

# CSE 312 – Section 5 Solutions

Spring 2026

## Review of Main Concepts

- **Independence:** Random variable  $X$  and event  $E$  are independent iff

$$\forall x, \quad \mathbb{P}(X = x \cap E) = \mathbb{P}(X = x)\mathbb{P}(E)$$

Random variables  $X$  and  $Y$  are independent iff

$$\forall x \forall y, \quad \mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

In this case, we have  $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$  (the converse is not necessarily true).

- **i.i.d. (independent and identically distributed):** Random variables  $X_1, \dots, X_n$  are i.i.d. (or iid) iff they are mutually independent and have the same probability mass function.
- **Independence of functions of a r.v.:** If  $X$  and  $Y$  are independent and  $g(\cdot), h(\cdot)$  are functions mapping real numbers to real numbers, then  $g(X)$  and  $h(Y)$  are independent. (See if you can prove this!) Thus, for example: if  $X$  and  $Y$  are independent, then so are  $X^2$  and  $Y^2$ .
- **Variance of Independent Variables:** If  $X$  and  $Y$  are independent random variables,  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ . Linearity of variance does **not** hold in general. It depends on independence, whereas linearity of expectation always holds. Note that this combined with the above shows that  $\forall a, b, c \in \mathbb{R}$  and if  $X$  is independent of  $Y$ ,  $\text{Var}(aX + bY + c) = a^2\text{Var}(X) + b^2\text{Var}(Y)$ .
- Review: Zoo of Discrete Random Variables

- a) **Uniform:**  $X \sim \text{Uniform}(a, b)$  ( $\text{Unif}(a, b)$  for short), for integers  $a \leq b$ , iff  $X$  has the following probability mass function:

$$p_X(k) = \frac{1}{b - a + 1}, \quad k = a, a + 1, \dots, b$$

$\mathbb{E}[X] = \frac{a+b}{2}$  and  $\text{Var}(X) = \frac{(b-a)(b-a+1)}{12}$ . This represents a random experiment in which each outcome from  $[a, b]$  is equally likely. For example, a single roll of a fair die is  $\text{Uniform}(1, 6)$ .

- b) **Bernoulli (or indicator):**  $X \sim \text{Bernoulli}(p)$  ( $\text{Ber}(p)$  for short) iff  $X$  has the following probability mass function:

$$p_X(k) = \begin{cases} p, & k = 1 \\ 1 - p, & k = 0 \end{cases}$$

$\mathbb{E}[X] = p$  and  $\text{Var}(X) = p(1 - p)$ . An example of a Bernoulli r.v. is one flip of a coin with  $\mathbb{P}(\text{head}) = p$ .

- c) **Binomial:**  $X \sim \text{Binomial}(n, p)$  ( $\text{Bin}(n, p)$  for short) iff  $X$  is the sum of  $n$  iid Bernoulli( $p$ ) random variables.  $X$  has probability mass function

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

$\mathbb{E}[X] = np$  and  $\text{Var}(X) = np(1-p)$ . An example of a Binomial r.v. is the number of heads in  $n$  independent flips of a coin with  $\mathbb{P}(\text{head}) = p$ . Note that  $\text{Bin}(1, p) \equiv \text{Ber}(p)$ . As  $n \rightarrow \infty$  and  $p \rightarrow 0$ , with  $np = \lambda$ , then  $\text{Bin}(n, p) \rightarrow \text{Poi}(\lambda)$ . If  $X_1, \dots, X_n$  are independent Binomial r.v.'s, where  $X_i \sim \text{Bin}(N_i, p)$ , then  $X = X_1 + \dots + X_n \sim \text{Bin}(N_1 + \dots + N_n, p)$ .

- d) **Geometric:**  $X \sim \text{Geometric}(p)$  ( $\text{Geo}(p)$  for short) iff  $X$  has the following probability mass function:

$$p_X(k) = (1-p)^{k-1} p, \quad k = 1, 2, \dots$$

$\mathbb{E}[X] = \frac{1}{p}$  and  $\text{Var}(X) = \frac{1-p}{p^2}$ . An example of a Geometric r.v. is the number of independent coin flips up to and including the first head, where  $\mathbb{P}(\text{head}) = p$ .

