CSE 312: Foundations of Computing II
Assignment #4
January 28, 2016
due: Friday, February 5, 2016, 1:30 p.m. (before lecture starts)

Instructions:

• **Answers:** When asked for a short answer (such as a single number), also *show and explain your work* briefly. If your answer is too large to evaluate easily, you can leave it in terms of factorials, combinations, etc., for instance 26^7 or 26!/7! or 26(25)!/

• **Harmonic numbers:** The *n*-th Harmonic number is defined as \( H_n = 1 + 1/2 + 1/3 + 1/4 + \cdots + 1/n \). As a function of \( n \), \( H_n \) grows very similarly to \( \ln n \). In fact, \( H_n = \ln n + \Theta(1) \). You may use this fact without proof.

• **Bundles:** This week’s turnin bundles: (A) problems 1–2, (B) problems 3–4, (C) problems 5–7. PRINT your full name in the upper left corner of each bundle’s top page, with your last name printed clearly in all CAPITAL LETTERS. Each bundle should be stapled separately. We don’t supply the stapler.

A1. For quality control during the manufacture of chips, a sample of 3 chips is selected randomly and uniformly, without replacement, from a batch containing 21 chips, and each of those 3 sample chips is tested. Suppose that 6 of the 21 chips in the batch are defective. Let random variable \( X \) be the number of chips in the sample that are defective. Calculate all answers exactly as simplified fractions.

  (a) Calculate the probability mass function for \( X \).
  (b) Find \( E\left[X\right] \) directly by applying the definition of expectation to the result from part (a).
  (c) Find \( E\left[X\right] \) again by using linearity of expectation.

Look back at your work and think about which way of computing the expectation was simpler, (a) + (b), or (c).

A2. Suppose two random subsets \( A \) and \( B \) of \( \{1, 2, \ldots, n\} \) are chosen independently, where every subset of \( \{1, 2, \ldots, n\} \) is equally likely. Let the random variable \( X \) be \( |A \cup B| \).

  (a) Calculate the probability mass function for \( X \).
  (b) Calculate \( E\left[X\right] \). (Hint: use linearity of expectation.)

B3. Let \( X \) be the number of heads seen when 3 coins are flipped independently. The first coin has probability \( p \) of heads, the second has probability \( q \) of heads, and the third has probability \( r \) of heads.

  (a) Calculate the probability mass function for \( X \) as a function of \( p, q, r \).
  (b) Find \( E\left[X\right] \) directly by applying the definition of expectation to the result from part (a). Simplify your answer as far as possible, which will be important for comparing your answer to (c) below.
  (c) Find \( E\left[X\right] \) again by using linearity of expectation.

Look back at your work and think about which way of computing the expectation was simpler, (a) + (b), or (c).

B4. An urn initially contains one red and one blue ball. At each stage, a ball is randomly chosen from the urn and then replaced along with another of the same color. For instance, if in the first stage a red ball is chosen, then at the beginning of the second stage there are 2 red balls and 1 blue ball in the urn. Let \( X \) be the first stage at which the ball chosen is blue. (For example, if the first two selections are red and the third blue, then \( X \) is 3.)
(a) Find the probability mass function for $X$.

(b) Find $P(X > i)$ as a function of $i$.

(c) What is the limiting value of the probability that a blue ball is not selected by the $n$-th stage, as $n$ tends to infinity?

(d) Show that $E[X]$ is unbounded. (Hint: see the instructions at the top of this homework.) Parts (c) and (d) form an interesting contrast.

C5. You have an array holding 5 numbers, and you run the following algorithm to find the maximum of those numbers:

```java
max = a[0];
for(i = 1; i < 5; i++){
    if(max < a[i]){ // (*)
        max = a[i]; // (**)
    }
}
```

To do a detailed analysis of the running time of this algorithm, you need to know how many times statement (**) is skipped. A related, perhaps simpler, question is this: assuming that the 5 numbers are all distinct and equally likely to be in any order (that is, they might as well be a random permutation of 1, 2, 3, 4, 5), let the random variable $X$ be the number of times the “if” test in (*) evaluates to false before the first time statement (**) is executed (if ever). For example, $X = 2$ means that the 0-th array element was at least as great as elements 1 and 2, but less than element 3, and $X = 4$ means that the 0-th element was at least as great as any of the others. Give your answers as simplified fractions.

(a) Find the probability mass function for $X$.

(b) Find $E[X]$.


C7. A certain bubble gum company includes a picture card of a famous computer scientist in each pack of bubble gum it sells. Needless to say, kids absolutely love collecting these exciting cards. A complete set of cards consists of $n$ different pictures. Suppose that every pack you buy is equally likely to contain any of the $n$ pictures.

(a) Alice buys a pack a day until the first time she gets a duplicate of a card she already has, and then gets bored and stops. Let random variable $X$ be the number of packs required to accomplish this objective. Find the probability mass function of $X$. (Hint: The probability mass function is $P(X = k)$ as a function of $k$. I find it helps me first to compute $P(X > k)$ as a function of $k$, then use this to compute the probability mass function.)

(b) Bob buys a pack a day until he has a complete set. Let random variable $Y$ be the number of packs required to accomplish this objective. What is $E[Y]$? (Hint: Define random variable $Y_i$ as the number of packs Bob buys from the time he first has $i-1$ different pictures until the first time he acquires one that he doesn’t have yet. Also, see the instructions at the top of this homework.)