COMPUTER SCIENCE E-20, SPRING 2015
Homework Problems
Expectation, Probability Review, Series

**Due Thursday, April 30, 2015 before 9PM EDT. Upload a PDF of your answers at https://canvas.harvard.edu/courses/1815/assignments/23127**

1. Recall that the expectation of a random natural number $X$ is the sum

$$\mathbb{E}[X] = \sum_{n=0}^{\infty} n\mathbb{P}[X = n]$$

(a) Show that $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.

**Solution:** Using the definition above, we have $\mathbb{E}[X+Y] = \sum_{n=0}^{\infty} n\mathbb{P}[X+Y = n]$. If $X + Y = n$, then there is some $k$ with $0 \leq k \leq n$ such that $X = k$ and $Y = n - k$. Thus

$P[X + Y = n] = \sum_{k=0}^{n} P[X = k, Y = n - k]$.

Plugging into the first equation gives us

$\mathbb{E}[X + Y] = \sum_{n=0}^{\infty} n \sum_{k=0}^{n} P[X = k, Y = n - k]$

Note that every (ordered) pair of natural numbers is represented exactly once in the sum above, multiplied by the sum of the numbers. So

$\mathbb{E}[X + Y] = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty}(x + y)P[X = x, Y = y]$

Distributing yields

$\mathbb{E}[X+Y] = (\sum_{x=0}^{\infty} x(\sum_{y=0}^{\infty} P[X = x, Y = y]))+(\sum_{y=0}^{\infty} y(\sum_{x=0}^{\infty} P[X = x, Y = y]))$

By the law of total probability, the above reduces to

$\mathbb{E}[X + Y] = (\sum_{x=0}^{\infty} xP[X = x]) + (\sum_{y=0}^{\infty} yP[Y = y]))$

And therefore we have

$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

(b) Show that it is not always the case that $\mathbb{E}[X^2] = \mathbb{E}[X]^2$.

**Solution:** Let $X$ be a r.v. with $P[X = -1] = \frac{1}{2}$ and $P[X = 1] = \frac{1}{2}$. Then $P[X^2 = 1] = 1$, so $\mathbb{E}[X^2] = 1$, but $\mathbb{E}[X] = 0 = \mathbb{E}[X]^2$.

2. Let $n_0 = n$ and having chosen $n_i$, pick a number $n_{i+1}$ between 0 and $n_i$ (inclusive) at random.

(a) Let $\ell$ be the least number such that $n_\ell = 0$. In terms of $n$, what is the expected value of $\ell$? (Hint: let $e_n$ be the expected value of $\ell$ when $n_0 = n$. Find an expression for $e_n$ in terms of $e_k$ for $0 \le k \le n$, and solve for $e_n$.)

**Solution:** Notice that for all $n > 0$, $e_n = 1 + \frac{1}{n+1}\sum_{i=0}^{n} e_i$. This is because the process starting at $n$ requires one step to choose some $k \le n$ at random, and then apes the process starting at $k$; hence $e_n$ is one more than the average of the $e_k$ for $k \le n$. We can simplify this to obtain $e_n = \frac{1}{n} + (1 + \frac{1}{n}\sum_{i=0}^{n-1} e_i)$. If $n-1 > 0$, then $1 + \frac{1}{n}\sum_{i=0}^{n-1} e_i = e_{n-1}$, so $e_n = \frac{1}{n} + e_{n-1}$ for $n > 1$. Expanding this, using the fact that $e_1 = 2$, for $n > 0$ we have

$$e_n = \frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \cdots \frac{1}{2} + 2$$

(b) What is the expected value of the sum $\sum_{i=1}^{\ell} n_i$?