- e) **Poisson:**  $X \sim \text{Poisson}(\lambda)$  ( $\text{Poi}(\lambda)$  for short) iff  $X$  has the following probability mass function:

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \dots$$

$\mathbb{E}[X] = \lambda$  and  $\text{Var}(X) = \lambda$ . An example of a Poisson r.v. is the number of people born during a particular minute, where  $\lambda$  is the average birth rate per minute. If  $X_1, \dots, X_n$  are independent Poisson r.v.'s, where  $X_i \sim \text{Poi}(\lambda_i)$ , then  $X = X_1 + \dots + X_n \sim \text{Poi}(\lambda_1 + \dots + \lambda_n)$ .

## Announcements & Plan for Section

### Announcements

- PSet 3 grades were released and can be viewed on Gradescope. Regrade requests will close on 5/2. We highly recommend taking a look through feedback on the psets, the common errors on Ed and the solutions that were posted (link on Ed).
- PSet 4 was due earlier this week.
- PSet 5 is released — start early.
- This week's focus: Random Variable Zoo + applications, independence and more practice with expectation/variance.
- A practice midterm will be posted today or tomorrow. Please attempt the problems before next week's section where we will review the solutions.

### Plan for Section

- **Content Review (Problem 1)**  
Quick review of concepts such as variance properties, linearity of expectation and independence.
- **Problem 2 (Pond Fishing)**  
Build intuition for the different zoo variables and what each of them represent.

- **Problem 4 (Variance of a Product)**

Gives students practice with the definition of variance, and understanding when linearity of variance holds.

- **Problem 5 (True or False?)**

Exam-style true or false questions to strengthen the definitions of zoo variables.

- **Problem 3 (Best Coach Ever!! - if time)**

Gives students intuition for when to apply which zoo variable, practices geometric and binomial variables.

## 1 Content Review Questions

a) True or false:  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$  for all random variables  $X$  and  $Y$

False. This property only holds if  $X$  and  $Y$  are independent.

b) What is  $\text{Var}(3X + 4)$ ?

- $3\text{Var}(X) + 4$
- $3\text{Var}(X)$
- $9\text{Var}(X)$
- $\text{Var}(X)$

$9\text{Var}(X)$  by the scaling property of variance

c) True or false:  $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$  for all random variables  $X$  and  $Y$ .

True. This is by the linearity of expectation.  $X$  and  $Y$  do not have to be independent.

d) What is  $\mathbb{E}[3X + 4]$ ?

- $3\mathbb{E}[X] + 4$
- $3\mathbb{E}[X]$
- $9\mathbb{E}[X]$
- $\mathbb{E}[X]$

$3\mathbb{E}[X] + 4$  by the linearity of expectation.

## 2 Pond Fishing

Suppose I am fishing in a pond with  $B$  blue fish,  $R$  red fish, and  $G$  green fish, where  $B + R + G = N$ . Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

a) How many of the next 10 fish I catch are blue, if I catch and release them

$$\text{Bin}\left(10, \frac{B}{N}\right)$$

$$\text{Ber}\left(\frac{B}{N}\right)$$

$$\text{Bin}\left(1, \frac{B}{N}\right)$$

Since this is the same as saying how many of my next 10 trials (fish) are a success (are blue), this is a binomial distribution. Specifically, since we are doing catch and release, the probability of a given fish being blue is  $\frac{B}{N}$  and each trial is independent. Thus:

$$\text{Bin}\left(10, \frac{B}{N}\right)$$

b) How many fish I had to catch until my first green fish, if I catch and release them

$$\text{Ber}\left(\frac{G}{N}\right)$$

$$\text{Bin}\left(1, \frac{G}{N}\right)$$

$$\text{Geo}\left(\frac{G}{N}\right)$$

Once again, each catch is independent, so this is asking how many trials until we see a success, hence it is a geometric distribution:

$$\text{Geo}\left(\frac{G}{N}\right)$$

c) How many red fish I catch in the next five minutes, if I catch on average  $r$  red fish per minute

