CSE 312 – Section 5

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, $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. 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, $\mathrm{Var}(aX + bY + c) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$.

- Review: Zoo of Discrete Random Variables

  a) **Uniform**: $X \sim \mathsf{Uniform}(a, b)$ ($\mathsf{Unif}(a, b)$ 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 $\mathrm{Var}(X) = \frac{(b-a)(b-a+2)}{12}$. 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 $\mathsf{Uniform}(1, 6)$.

  b) **Bernoulli (or indicator)**: $X \sim \mathsf{Bernoulli}(p)$ ($\mathsf{Ber}(p)$ 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 $\mathrm{Var}(X) = p(1 - p)$. An example of a Bernoulli r.v. is one flip of a coin with $\mathbb{P}(\text{head}) = p$.

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

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

$\mathbb{E}[X] = np$ and $\mathrm{Var}(X) = np(1-p)$. 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 $\mathsf{Bin}(1, p) \equiv \mathsf{Ber}(p)$. As $n \to \infty$ and $p \to 0$, with $np = \lambda$, then $\mathsf{Bin}(n, p) \to \mathsf{Poi}(\lambda)$. If $X_1, \ldots, X_n$ are independent Binomial r.v.'s, where $X_i \sim \mathsf{Bin}(N_i, p)$, then $X = X_1 + \ldots + X_n \sim \mathsf{Bin}(N_1 + \ldots + N_n, p)$.

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

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

$\mathbb{E}[X] = \frac{1}{p}$ and $\mathrm{Var}(X) = \frac{1-p}{p^2}$. 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**: $X \sim \mathsf{Poisson}(\lambda)$ ($\mathsf{Poi}(\lambda)$ for short) iff X has the following probability mass function:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

○ $$\text{Poi}(5R)$$

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

○ $$\text{Poi}(5r)$$

d) Whether or not my next fish is blue

○ $$\text{Poi}(5B)$$

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

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

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?

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?

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?

4 Variance of a Product

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

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 $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$.

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

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

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

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

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

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

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

6 Fun with Poissons

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

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

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

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 $X \sim \mathsf{Geo}(p)$. Show that X is memoryless.

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?

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

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.

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.

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.

a) Let D be the day the *first* node fails. Find the probability mass function of D. Start by computing the probability that $D > d$.

b) Find the exact probability that the entire system shuts down on day k.

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.

a) TA1 takes an amount of time (in hours) to grade that is Poisson with parameter λ.

b) TA2 takes an amount of time (in hours) to grade that is Binomial with parameters n and p.

c) 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?