SDS 321

Lecture 8

Discrete Random Variables

What is a Random Variable?

- What is a variable? (how is it different from a constant?)
- A random variable is a variable defined in the context of a random experiment.

- Example 1: Experiment: Toss a coin 3 times
  Random Variable: X = number of heads in 3 coin tosses

- Example 2: Experiment: Choose a student at random from this class
  Random Variable: Y = the student's height
  Random Variable: X = number of siblings the student has
  Random Variable: Z = student's pet ownership status (Y/N or 1/0)

More Examples:

- X = outcome of a coin toss
- Y = the sum of two rolls of a die
- X = the number of die rolls it takes to get a six.

Discrete and Continuous Random Variables

- A random variable is discrete if its range (the values it can take) is finite or at most countably infinite.

  - X = sum of two rolls of a die. $X \in \{2,...,12\}$.
  - X = number of heads in 100 coin tosses. $X \in \{0,...,100\}$
  - X = number of coin tosses to get a head. $X \in \{1,2,3,...\}$

- A random variable is continuous if its range is uncountably infinite. A continuous random variable often is a result of a measurement, and often (but not always) has units.
- Example: height, blood pressure, amount of rainfall on a day, literacy percentage in a country

- We can create discrete random variables from continuous random variables.
- Example?

Discrete Random Variables

Probability Mass Function (PMF)
and
Common Discrete Probability Distributions

Probability Mass Function (PMF)

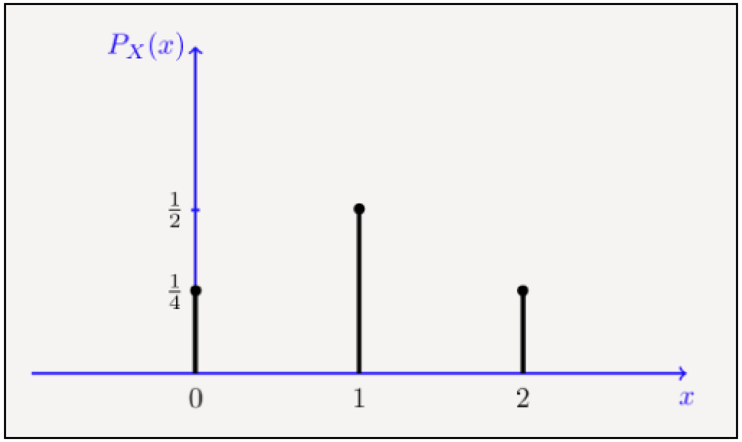
- A Probability Mass Function (PMF) assigns probabilities to the different values that a discrete random variable can take.

- A PMF can be presented in the form of a table (not always possible), or can be expressed as a function (not always possible).

- Example (Table form):
  X = number of heads in 2 fair coin tosses
  X can take values in {0, 1, 2}.

  P(X = 0) = P({TT}) = 1/4.
  P(X = 1) = P({HT,TH}) = 1/2.
  P(X = 2) = P({HH}) = 1/4.

X	**0**	**1**	**2**
P(X = x)	1/4	1/2	1/4



Probability Mass Function (PMF)

- Example (Function form):

  Sometimes the table form is not possible because the number of values a random variable can take is too many to list in table form.

  X = number of rolls of a fair die until I get the first 6
  X can take values = ?

  PMF is $P(X = x) = p*(1-p)^{x-1}$, where X = 1, 2, 3, …
  where X is the total number of rolls and p=probability of getting a 6.

- In general, the probability that a random variable X takes a value x is written as $p_X(x)$, or $P_X(x)$ or $P_X(X = x)$ etc.

- A random variable is always written in upper-case and the numerical value we are trying to evaluate the probability for is written in lower case.

Properties of PMF's

1) $P(X = x) \geq 0$ for all values of x

2) $$\sum_{x} P(X = x) = 1$$

$$P(X \in S) = \sum_{x \in S} P(X = x)$$

To compute $P(X = x)$

- Collect all possible outcomes that give $\{X = x\}$.
- Add their probabilities to get $P(X = x)$.

