the following online, randomized algorithm for estimating the proportion of ones in a bit sequence $(b_1, \ldots, b_n) \in \{0, 1\}^n$. At time t, the algorithm is called with inputs (t, X_{t-1}, b_t) and it outputs a single bit, $X_t \in \{0, 1\}$ (and we define $X_0 = 0$). The algorithm chooses its output as follows. If $t = 1$, it returns b_1. For $t > 1$, it tosses a biased coin: with probability $1/t$ it returns b_t, otherwise it returns X_{t-1}.

(a) Show that the output produced by the algorithm at each step is an unbiased estimate of the proportion of ones in the sequence so far, that is,

$$\mathbb{E}[X_t] = \frac{1}{t}\sum_{i=1}^{t} b_i.$$

(b) What is the variance of X_t?

(c) Suppose that, instead of outputting a single bit the algorithm outputs d bits. That is, at time t, the algorithm is called with inputs (t, X_{t-1}, b_t) and it outputs $X_t = (X_{t,1}, \ldots, x_{t,d}) \in \{0,1\}^d$ (and we define $X_0 = (0, \ldots, 0) \in \{0,1\}^d$). Describe a randomized algorithm for which the output produced by the algorithm at each step leads to an unbiased estimate of the proportion of ones in the sequence so far, in the sense that

$$\mathbb{E}[Y_t] = \frac{1}{t}\sum_{i=1}^{t} b_i,$$

where

$$Y_t = \frac{1}{d}\sum_{i=1}^{d} X_{t,i}.$$

What is the variance of Y_t? (Your new algorithm must be better than the old algorithm: its variance must be no larger, and for some sequences $b_1, \ldots, b_t$ its variance must be smaller.)

5. We throw m balls into 3 bins, uniformly and independently. Use a Chernoff bound on the tail of the binomial distribution to give an upper bound on the probability that the first bin contains more balls than the total number in the other two.

6. Consider the event that every bin receives exactly k balls when kn balls are thrown randomly into n bins.

(a) Determine the exact probability of this event.

(b) Compute the probability under the Poisson approximation.

(c) The value upon dividing the expression in (b) by that in (a) equals the probability that a Poisson random variable with parameter λ takes on some value r. Explain briefly but precisely why this quotient matches a Poisson distribution. Also, what are the values of λ and r?