$\text{Poi}(5R)$

$\text{Bin}\left(5, \frac{R}{N}\right)$

$\text{Poi}(5r)$

This is asking for the number of occurrences of an event given an average rate, which is the definition of the Poisson distribution. Since we're looking for events in the next 5 minutes, that is our time unit, so we have to adjust the average rate to match ( $r$  per minute becomes  $5r$  per 5 minutes).

$\text{Poi}(5r)$

d) Whether or not my next fish is blue

$\text{Poi}(5B)$

$\text{Bin}\left(1, \frac{R}{N}\right)$

$\text{Ber}\left(\frac{B}{N}\right)$

This is the same as the binomial case, but it's only one trial, so it is necessarily Bernoulli.

$\text{Ber}\left(\frac{B}{N}\right)$

### 3 Best Coach Ever!!

You are a hardworking boxer. Your coach tells you that the probability of your winning a boxing match is 0.2 independently of every other match.

a) How many matches do you expect to fight until you win 10 times and what kind of random variable is this?

The number of matches you have to fight until you win 10 times can be modeled by  $\sum_{i=1}^{10} X_i$  where  $X_i \sim \text{Geometric}(0.2)$  is the number of matches you have to fight to go from  $i - 1$  wins to  $i$  wins, including the match that gets you your  $i^{\text{th}}$  win, where every match has a 0.2 probability of success. Recall  $\mathbb{E}[X_i] = \frac{1}{0.2} = 5$ .  $\mathbb{E}[\sum_{i=1}^{10} X_i] =$

$$\sum_{i=1}^{10} \mathbb{E}[X_i] = \sum_{i=1}^{10} \frac{1}{0.2} = 10 \cdot 5 = 50.$$

- b) You only get to play 12 matches every year. To win a spot in the Annual Boxing Championship, a boxer needs to win at least 10 matches in a year. What is the probability that you will go to the Championship this year and what kind of random variable is the number of matches you win out of the 12?

You can go to the championship if you win more than or equal to 10 times this year. Let  $Y$  be the number of matches you win out of the 12 matches. Note that  $Y \sim \text{Binomial}(12, 0.2)$ . Since the max number you can win is 12 (there are 12 matches), we are looking for  $\mathbb{P}(10 \leq Y \leq 12)$ . Thus, since  $Y$  is discrete, we are interested in

$$\mathbb{P}(Y = 10) + \mathbb{P}(Y = 11) + \mathbb{P}(Y = 12) = \sum_{i=10}^{12} \binom{12}{i} 0.2^i (1 - 0.2)^{12-i}$$

- c) Let  $p$  be your answer to part (b). How many times can you expect to go to the Championship in your 20 year career?

The number of times you go to the championship can be modeled by  $Y \sim \text{Binomial}(20, p)$  as whether or not we go to the championship every year is independent of other years, and since we are trying to find how many of these 20 trials is a success. So,  $\mathbb{E}[Y] = 20 \cdot p$ .

## 4 Variance of a Product

Let  $X, Y, Z$  be independent random variables with means  $\mu_X, \mu_Y, \mu_Z$  and variances  $\sigma_X^2, \sigma_Y^2, \sigma_Z^2$ , respectively. Find  $\text{Var}(XY - Z)$ .

