# CSE 312 Section 4

#### **Random Variables and Expectation**

#### Administrivia

#### **Announcements & Reminders**

• HW2

- Grades released on gradescope check your submission to read comments
- Regrade requests open ~24 hours after grades are released and close after a week

#### • HW3

- Written and Coding due yesterday, Wednesday 4/17 @ 11:59pm
  - Late deadline Saturday 4/22 @ 10pm
- # late days used for HW3 = Max(late days for written, late days for coding)
  - If you submit one day late for homework but two days late for coding, it counts for two late days total
- HW4
  - Released on the course website
  - Due Wednesday 4/24 @ 11:59pm
  - This is the last homework before the midterm!

## **Review & Questions**



### Any lingering questions from this last week?

Each week in section, we'll be reviewing the main concepts from this week and putting them into action by going through some practice problems together. But before we get into that review, we'll try to start off each section with some time for you to ask questions. Was anything particularly confusing this week? Is there anything we can clarify before we dive into the review? This is your chance to clear things up!

#### **Kahoot for content review!** *see task 1 from section handout*

#### **Review of Main Concepts**

- **Random Variable (rv):** A numeric function  $X : \Omega \to \mathbb{R}$  of the outcome.
- **Range/Support:** The support/range of a random variable *X*, denoted  $\Omega_X$ , is the set of all possible values that *X* can take on.
- **Discrete Random Variable (drv):** A random variable taking on a countable (either finite or countably infinite) number of possible values.
- **Probability Mass Function (pmf)** for a discrete random variable *X*: a function  $p_X : \Omega_X \to [0,1]$  with  $p_X(x) = \mathbb{P}(X = x)$  that maps possible values of a discrete random variable to the probability of that value happening, such that  $\sum_x p_X(x) = 1$ .
- **Cumulative Distribution Function (CDF)** for a random variable *X*: a function  $F_X$ :  $\mathbb{R} \to \mathbb{R}$  with  $F_1(x) = \mathbb{P}(Y \le x)$ 
  - $: \mathbb{R} \to \mathbb{R} \text{ with } F_X(x) = \mathbb{P}(X \le x)$

#### **Review of Main Concepts**

• **Expectation (expected value, mean, or average):** The expectation of a discrete random variable is defined to be

$$\mathbb{E}[X] = \sum_{x} x \, p_X(x) = \sum_{x} x \, \mathbb{P}(X = x)$$

The expectation of a function of a discrete random variable g(X) is

$$\mathbb{E}[g(X)] = \sum_{x} g(x) \, p_X(x)$$

# **Problem 2 – Identify that Range!**



Identify the support/range  $\Omega_X$  of the random variable X, if X is...

- a) The sum of two rolls of a six-sided die.
- b) The number of lottery tickets I buy until I win it.
- c) The number of heads in *n* flips of a coin with  $0 < \mathbb{P}(\text{head}) < 1$ .
- d) The number of heads in *n* flips of a coin with  $\mathbb{P}(\text{head}) = 1$ .
- e) The number of whole minutes I wait at the bus stop for the next bus.

Work on this problem with the people around you, and then we'll go over it together!

Identify the support/range  $\Omega_X$  of the random variable X, if X is...

a) The sum of two rolls of a six-sided die.

Min:

Max:

In between?

Identify the support/range  $\Omega_X$  of the random variable X, if X is...

a) The sum of two rolls of a six-sided die.

X takes on every integer value between the min sum 2, and the max sum 12.

 $\Omega_X = \{2, 3, \dots, 12\}$ 

Identify the support/range  $\Omega_X$  of the random variable X, if X is...

b) The number of lottery tickets I buy until I win it.

Min:

Max:

In between?

Identify the support/range  $\Omega_X$  of the random variable X, if X is...

b) The number of lottery tickets I buy until I win it.

X takes on all positive integer values (I may never win the lottery).

 $\Omega_X=\{1,2,\dots\}$ 

Identify the support/range  $\Omega_X$  of the random variable X, if X is... c) The number of heads in n flips of a coin with  $0 < \mathbb{P}(\text{head}) < 1$ .

Min:

Max:

In between?

Identify the support/range  $\Omega_X$  of the random variable X, if X is...

c) The number of heads in *n* flips of a coin with  $0 < \mathbb{P}(\text{head}) < 1$ .

X takes on every integer value between the min number of heads 0, and the max n.

 $\Omega_X = \{1, 2, \dots, n\}$ 

Identify the support/range  $\Omega_X$  of the random variable X, if X is... d) The number of heads in n flips of a coin with  $\mathbb{P}(\text{head}) = 1$ .

Min:

Max:

In between?

Identify the support/range  $\Omega_X$  of the random variable X, if X is... d) The number of heads in n flips of a coin with  $\mathbb{P}(\text{head}) = 1$ .

