Homework 4 (Due Sep 19, 2022)

Jack Hyatt
MATH 574 - Discrete Mathamatics - Fall 2022

April 24, 2023

1. Let X be a random variable on a sample space S such that $X(s) \geq 0$ for all $s \in S$. Prove that for every number $a > 0$, $p(X \geq a) \leq \frac{E(X)}{a}$. This is called **Markov's inequality**. [1]

   Let $X = X_< + X_\geq$, where $X_<$ is the subset of X where all elements are less than a and $X_\geq$ is the subset of X where all elements are greater than or equal a. So then

   $$P(X \geq a) = \sum_{r \in X_\geq} p(X = r)$$

   (Since $r \geq a$, r/a is greater than 1)

   $$\leq \sum_{r \in X_\geq} \frac{r}{a} \cdot p(X = r)$$

   $$\leq \frac{1}{a} \sum_{r \in X_\geq} r \cdot (X = r)$$

   $$\leq \frac{1}{a} \sum_{r \in X} r \cdot (X = r) = \frac{E(X)}{a}$$

   □

2. A biased coin has probability p of getting heads. Let X be the number of flips it takes to get exactly n heads.

   (a) Use the linearity of expectation to prove that $E(X) = n/p$. [2]

   Let X_i to be the number of flips it takes to get the ith heads after getting the $(i-1)$th heads. Then each E_i has a geometric distribution. There is an n amount of X_i, and using LoE, $E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} 1/p = \frac{n}{p}$.

[1]Hint: Use the formula $E(X) = \sum_{r \in X(S)} rP(X = r)$ and split it into $r < a$ and $r \geq a$.

[2]Hint: Define the random variable X_i to be the number of flips it takes to get the ith heads after getting the $(i-1)$th heads.

(b) Using part (a), give a double counting proof of the following:

$$\sum_{m=n}^{\infty} m\binom{m-1}{n-1}p^n(1-p)^{m-n} = n/p.$$

RHS: This is the expected number of flips it takes to get exactly n heads.
LHS: The $p^n(1-p)^{m-n}$ terms represent the probability of getting n heads in m flips. $\binom{m-1}{n-1}$ represents the ways you can order getting $n-1$ heads in $m-1$ flips, with the mth flip being heads. The m term will represent the value of the random variable for the outcome, the number of flips. Since it is summing from n to ∞, the sum covers all possible lengths for getting n heads. This is the formula for expected value, thus the left hand side is also the expected number of flips it takes to get exactly n heads.

3. A game is played where the player rolls 2 fair 6-sided dice. The player must pay \$1 to play the game. The player wins \$2 if the product of the two dice is an odd number, and \$1 if the sum of the two dice is an odd number.

(a) What is the player's expected net profit for this game?

This is the sum of the outcomes multiplied by their probabilities. The probability of the product of two dice being odd is 1/4, and the probability of two dice summing to an odd number is also 1/2 disjointed from the product being odd. So there is a 1/4 chance the player wins nothing. Therefore, the expected value is $1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{4} = 0$.

(b) What is the variance of the player's net profit?

We will use the formula $V(X) = E(X^2) - E(X)^2$. Thus we need to know X^2, which will just square the outcomes. This giving us 1,0,1, in the same order as before. This gives $E(X^2) = 1/2$. Thus $V(X) = E(X^2) - E(X)^2 = 1/2 - 0^2 = 1/2$.

4. Prove that for an integer $n \geq 1$, $\sum_{k=1}^{n} k = \binom{n+1}{2}$.

Proof. Since n is a natural number, the sum from 1 to n is well known and called the triangular numbers. The formula for triangular numbers from 1 to n is $\frac{n(n+1)}{2}$.
The right hand side equates to $\frac{(n+1)!}{2!(n-1)!}$ by definition of the Binomial Coefficient. Simplifying that fraction results in $\frac{n(n+1)}{2}$, which is the same as the left hand side. $\square$

5. A random subset of $\{1, \dots, n\}$ is chosen using the following process: for each element $i \in [n]$ we include i in the subset with probability 1/2. Let X be the random variable equal to the sum of the elements of the subset. Let Y be the random variable equal to the largest element in the subset.

(a) Compute $E(X)$.