Example:

You toss two fair coins. What is the probability that you see at least 1 Head?

If X = # Heads,

We want $P(X \geq 1) = P(HT) + P(TH) + P(HH) = 3/4$

The Uniform Distribution

Consider the roll of a fair die. You are interested in X = number on the roll.

- X can take how many different values? {1, 2, 3, 4, 5, 6}
- What are the probabilities of taking on those values?

$$P(X=1)=P(X=2)=P(X=3)=P(X=4)=P(X=5)=P(X=6) = 1/6$$

- For a uniform random variable X, each value that X can take has equal probability mass.
- If a uniform random variable X takes on k different values, then the probability mass at each of those values is 1/k.
- So in this example, X is a Uniform Random Variable.

The Bernoulli Distribution

Consider the toss of a biased coin, which gives a head with probability p.

- A **Bernoulli** random variable X takes two values: 1 if a head comes up and 0 if not. $X = \begin{cases} 1 & \text{If head} \\ 0 & \text{If tail} \end{cases}$
- The PMF is given by: $p_X(x) = \begin{cases} p & \text{If } x = 1 \\ 1 - p & \text{If } x = 0 \end{cases}$
- Examples of a Bernoulli
  - A person can be healthy or sick with a certain disease.
  - A test result for a disease can be positive or negative.
  - It may rain one day or not.

If X ~ Ber(p), read as Bernoulli with parameter “p”
then P(X = 1) = p, and P(X = 0) = 1-p

The Binomial Distribution

- We can use independent Bernoulli RVs to create a Binomial RV.
- Let $X_{1,}$ X_2, … X_n be independent Bernoulli RVs with parameter p.
- Let Y = the sum $X_1 + X_2 + \ldots + X_n$
- Then Y is called a Binomial random variable, with parameters n and p.
- $Y \sim Bin(n, p)$

Example:
Let $X_{1,}$ X_2, … X_n be the outcomes of "n" independent tosses of a coin with P(X=1) = p, and P(X=0) = 1-p, where Heads=1 and Tails=0.

- So $X_1 \sim Ber(p)$, $X_2 \sim Ber(p)$, … , $X_n \sim Ber(p)$
- $Y = X_1 + X_2 + \ldots + X_n$ is the total # heads in "n" tosses of this coin.
- $Y \sim Bin(n, p)$.
- What is its PMF? Let's first look at some examples.