First recall that  $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \implies \mathbb{E}[X^2] = \text{Var}(X) + \mathbb{E}[X]^2 = \sigma_X^2 + \mu_X^2$ , and same for  $Y$ .

$$\begin{aligned} \text{Var}(XY) &= \mathbb{E}[X^2Y^2] - \mathbb{E}[XY]^2 \quad (\text{by theorem in class}) \\ &= \mathbb{E}[X^2]\mathbb{E}[Y^2] - (\mathbb{E}[X]\mathbb{E}[Y])^2 \quad (\text{by independence}) \\ &= \mathbb{E}[X^2]\mathbb{E}[Y^2] - \mathbb{E}[X]^2\mathbb{E}[Y]^2 \\ &= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2 \end{aligned}$$

By independence,

$$\begin{aligned} \text{Var}(XY - Z) &= \text{Var}(XY) + \text{Var}(-Z) = \text{Var}(XY) + \text{Var}(Z) \\ &= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2 + \sigma_Z^2 \end{aligned}$$

## 5 True or False?

Identify the following statements as true or false (true means always true). Justify your answer.

- a) For any random variable  $X$ , we have  $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$ .

True, since  $0 \leq \text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$ , since the squaring necessitates the result is non-negative.

- b) Let  $X, Y$  be random variables. Then,  $X$  and  $Y$  are independent if and only if  $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ .

False. The forward implication is true, but the reverse is not. For example, if  $X$  is the discrete uniform random variable on the set  $\{-1, 0, 1\}$  such that  $\mathbb{P}(X = -1) = \mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{3}$ , and  $Y = X^2$ , we have  $\mathbb{E}[X] = 0$ , so  $\mathbb{E}[X]\mathbb{E}[Y] = 0$ . However, since  $X = X^3$ ,  $\mathbb{E}[XY] = \mathbb{E}[X^3] = \mathbb{E}[X] = 0$ , we have that  $\mathbb{E}[X]\mathbb{E}[Y] = 0 = \mathbb{E}[XY]$ . However,  $X$  and  $Y$  are not independent; indeed,  $\mathbb{P}(Y = 0 \mid X = 0) = 1 \neq \frac{1}{3} = \mathbb{P}(Y = 0)$ .

- c) Let  $X \sim \text{Binomial}(n, p)$  and  $Y \sim \text{Binomial}(m, p)$  be independent. Then,  $X + Y \sim \text{Binomial}(n + m, p)$ .

True.  $X$  is the sum of  $n$  independent Bernoulli trials, and  $Y$  is the sum of  $m$ . So  $X + Y$  is the sum of  $n + m$  independent Bernoulli trials, so  $X + Y \sim \text{Binomial}(n + m, p)$ .

- d) Let  $X_1, \dots, X_{n+1}$  be independent Bernoulli( $p$ ) random variables. Then,  $\mathbb{E}[\sum_{i=1}^n X_i X_{i+1}] = np^2$ .

True. Notice that  $X_i X_{i+1}$  is also Bernoulli (only takes on 0 and 1), but is 1 iff both are 1, so  $X_i X_{i+1} \sim \text{Bernoulli}(p^2)$ . The statement holds by linearity, since  $\mathbb{E}[X_i X_{i+1}] = p^2$ .

- e) Let  $X_1, \dots, X_{n+1}$  be independent Bernoulli( $p$ ) random variables. Then,  $Y = \sum_{i=1}^n X_i X_{i+1} \sim \text{Binomial}(n, p^2)$ .

False. They are all Bernoulli( $p^2$ ) as determined in the previous part, but they are not independent. Indeed,  $\mathbb{P}(X_1 X_2 = 1 \mid X_2 X_3 = 1) = \mathbb{P}(X_1 = 1) = p \neq p^2 = \mathbb{P}(X_1 X_2 = 1)$ .

- f) If  $X \sim \text{Bernoulli}(p)$ , then  $nX \sim \text{Binomial}(n, p)$ .

False. The range of  $X$  is  $\{0, 1\}$ , so the range of  $nX$  is  $\{0, n\}$ .  $nX$  cannot be  $\text{Bin}(n, p)$ , otherwise its range would be  $\{0, 1, \dots, n\}$ .

- g) If  $X \sim \text{Binomial}(n, p)$ , then  $\frac{X}{n} \sim \text{Bernoulli}(p)$ .