**Solution:** The expected value is $n$. We proceed similarly to part (a). For $n = 0$ we have $s_0 = 0$. For $n \ge 0$, let $s_n = \sum_{i=1}^{l} n_i$ for a given $n_0 = n$. Then $s_n = \frac{1}{n+1}\sum_{k=0}^{n}(s_k + k)$. Solving for $s_n$:

$s_n = \frac{1}{n+1}(s_n + n + \sum_{k=0}^{n-1}(s_k + k))$
$(n+1)s_n - s_n = n + \sum_{k=0}^{n-1}(s_k + k)$
$s_n = 1 + \frac{1}{n}\sum_{k=0}^{n-1}(s_k + k)$

We claim that the expression above is equivalent to $s_n = n$. As a base case, we have $s_0 = 0$. Suppose the claim holds up to $s_m$. Then

$s_{m+1} = 1 + \frac{1}{m+1}\sum_{k=0}^{m}(s_k + k)$
$= 1 + \frac{1}{m+1}(s_m + m + \sum_{k=0}^{m-1}(s_k + k))$
$= 1 + \frac{1}{m+1}(s_m + m + ms_m - m)$
$= 1 + \frac{1}{m+1}((1 + m)s_m)$
$= 1 + s_m = m + 1.$

3. (a) Taylor, a computer science concentrator, must complete 3 problem sets before doing laundry. Each problem set requires 2 days with probability $\frac{2}{3}$ and 3 days with probability $\frac{1}{3}$. What is the expected number of days Taylor delays laundry?

   (For example, if the first problem set requires 3 days and the second and third problem set each require 2 days, then Taylor delays laundry for 7 days.)

   (b) Blake, an English concentrator, must complete an essay before doing laundry. The length of the essay is equal to the sum of the numbers rolled on 2 fair, 6-sided dice. If Blake can write 1 page each day, what is the expected number of days Blake delays laundry?

   (For example, if the rolls are 5 and 3, then Blake delays laundry for 8 days.)

   **Solution**

   (a) Let $X$ represent the time it takes to complete a single problem set. Then $E(X) = 2(\frac{2}{3}) + 3(\frac{1}{3}) = \frac{7}{3}$. For 3 problem sets, $E(X + X + X) = 3E(X) = 7$ days

   (b) Let $X$ represent the value of a single dice roll. Then $E(X) = 1(\frac{1}{6}) + 2(\frac{1}{6}) + 3(\frac{1}{6}) + 4(\frac{1}{6}) + 5(\frac{1}{6}) + 6(\frac{1}{6}) = \frac{7}{2}$. Then the expected value of 2 rolls is $E(2X) = 2E(X) = 7$ days.

4. Evariste Galois (a great mathematician who actually died in a duel at age 19) is fighting a duel to the death with Gaston. They fire alternately, with Evariste going first: EGEGEG... Evariste is a terrible shot, and has only one chance in five of killing Gaston with any given shot. Gaston is a somewhat better marksman, and has one chance in four of killing Evariste with any given shot. Either a shot kills the opponent, or it misses completely.

   (a) By summing a geometric series, determine the probability that Evariste wins the duel.

   *Hint: Evariste wins if both duelists miss $r$ times, and he then hits.*

   (b) To check your answer, sum a series to calculate directly the probability that Gaston wins.

   **Solution**

   (a) Let's say that Evariste wins after both duelists shoots $r$ times. $r$ can range from 0 to infinity (the duel can potentially go on forever). The probability that Evariste wins is the probability that both Eraiste and

Gaston misses for $r$ shots and that Evariste kills Gaston at the $r+1$ shot.

$$\sum_{r=0}^{\infty} \frac{1}{5} \cdot (\frac{3}{4} \cdot \frac{4}{5})^r = \frac{1}{5} \cdot \sum_{r=0}^{\infty} (\frac{3}{5})^r = \frac{1}{5} \cdot (\frac{1}{1-\frac{3}{5}}) = \frac{1}{2}$$

(b) Probability that Gaston wins is the probability that both Eraiste and Gaston misses for $r$ shots, then Eraiste misses the $r+1$ shot, but Gaston doesn't:

$$\frac{4}{5} \cdot \frac{1}{4} \cdot \sum_{r=0}^{\infty} (\frac{3}{5})^r = \frac{1}{5} \cdot \sum_{r=0}^{\infty} (\frac{3}{5})^r = \frac{1}{5} \cdot (\frac{1}{1-\frac{3}{5}}) = \frac{1}{2}$$