Since  $\mathbb{P}(\text{head}) = 1$ , we are guaranteed to get *n* heads in *n* flips..

 $\Omega_X = \{n\}$ 

Identify the support/range  $\Omega_X$  of the random variable X, if X is...

e) The number of whole minutes I wait at the bus stop for the next bus.

Min:

Max:

In between?

Identify the support/range  $\Omega_X$  of the random variable X, if X is...

e) The number of whole minutes I wait at the bus stop for the next bus.

The number of whole minutes is discrete and will take on values between the minimum waiting time (0, the bus is here), and the maximum waiting time ( $\infty$ , the bus never gets here).

 $\Omega_X=\{0,1,\dots\}$ 

# Problem 4 – Kit Kats Again



We have *N* candies in the jar. We have *K* kit kats in the jar.

We are drawing without replacement until we have k kit kats.

 $k \leq K \leq N$ 

Let X be the number of draws until the kth kit kat (this includes the kth kit kat).

What is  $\Omega_X$ , the range of X? What is  $p_X(n) = \mathbb{P}(X = n)$ ?

Work on finding the range of *X* with the people around you!

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?

Min:

Max:

In between:

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?

Min: **k** (we were just really lucky and kept picking kit kats one after the other until we had all k that we wanted)

Max:

In between:

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?

Min: **k** (we were just really lucky and kept picking kit kats one after the other until we had all k that we wanted)

Max: **N** – **K** + **k** (ended up pick out all the non kit kats (N - K) in the process of picking out our k kit kats)

In between:

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?

Min: **k** (we were just really lucky and kept picking kit kats one after the other until we had all k that we wanted)

Max: **N** – **K** + **k** (ended up pick out all the non kit kats (N - K) in the process of picking out our k kit kats)

In between: Any integer value in between

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?

 $\Omega_X = \{k, k+1, \dots, N-K+k\}$ 

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?

Work on finding the PMF of X with the people around you!

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?

Consider all *N* candies to be arranged randomly in a row. This is the order in which we will be drawing our *n* candies.

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?

Consider all *N* candies to be arranged randomly in a row. This is the order in which we will be drawing our *n* candies.

We know that the *n*th candy is a kitkat (the *k*th kitkat) to be chosen. This *n*th candy divides the row into 2 sections.

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?

Left: (?) candies and (?) kit kats nth spot: kth kit kat Right: (?) candies and (?) kit kats

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?

Left: **n** - **1** candies and **k** - **1** kit kats nth spot: kth kit kat Right: (?) candies and (?) kit kats

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

```
What is \Omega_X, the range of X? \Omega_X = \{k, k + 1, ..., N - K + k\}
What is p_X(n) = \mathbb{P}(X = n)?
```

Left: **n** - **1** candies and **k** - **1** kit kats nth spot: kth kit kat Right: **N** - **n** candies and **K** - **k** kit kats Sample Space:

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?

Left: n - 1 candies and k - 1 kit katsEvent: number of ways to arrange the candies to have thisnth spot: kth kit katoutcomeRight: N - n candies and K - k kit katsSample Space: number of ways to arrange the candies

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the *k*th kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?

Left: **n** - **1** candies and **k** - **1** kit kats nth spot: kth kit kat  $|E|: \binom{n}{k}$ |S|:  $\binom{N}{K}$ 

$$|\mathsf{E}|: \binom{n-1}{k-1}\binom{N-n}{K-k}$$
$$|\mathsf{S}|: \binom{N}{K}$$

### 4 – Kit Kats Again

We have N candies in the jar. We have K kit kats in the jar. We are drawing without replacement until we have k kit kats.  $k \le K \le N$ 

Let *X* be the number of draws until the *k*th kit kat (this includes the last kit kat).

What is  $\Omega_X$ , the range of X?  $\Omega_X = \{k, k + 1, ..., N - K + k\}$ What is  $p_X(n) = \mathbb{P}(X = n)$ ?  $p_X(n) = \mathbb{P}(X = n) = \frac{\binom{n-1}{k-1}\binom{N-n}{K-k}}{\binom{N}{K}}$  if  $n \in \Omega_X$  and 0 otherwise

Justification: Consider all N candies to be arranged randomly in a row. We will treat the first n candies to be the "chosen" candies. We know that the nth candy is a kitkat (the kth kitkat) to be chosen. This nth candy divides the row into 2 sections. On the left, are the n - 1 candies that were chosen (all candies chosen except the last kitkat). On the right, are the candies remaining in the jar. So the first term in the numerator is choosing the k - 1 spots where kitkats will be chosen from the n - 1 spots. The second term is choosing the K - k spots for the kitkats that remain in the jar among the remaining N - n candies. The denominator is the total number of possible ways to place the K kitkats among N candies.