False. Again, the range of  $X$  is  $\{0, 1, \dots, n\}$ , so the range of  $\frac{X}{n}$  is  $\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ . Hence it cannot be  $\text{Ber}(p)$ , otherwise its range would be  $\{0, 1\}$ .

h) For any two independent random variables  $X, Y$ , we have  $\text{Var}(X - Y) = \text{Var}(X) - \text{Var}(Y)$ .

False.  $\text{Var}(X - Y) = \text{Var}(X + (-Y)) = \text{Var}(X) + (-1)^2 \text{Var}(Y) = \text{Var}(X) + \text{Var}(Y)$ .

## 6 Fun with Poissons

Let  $X \sim \text{Poisson}(\lambda_1)$  and  $Y \sim \text{Poisson}(\lambda_2)$ , and  $X$  and  $Y$  are independent.

a) Show that  $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$ . (We did this problem in class.)

To show that a random variable is distributed according to a particular distribution, we must show that they have the same PMF. Thus, we are trying to show that  $\mathbb{P}(X + Y = n) = e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^n}{n!}$

$$\begin{aligned}
 \mathbb{P}(X + Y = n) &= \sum_{k=0}^n \mathbb{P}(X = k \cap Y = n - k) \\
 &= \sum_{k=0}^n \mathbb{P}(X = k) \mathbb{P}(Y = n - k) && \text{[X and Y are independent]} \\
 &= \sum_{k=0}^n e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} \\
 &= e^{-(\lambda_1 + \lambda_2)} \sum_{k=0}^n \frac{\lambda_1^k}{k!} \frac{\lambda_2^{n-k}}{(n-k)!} \\
 &= e^{-(\lambda_1 + \lambda_2)} \sum_{k=0}^n \frac{1}{k!(n-k)!} \lambda_1^k \lambda_2^{n-k} \\
 &= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} \lambda_1^k \lambda_2^{n-k} \\
 &= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} \sum_{k=0}^n \binom{n}{k} \lambda_1^k \lambda_2^{n-k} \\
 &= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} (\lambda_1 + \lambda_2)^n && \text{[Binomial Theorem]}
 \end{aligned}$$

b) Show that  $\mathbb{P}(X = k \mid X + Y = n) = \mathbb{P}(W = k)$  where  $W \sim \text{Bin}(n, \frac{\lambda_1}{\lambda_1 + \lambda_2})$

$$\begin{aligned}
\mathbb{P}(X = k \mid X + Y = n) &= \frac{\mathbb{P}(X = k \cap X + Y = n)}{\mathbb{P}(X + Y = n)} \\
&= \frac{\mathbb{P}(X = k \cap Y = n - k)}{\mathbb{P}(X + Y = n)} \\
&= \frac{\mathbb{P}(X = k)\mathbb{P}(Y = n - k)}{\mathbb{P}(X + Y = n)} \quad [\text{X and Y are independent}] \\
&= \frac{e^{-\lambda_1} \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!}}{e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^n}{n!}} \\
&= \frac{\frac{\lambda_1^k}{k!} \cdot \frac{\lambda_2^{n-k}}{(n-k)!}}{\frac{(\lambda_1 + \lambda_2)^n}{n!}} \\
&= \frac{n!}{k!(n-k)!} \cdot \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n} \\
&= \binom{n}{k} \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^k (\lambda_1 + \lambda_2)^{n-k}} \\
&= \binom{n}{k} \left( \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left( \frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-k} \\
&= \binom{n}{k} p^k (1-p)^{n-k}, \text{ where } p = \frac{\lambda_1}{\lambda_1 + \lambda_2}
\end{aligned}$$

## 7 Memorylessness

We say that a random variable  $X$  is memoryless if  $\mathbb{P}(X > k + i \mid X > k) = \mathbb{P}(X > i)$  for all non-negative integers  $k$  and  $i$ . The idea is that  $X$  does not *remember* its history. Let  $X \sim \text{Geo}(p)$ . Show that  $X$  is memoryless.