Let $X_i = i$ if i is in the subset, and 0 otherwise. Then $X = \sum_{i=1}^{n} X_i$. So using LoE, $E(X) = \sum_{i=1}^{n} E(X_i)$. And trivially, $E(X_i) = i/2$. Therefore, $E(X) = \frac{1}{2}\sum_{i=1}^{n} i = \frac{n(n+1)}{4}$.

(b) Show that X and Y are not independent.

Let $X = 2$ and $Y = 1$. This has a 0 chance of happening since if $X = 2$, then the subset must be $\{2\}$, and so Y cannot be 1. But $P(Y = 1) > 0$ and $P(X = 2) > 0$, so their product is not 0. $\square$

6. Let X be a random variable that has geometric distribution with probability of success p. In this question we will show that $V(X) = \frac{1-p}{p^2}$.

(a) For r with $|r| < 1$, prove that

$$\sum_{k=1}^{\infty} k^2 r^{k-1} = \frac{1+r}{(1-r)^3}.$$

You may use results proven in class.

Assume $|r| < 1$, then

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \implies$$

$$\frac{d}{dr}(\sum_{k=0}^{\infty} r^k) = \frac{d}{dr}(\frac{1}{1-r}) \implies$$

$$\sum_{k=1}^{\infty} kr^{k-1} = \frac{1}{(1-r)^2} \implies$$

$$\sum_{k=1}^{\infty} kr^{k} = \frac{r}{(1-r)^2} \implies$$

$$\frac{d}{dr}(\sum_{k=1}^{\infty} kr^k) = \frac{d}{dr}(\frac{r}{(1-r)^2}) \implies$$

$$\sum_{k=1}^{\infty} k^2 r^{k-1} = \frac{1+r}{(1-r)^3} \quad \square$$

(b) Use part (a) to prove $V(X) = \frac{1-p}{p^2}$.

$V(X) = E(X^2) - E(X)^2$. Since X has a geometric distribution, $E(X) = 1/p$.

We now need to calculate $E(X^2)$. Squaring X will just square the outcomes, but leaves the probabilities alone. So

$$E(X^2) = \sum_{r \in X^2} rp(X = r) = \sum_{r=1}^{\infty} r^2 p(1-p)^{r-1} = p \sum_{r=1}^{\infty} r^2 (1-p)^{r-1}$$

Using the fact from part (a), we get

$$p \sum_{r=1}^{\infty} r^2 (1-p)^{r-1} = p \cdot \frac{1 + 1 - p}{(1 - (1-p))^3} = \frac{2-p}{p^2}$$

So $V(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$ □

7. Suppose that the number of cans of soda pop filled in a day at a bottling plant is a random variable with an expected value of 10,000 and a variance of 1000.

   (a) Use Markov's inequality (#1 on this homework) to obtain an upper bound on the probability that the plant will fill more than 11,000 cans on a particular day.

   Plugging in the values for Markov's inequality, we get $P(X \geq 11,001) \leq \frac{10,000}{11,001} \approx 0.909$.

   (b) Use Chebyshev's inequality to obtain a lower bound on the probability that the plant will fill between 9000 and 11,000 cans on a particular day.

   Since the variance is 1000, the standard deviation is $\sqrt{1000}$. So the number of standard deviations away from the expected value we are looking for is $(10000 - 9000)/\sigma = \sqrt{1000}$.
   Chebyshev's inequality is $p(|X - E(X)| \geq k\sigma) \leq \frac{1}{k^2}$. Since $k = \sqrt{1000}$, we can just let $k = \sigma$. So the inequality now turns into

$$p(|X - E(X)| \geq \sigma^2) \leq \frac{1}{\sigma^2} \implies p(|X - E(X)| \geq V(X)) \leq \frac{1}{V(X)} = \frac{1}{1000}$$

8. A biased coin has probability $p = .99$ for heads. Suppose we flip the coin 1000 times. Use Chebyshev's formula to give an upper bound for the probability that we get heads at most 900 times.

   Let X be the random variable where it equals the number of heads flipped. Then $X \sim B(n,p)$. So $E(X) = 1000 \cdot 0.99 = 990$. $V(X) = np(1-p) = 990 \cdot .01 = 9.9$. Therefore, $p(|X - E(X)| \geq 900) \leq \frac{9.9}{900^2}$