

# CSE 312 – Problem Set 5

Due Wednesday, May 6, at 11:59pm

Spring 2026

## Instructions

For details on expectations for solutions, collaboration policy, late policy, etc, see the instructions on problem set 1.

**Solutions submission.** You must submit your solution via Gradescope. In particular:

- Problem 1 will be done on Gradescope.
- Submit the solutions to problems 2-5 under “PSet 5 [Written]” as on previous problem sets.

## 1 Gradescope Questions (22 points)

Please do these [questions on gradescope](#).

## 2 A simple investing strategy (10 points)

A certain investor buys  $n$  different stocks, with each independently having probability  $p$  of going up in value each day, where  $0 < p < 1$ . At the end of each day, he keeps all the stocks that went up and sells all of the others. He repeats this process each day, until he has sold all of his stocks. We assume that each stock has probability  $p$  of going up on any given day and whether or not that happens is independent of what happens on other days or what happens with other stocks.

Let  $X_t$  be the random variable representing the number of stocks the investor owns at the end of  $t$  days, so  $X_0 = n$ . Explain each of your answers to the following questions (5 points each)

- What is the distribution of  $X_t$  for an arbitrary integer  $t$ ? In other words, what is  $\mathbb{P}[X_t = k]$ ? What distribution in our zoo is this, and what are its parameters?  
**Hint:** Think about the probability that a particular stock ‘survives’ until after day  $t$ .
- What is the probability that he owns at least one stock after  $t$  days? Your answer should not have any summations in it.

## 3 $X$ and $Y$ (15 points)

Let  $X$  be a Geometric RV with parameter  $p$ , let  $Y$  be a Poisson RV with parameter  $\lambda$ , and  $Z = \max(X, Y)$ . Assume that  $X$  and  $Y$  are independent. For each of the following problems, your final

answers should not have summations. You may want to use the Taylor series expansion of  $e^x$ , that is, that for any  $x$ ,

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

- a) (3 points) What is  $\mathbb{P}[X > k]$ ? (Think about it from first principles.)
- b) (6 points) Compute  $\mathbb{P}[X > Y]$ . (Consider the Law of Total Probability)

$$\mathbb{P}[X > Y] = \sum_{k=0}^{\infty} \mathbb{P}[X > Y \mid Y = k] \cdot \mathbb{P}[Y = k].$$

- c) (1 point) Compute  $\mathbb{P}[Z \geq X]$ .
- d) (5 points) Compute  $\mathbb{P}[Z \leq Y]$ .

## 4 Round-Robin Tennis (12 points)

A random number of  $N$  tennis players show up for a local round-robin tournament. To complete the tournament, each player must play exactly one match against every other player. Thus, a total of  $M = \binom{N}{2} = \frac{N(N-1)}{2}$  matches are played.

Answer each of the following questions (4 points each). Make sure that each of your answers are **not** in the form of a summation. **In each case include the expectation and variance of  $N$  as part of your answer.**

It may be useful to recall that for any random variable  $E(X^2) = \text{Var}(X) + [E(X)]^2$ .

- a) What is the expected value of  $M$  if  $N$  equals some fixed positive integer  $c$  with probability 1? (Some parts of your answer may depend on  $c$ .)
- b) What is the expected value of  $M$  if  $N$  has a Poisson distribution with parameter  $\lambda$ ? (Your answer will be a function of  $\lambda$ .)
- c) What is the expected value of  $M$  if  $N = 10X + 7$ , where  $X$  is a Bernoulli random variable with parameter  $p$ ? (Your answer will be a function of  $p$ .)

## 5 Sample Sampling Algorithm (18 points)

Consider the following algorithm for generating a random sample  $S$  of size  $n$  from the set of integers  $\{1, 2, \dots, N\}$ , where  $0 < n < N$ .

```

function sample( $N, n$ )
     $S \leftarrow \emptyset$  //  $S$  is a set of distinct integers, initially an empty set
    while  $|S| < n$  do
         $x \leftarrow \text{RollDie}(N)$  //  $x$  is the outcome of rolling a fair  $N$ -sided die
         $S \leftarrow S \cup \{x\}$  // if  $x$  is already in  $S$  it doesn't change
    return  $S$ 

```

Let  $I$  be the number of die rolls until  $S$  is returned. Also, let  $I_i$  be the random variable which describes the number of rolls it takes from the time the set  $S$  has  $i - 1$  values to the first time a new value is added after that (i.e., the set  $S$  has  $i$  values). Answer the following questions (6 points each)

- a) What type of random variable from our zoo is  $I_i$  and what is/are the relevant parameter(s) for that random variable?
- b) What is  $I$  in terms of the random variables  $I_i$ ? Calculate  $\mathbb{E}[I]$ , expressing the result as a summation that depends on both  $N$  and  $n$ .
- c) What is  $\text{Var}[I]$ ? You can leave your answer in summation form.