Homework Assignment 7

Due Sunday Nov 3rd by 3pm

SDS 321 Intro to Probability and Statistics

1. **(4+1 pts)(Markov's inequality:)** Let X be a random variable that takes only non-negative values.

   (a) Prove that for any $s > 0$,

   $$P(X \geq s) \leq \frac{E[X]}{s} \tag{1}$$

   Hint: use the total expectation theorem!

   $E[X] = E[X|\{X \geq s\}]P(X \geq s) + \underbrace{E[X|\{X < s\}]P(X < s)}_{\geq 0 \text{ Since } X \geq 0} \geq E[X|\{X \geq s\}]P(X \geq s)$. But $E[X|X \geq s] \geq s$, since your conditional PMF or PDF is now defined on all values of X which are bigger than s. So the expectation will be bigger than s. So now we have $E[X] \geq sP(X \geq s)$, and so $P(X \geq s) \leq E[X]/s$. 1 pt for using the total expectation theorem correctly. 1pt for using $X \geq 0$. 1pt for $E[X|X \geq s] \geq s$. 1pt for the last step. full points for any correct proof.

   (b) You are tossing 500 biased coins, each having a probability 1/5 to come up heads. Using the Markov inequality above, can you give an upper bound on the probability of seeing 400 heads? X is a Binomial(500,1/5) r.v. and so $E[X] = 100$. $P(X > 400) \leq E[X]/400 = 100/400 = 0.25$

2. **(2.5+2.5 points)** You have a fair coin and two biased coins. The first biased coin lands heads with probability .3 whereas the second lands head with probability .8. You choose one of these three coins at random and now keep tossing it until you see a head.

   (a) What is the expected number of tosses until you see a head? By the law of total expectation, we'll take two variables: X is the number of flips until a head, Y is the coin chosen ($Y = 1$ being fair, $Y = 2$ being biased with success probability .3, and $Y = 3$ being biased with success probability .8). Given $Y = i$, X is a geometric distribution with the corresponding success probability.

   $$E[X] = \sum_{i=1}^{3} E[X|Y = i]P(Y = i)$$
   $$= \frac{1}{3}\left(E[X|Y = 1] + E[X|Y = 2] + E[X|Y = 3]\right)$$

   $$E[X|Y = 1] = \frac{1}{1/2} = 2$$

   $$E[X|Y = 2] = \frac{1}{.3} = 3.333$$

   $$E[X|Y = 3] = \frac{1}{.8} = 1.25$$

   $$E[X] = (2 + 3.333 + 1.25)/3 = 2.19$$

(b) What is the expected number of heads in 20 tosses? This is identical to the previous problem, except that X is now the number of heads in 20 tosses, so $X|Y = i$ is a binomial distribution with correpsonding success probability and $n = 20$:

$$E[X|Y = 1] = 20 \cdot .5 = 10$$
$$E[X|Y = 3] = 20 \cdot .3 = 6$$
$$E[X|Y = 4] = 20 \cdot .8 = 16$$
$$E[X] = \frac{1}{3}(10 + 6 + 16) = 32/3$$

3. **(3+1 pts)** Consider the following PMF for random variables X and Y.

X \ Y	1	2	P(X=x)
0	1/6	1/3	1/2
1	1/6	1/3	1/2
P(Y=y)	1/3	2/3	1

Table 1: Joint PMF of X and Y

(a) Calculate the missing values so that X and Y are independent. There may be multiple PMFs that satisfy this. Just provide one. also $P(X = 0, Y = 1) = 1/3$ and $P(X = 1, Y = 0) = 1/6$ works.

(b) Now compute $Var(X + Y)$. $Var(X + Y) = Var(X) + Var(Y)$. $Var(X) = EX^2 - EX^2 = 1/3 + 8/3 - (5/3)^2 = 2/9$ and $Var(Y) = 1/2 - 1/4 = 1/4$. So $Var(X + Y) = 1/4 + 2/9$.

4. **(3+1+2 pts)** (**Coupon Collector problem:**) There are 10 coupons of different colors in a box. You draw one coupon at random, note its color and put it back in the box. At any draw, each coupon is equally likely to be drawn.

(a) Say you have already seen $i - 1$ distinct coupon colors. Now, let X_i be the number of coupons you draw (with replacement) until you see a coupon whose color you have not seen before. Trivially, $X_1 = 1$, since the first coupon drawn always has a new color. For $i = 1, \ldots, 10$,

i. What is the distribution of X_i? After seeing $i-1$ colors the probability of seeing a new color when you pick a coupon with replacement is $(10 - (i - 1))/10$. So $X_i \sim geometric\left(\frac{11-i}{10}\right)$. One point for understanding its a geometric. Two points for the correct parameter.

ii. What is $E[X_i]$? $E[X_i] = \frac{10}{11-i}$. One point for plugging into the right formula.

(b) Let Y is the total number of draws needed to see all 10 colors. Show that $E[Y] = 10\left(\sum_{i=1}^{10} \frac{1}{i}\right)$. *Hint: write Y in terms of X_i's.*

(c) What is $E[X_i]$? $Y = \sum_i X_i$. So by linearity of expectation: $E[Y] = \sum_{i=1}^{10} \frac{10}{11-i} = 10 \sum_{i=1}^{10} \frac{1}{i}$. One point for understanding its a linearity of expectation problem. One point plugging into the right form.