# **Problem 5 – Hungry Washing Machine**



You have 10 pairs of socks (so 20 socks in total), with each pair being a different color. You put them in the washing machine, but the washing machine eats 4 of the socks chosen at random. Every subset of 4 socks is equally probable to be the subset that gets eaten. Let *X* be the number of complete pairs of socks that you have left.

- a) What is the range of X,  $\Omega_X$  (the set of possible values it can take on)? What is the probability mass function of X?
- b) Find  $\mathbb{E}[X]$  from the definition of expectation.

Work on this problem with the people around you, and then we'll go over it together!

a) What is the range of X,  $\Omega_X$  (the set of possible values it can take on)? What is the probability mass function of X?

a) What is the range of X,  $\Omega_X$  (the set of possible values it can take on)? What is the probability mass function of X?

The washing machine eats 4 socks every time. It can either eat a single sock from 4 pairs of socks, leaving us with 6 complete pairs, or a single sock from 2 pairs and a matching pair, leaving us with 7 complete pairs, or 2 pairs of matching socks, leaving us with 8 complete pairs.

 $\Omega_X = \{6,7,8\}$ 

a) What is the range of X,  $\Omega_X$  (the set of possible values it can take on)? What is the probability mass function of X?

The washing machine eats 4 socks every time. It can either eat a single sock from 4 pairs of socks, leaving us with 6 complete pairs, or a single sock from 2 pairs and a matching pair, leaving us with 7 complete pairs, or 2 pairs of matching socks, leaving us with 8 complete pairs.

 $\Omega_X = \{6,7,8\}$ 

We are dealing with a sample space with equally likely outcomes. As such, we can compute use the formula  $\mathbb{P}(E) = \frac{|E|}{|\Omega|}$ . We know that  $|\Omega| = \binom{20}{4}$  because the washing machine picks a set of 4 socks out of 20 possible socks.

a) What is the range of X,  $\Omega_X$  (the set of possible values it can take on)? What is the probability mass function of X?

To define the pmf of *X*, we consider each value in the range of *X*.

For k = 6, we first pick 4 out of 10 pairs of socks from which we will eat a single sock  $\binom{10}{4}$  ways), and for each of these 4 pairs we have two socks to pick from  $\binom{2}{1}^4$  ways). Using the product rule, we get  $|X = 6| = \binom{10}{4} 2^4$ 

For k = 7, we first pick 1 out of 10 pairs of socks to eat in its entirety  $\binom{10}{1}$  ways), and then 2 out of the 9 remaining pairs from which we will eat a single sock  $\binom{9}{2}$  ways), and for each of these 2 pairs we have two socks to pick from  $\binom{2}{1}^2$  ways). Using the product rule, we get  $|X = 7| = 10\binom{9}{4}2^2$ 

For k = 8, we pick 2 out of 10 pairs of socks to eat  $\binom{10}{2}$  ways). We get  $|X = 8| = \binom{10}{2}$ 

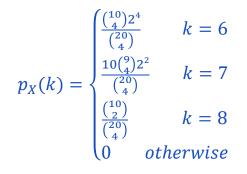
a) What is the range of X,  $\Omega_X$  (the set of possible values it can take on)? What is the probability mass function of X?

To define the pmf of X, we consider each value in the range of X.

For k = 6, we first pick 4 out of 10 pairs of socks from which we will eat a single sock  $\binom{10}{4}$  ways), and for each of these 4 pairs we have two socks to pick from  $\binom{2}{1}^4$  ways). Using the product rule, we get  $|X = 6| = \binom{10}{4} 2^4$ 

For k = 7, we first pick 1 out of 10 pairs of socks to eat in its entirety  $\binom{10}{1}$  ways), and then 2 out of the 9 remaining pairs from which we will eat a single sock  $\binom{9}{2}$  ways), and for each of these 2 pairs we have two socks to pick from  $\binom{2}{1}^2$  ways). Using the product rule, we get  $|X = 7| = 10\binom{9}{4}2^2$ 

For k = 8, we pick 2 out of 10 pairs of socks to eat  $\binom{10}{2}$  ways). We get  $|X = 8| = \binom{10}{2}$ 

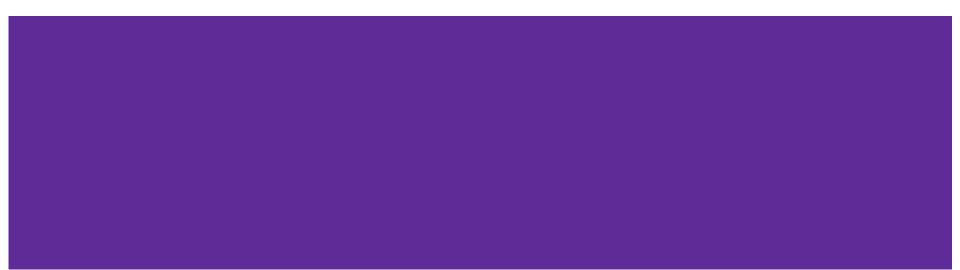


b) Find  $\mathbb{E}[X]$  from the definition of expectation.