Let's note that if  $X \sim \text{Geo}(p)$ , then  $\mathbb{P}(X > k) = \mathbb{P}(\text{no successes in the first } k \text{ trials}) = (1-p)^k$ .

$$\begin{aligned}
\mathbb{P}(X > k + i \mid X > k) &= \frac{\mathbb{P}(X > k \mid X > k + i) \mathbb{P}(X > k + i)}{\mathbb{P}(X > k)} \quad [\text{Defn of Cond. Probability}] \\
&= \frac{\mathbb{P}(X > k + i)}{\mathbb{P}(X > k)} \quad [\mathbb{P}(X > k \mid X > k + i) = 1] \\
&= \frac{(1-p)^{k+i}}{(1-p)^k} \quad [\mathbb{P}(X > k) = (1-p)^k] \\
&= (1-p)^i \\
&= \mathbb{P}(X > i)
\end{aligned}$$

## 8 Poisson Practice

Seattle averages 3 days with snowfall per year. Suppose the number of days with snowfall follows a Poisson distribution.

- a) What is the probability of getting exactly 5 days of snow in a year?

$$\text{Let } X \sim \text{Poi}(3) \text{ Then } p_X(5) = \frac{3^5 e^{-3}}{5!} \approx .1008$$

- b) According to the Poisson model, what is the probability of getting 367 days of snow?

Let  $X \sim \text{Poi}(3)$ . Then  $p_X(367) = \frac{3^{367} e^{-3}}{367!} \approx 1.8 \times 10^{-610}$ , that's a very small estimate, but of course the true probability is 0. Recall that using a Poisson distribution is a modeling assumption, it may produce nonzero probabilities for events that are practically impossible.

## 9 How many 6's?

Suppose that a fair 8-sided die is rolled repeatedly, with each roll independent of the others. Let  $Z$  be the number of rolls until (and including) the first time either a 2 or a 3 is rolled, and let  $W$  be the number of 6's rolled until the first 2 or 3 is rolled. So, for example if the sequence of die values until the first 2 or 3 is 6,5,4,8,7,6,7,1,2, then  $Z$  is 9 and  $W$  is 2.

Define

$$p(j) := \begin{cases} \mathbb{P}(W = j \mid Z = i) & j \in \{0, 1, \dots, i-1\} \\ 0 & \text{otherwise} \end{cases}$$

Show that  $p(j)$  is the probability mass function of a binomially distributed random variable and determine its parameters  $n$  and  $p$ . Prove that your answer is correct mathematically.

$$\begin{aligned} \mathbb{P}(W = j \mid Z = i) &= \frac{\mathbb{P}(W = j, Z = i)}{\mathbb{P}(Z = i)} \\ &= \frac{\binom{i-1}{j} \left(\frac{1}{8}\right)^j \left(\frac{5}{8}\right)^{i-1-j} \frac{2}{8}}{\left(\frac{6}{8}\right)^{i-1} \frac{2}{8}} \\ &= \frac{\binom{i-1}{j} \left(\frac{1}{8}\right)^j \left(\frac{5}{8}\right)^{i-1-j}}{\left(\frac{6}{8}\right)^j \left(\frac{6}{8}\right)^{i-1-j}} \\ &= \binom{i-1}{j} \left(\frac{1}{6}\right)^j \left(\frac{5}{6}\right)^{i-1-j} \end{aligned}$$

since all the 8's in the denominators (on top and bottom) cancel out. Note that the numerator of (9) is the probability of the event that we have  $j$  6's in the first  $i-1$  rolls (since  $Z = i$  tells us that the  $i$ th roll must be a 2 or 3). Each particular sequence where  $j$  of the first  $i-1$  rolls are 6's,  $i-j-1$  of the first  $i-1$  rolls are 1,4,5,7 or 8 and the final roll is a 2 or 3 has

probability

$$\mathbb{P}(W = j, Z = i) = \left(\frac{1}{8}\right)^j \left(\frac{5}{8}\right)^{i-1-j} \frac{2}{8}$$

and there are  $\binom{i-1}{j}$  such sequences.