b) Find  $\mathbb{E}[X]$  from the definition of expectation.

$$\mathbb{E}[X] = \sum_{k \in \Omega_X} k \cdot p_X(k) = 6 \cdot \frac{\binom{10}{4} 2^4}{\binom{20}{4}} + 7 \cdot \frac{10\binom{9}{2} 2^2}{\binom{20}{4}} + 8 \cdot \frac{\binom{10}{2}}{\binom{20}{4}} = \frac{120}{19}$$

# Problem 7 – Frogger



A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability  $p_1$ , to the left with probability  $p_2$ , and doesn't move with probability  $p_3$ , where  $p_1 + p_2 + p_3 = 1$ . After 2 seconds, let X be the location of the frog.

- a) Find  $p_X(k)$ , the probability mass function for X.
- b) Compute  $\mathbb{E}[X]$  from the definition.

Work on this problem with the people around you, and then we'll go over it together!

A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability  $p_1$ , to the left with probability  $p_2$ , and doesn't move with probability  $p_3$ , where  $p_1 + p_2 + p_3 = 1$ . After 2 seconds, let X be the location of the frog.

a) Find  $p_X(k)$ , the probability mass function for X.

A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability  $p_1$ , to the left with probability  $p_2$ , and doesn't move with probability  $p_3$ , where  $p_1 + p_2 + p_3 = 1$ . After 2 seconds, let X be the location of the frog.

a) Find  $p_X(k)$ , the probability mass function for X.

Let *L* be a left step, *R* be a right step, and *N* be no step.

The range of *X* is  $\{-2, -1, 0, 1, 2\}$ . We can compute:  $p_X(-2) = \mathbb{P}(X = -2) = \mathbb{P}(LL) = p_2^2$   $p_X(-1) = \mathbb{P}(X = -1) = \mathbb{P}(LN \cup NL) = 2p_2p_3$   $p_X(0) = \mathbb{P}(X = 0) = \mathbb{P}(NN \cup LR \cup RL) = p_3^2 + 2p_1p_2$   $p_X(1) = \mathbb{P}(X = 1) = \mathbb{P}(RN \cup NR) = 2p_1p_3$  $p_X(2) = \mathbb{P}(X = 2) = \mathbb{P}(RR) = p_1^2$ 

A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability  $p_1$ , to the left with probability  $p_2$ , and doesn't move with probability  $p_3$ , where  $p_1 + p_2 + p_3 = 1$ . After 2 seconds, let X be the location of the frog.

a) Find  $p_X(k)$ , the probability mass function for X.

Let *L* be a left step, *R* be a right step, and *N* be no step.

The range of *X* is  $\{-2, -1, 0, 1, 2\}$ . We can compute:  $p_X(-2) = \mathbb{P}(X = -2) = \mathbb{P}(LL) = p_2^2$   $p_X(-1) = \mathbb{P}(X = -1) = \mathbb{P}(LN \cup NL) = 2p_2p_3$   $p_X(0) = \mathbb{P}(X = 0) = \mathbb{P}(NN \cup LR \cup RL) = p_3^2 + 2p_1p_2$   $p_X(1) = \mathbb{P}(X = 1) = \mathbb{P}(RN \cup NR) = 2p_1p_3$  $p_X(2) = \mathbb{P}(X = 2) = \mathbb{P}(RR) = p_1^2$ 

 $p_X = \begin{cases} p_2^2 & k = -2 \\ 2p_2p_3 & k = -1 \\ p_3^2 + 2p_1p_2 & k = 0 \\ 2p_1p_3 & k = 1 \\ p_1^2 & k = 2 \\ 0 & otherwise \end{cases}$ 

A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability  $p_1$ , to the left with probability  $p_2$ , and doesn't move with probability  $p_3$ , where  $p_1 + p_2 + p_3 = 1$ . After 2 seconds, let X be the location of the frog.

b) Compute  $\mathbb{E}[X]$  from the definition.

A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability  $p_1$ , to the left with probability  $p_2$ , and doesn't move with probability  $p_3$ , where  $p_1 + p_2 + p_3 = 1$ . After 2 seconds, let X be the location of the frog.

b) Compute  $\mathbb{E}[X]$  from the definition.

 $\mathbb{E}[X] = (-2)(p_2^2) + (-1)(2p_2p_3) + (0)(p_3^2 + 2p_1p_2) + (1)(2p_1p_3) + (2)(p_1^2) = 2(p_1 - p_2)$ 

# That's All, Folks!

Thanks for coming to section this week! Any questions?