We can re-write the last line of our working above as

$$\mathbb{P}(W = j | Z = i) = \binom{i-1}{j} \left(\frac{1}{6}\right)^j \left(1 - \frac{1}{6}\right)^{i-1-j}$$

which is a binomial random variable with  $p = \frac{1}{6}$ ,  $n = i - 1$ .

## 10 Practice with LOTUS

Suppose that  $X$  is a Binomial random variable with parameters  $n$  and  $p$ . Use LOTUS to show that

$$\mathbb{E} \left[ \frac{1}{X+1} \right] = \frac{1 - (1-p)^{n+1}}{(n+1)p}.$$

Hint: Use the result of Problem 2 from Section 2, namely that  $\frac{1}{k+1} \binom{n}{k} = \frac{1}{n+1} \binom{n+1}{k+1}$  and then manipulate the expression until you can use the Binomial Theorem.

By LOTUS, the expected value of a function of a random variable  $g(X)$  is  $\mathbb{E}[g(X)] = \sum_k g(k) \mathbb{P}(X = k)$ . Applying this to  $g(X) = \frac{1}{X+1}$ , we get:

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{X+1} \right] &= \sum_{k=0}^n \frac{1}{k+1} \mathbb{P}(X = k) \\ &= \sum_{k=0}^n \frac{1}{k+1} \binom{n}{k} p^k (1-p)^{n-k} \end{aligned}$$

Using the hint, we substitute  $\frac{1}{k+1} \binom{n}{k} = \frac{1}{n+1} \binom{n+1}{k+1}$ :

$$\mathbb{E} \left[ \frac{1}{X+1} \right] = \sum_{k=0}^n \frac{1}{n+1} \binom{n+1}{k+1} p^{k+1} (1-p)^{n-k}$$

We can factor out  $\frac{1}{n+1}$ . To make the exponent of  $p$  match the bottom term in the binomial coefficient  $\binom{n+1}{k+1}$ , we multiply and divide the expression by  $p$ :

$$\mathbb{E} \left[ \frac{1}{X+1} \right] = \frac{1}{(n+1)p} \sum_{k=0}^n \binom{n+1}{k+1} p^{k+1} (1-p)^{n-k}$$

Next, we perform a change of variables. Let  $j = k + 1$ . When  $k$  goes from 0 to  $n$ ,  $j$  goes from 1 to  $n + 1$ . Notice that  $n - k = n - (j - 1) = n + 1 - j$ . Substituting  $j$  into our summation:

$$\mathbb{E} \left[ \frac{1}{X+1} \right] = \frac{1}{(n+1)p} \sum_{j=1}^{n+1} \binom{n+1}{j} p^j (1-p)^{n+1-j}$$

By the Binomial Theorem, we know that  $\sum_{j=0}^{n+1} \binom{n+1}{j} p^j (1-p)^{n+1-j} = (p + (1-p))^{n+1} = 1$ .

Our summation is almost exactly this Binomial expansion, except it is missing the  $j = 0$  term. We can rewrite our summation as the full Binomial expansion minus that  $j = 0$  term:

$$\begin{aligned} \sum_{j=1}^{n+1} \binom{n+1}{j} p^j (1-p)^{n+1-j} &= \sum_{j=0}^{n+1} \left[ \binom{n+1}{j} p^j (1-p)^{n+1-j} \right] - \binom{n+1}{0} p^0 (1-p)^{n+1-0} \\ &= 1 - (1-p)^{n+1} \end{aligned}$$

Substituting this back into our expected value equation yields the final result:

$$\mathbb{E} \left[ \frac{1}{X+1} \right] = \frac{1}{(n+1)p} (1 - (1-p)^{n+1}) = \frac{1 - (1-p)^{n+1}}{(n+1)p}$$

## 11 Parallel Server Failures

A computing system relies on  $m$  independent nodes. The lifetime (in days) of each node is modeled by a Geometric distribution with parameter  $p$  (meaning each node has a probability  $p$  of failing on any given day). The system completely shuts down only when *all*  $m$  nodes have failed.

- Let  $D$  be the day the *first* node fails. Find the probability mass function of  $D$ . Start by computing the probability that  $D > d$ .
- Find the exact probability that the entire system shuts down on day  $k$ .

- First Failure (Minimum):** Let  $X_i$  be the lifetime of node  $i$ . We want the distribution of  $D = \min(X_1, X_2, \dots, X_m)$ . The event  $D > d$  means all  $m$  nodes survive past day  $d$ .

$$\begin{aligned} \mathbb{P}(D > d) &= \mathbb{P}(X_1 > d, X_2 > d, \dots, X_m > d) \\ &= \prod_{i=1}^m \mathbb{P}(X_i > d) \end{aligned}$$

Since  $\mathbb{P}(X_i > d) = (1-p)^d$ :

$$\mathbb{P}(D > d) = ((1-p)^d)^m = (1-p)^{dm}$$

This represents the tail probability of a new Geometric distribution since it has the form (probability of no failure in one day) <sup>$d$</sup> . In particular, on any given day, all  $m$  nodes survive with probability  $(1-p)^m$ , so the probability that at least one node fails on a given day is

$$1 - (1-p)^m$$

Thus, we can think of each day as an independent trial where the probability of “success” (failure of at least one machine) is  $1 - (1-p)^m$ . Therefore,  $D \sim \text{Geometric}(1 - (1-p)^m)$ .

- System Shutdown (Maximum):** Let  $Y$  be the day the entire system shuts down, so  $Y = \max(X_1, X_2, \dots, X_m)$ . The probability that all nodes fail *by* day  $k$  is the cumulative

distribution function  $F_Y(k)$ :

$$\begin{aligned} F_Y(k) &= \mathbb{P}(\max(X_1, X_2, \dots, X_m) \leq k) \\ &= \mathbb{P}(X_1 \leq k, X_2 \leq k, \dots, X_m \leq k) \\ &= \prod_{i=1}^m \mathbb{P}(X_i \leq k) \\ &= [1 - (1 - p)^k]^m \end{aligned}$$

The probability the system shuts down *exactly* on day  $k$  requires subtracting the probability that it shut down on any day prior:

$$\begin{aligned} \mathbb{P}(Y = k) &= \mathbb{P}(Y \leq k) - \mathbb{P}(Y \leq k - 1) \\ &= [1 - (1 - p)^k]^m - [1 - (1 - p)^{k-1}]^m \end{aligned}$$

## 12 Grading time

Suppose that homeworks are graded by TA1 with probability 0.3, by TA2 with probability 0.5 and by TA3 with probability 0.2.

- TA1 takes an amount of time (in hours) to grade that is Poisson with parameter  $\lambda$ .
- TA2 takes an amount of time (in hours) to grade that is Binomial with parameters  $n$  and  $p$ .
- TA3 takes an amount of time (in hours) to grade that is Geometric with parameter  $p$ .

What is the probability that TA1 did the grading given that the number of hours it took was  $h$ ?

Let  $H$  be the number of hours it took to grade. By Bayes Theorem,

$$\begin{aligned} \mathbb{P}(\text{TA1 graded} | H = h) &= \frac{\mathbb{P}(H = h | \text{TA1 graded}) \cdot \mathbb{P}(\text{TA1 graded})}{\mathbb{P}(H = h)} \\ &= \frac{e^{-\lambda} \frac{\lambda^h}{h!} \cdot 0.3}{\mathbb{P}(H = h)} \\ &= \frac{e^{-\lambda} \frac{\lambda^h}{h!} \cdot 0.3}{0.3 \cdot e^{-\lambda} \frac{\lambda^h}{h!} + 0.5 \cdot \binom{n}{h} p^h (1 - p)^{n-h} + 0.2 \cdot (1 - p)^{h-1} p}, \end{aligned}$$

where we used the Law of Total Probability to compute the denominator in the last step.