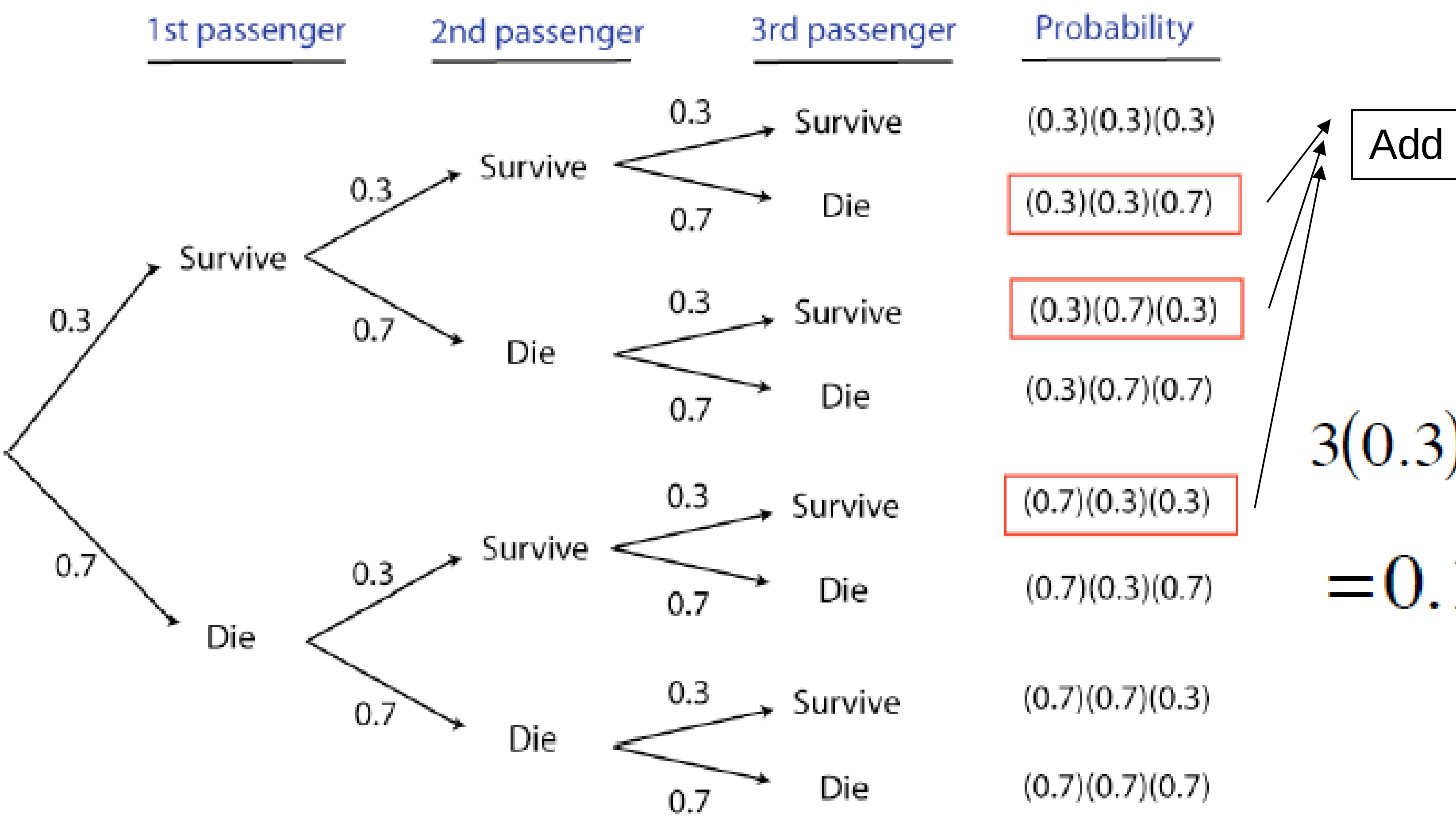
Binomial Distribution: Example 1

Toss a fair coin 10 times. Let Head=1 and Tail=0. What is the probability that the sum of outcomes on the 10 tosses = 6?

- So we need 6 Heads and 4 Tails.
- Number of ways of getting 6 Heads in 10 coin tosses is C(10,6)
- Probability of 6 Heads and 4 Tails = $(1/2)^{10}$
- So if Y= the sum, then
- $P(Y = 6) = C(10,6) \times (1/2)^{10}$
- What if the coin is biased, with P(Head) = 0.8?
- $P(Y = 6) = C(10,6) \times (0.8)^6(0.2)^4$

Binomial Distribution: Example 2

Suppose the probability of survival for every passenger on the Titanic was 0.30. What is the probability that 2 out of 3 randomly chosen passengers survived?



$$3(0.3)^2(0.7)$$

$$=0.189$$

The Binomial Distribution

Creating the Binomial Formula:

Probability that two out of three randomly chosen passengers survived the Titanic

Suppose Y = number of survivors out of 3 passengers

$$P(Y = 2) = \binom{3}{2}(0.3)^2(1-0.3)^{3-2}$$

- $\binom{3}{2}$: Number of ways to get 2 survivors out of 3 passengers
- $(0.3)^2$: Probability of 2 survivors
- $(1-0.3)^{3-2}$: Probability of 1 death

$$= \frac{3!}{2! \times 1!}(0.3)^2(0.7)^1$$

$$= 3(0.3)^2(0.7) = 0.189$$

The Binomial Distribution

So in general, the PMF of the Binomial random variable Y with parameters n and p is:

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ for } k \in \{0, 1, 2, \ldots, n\}$$

Is this a valid PMF?

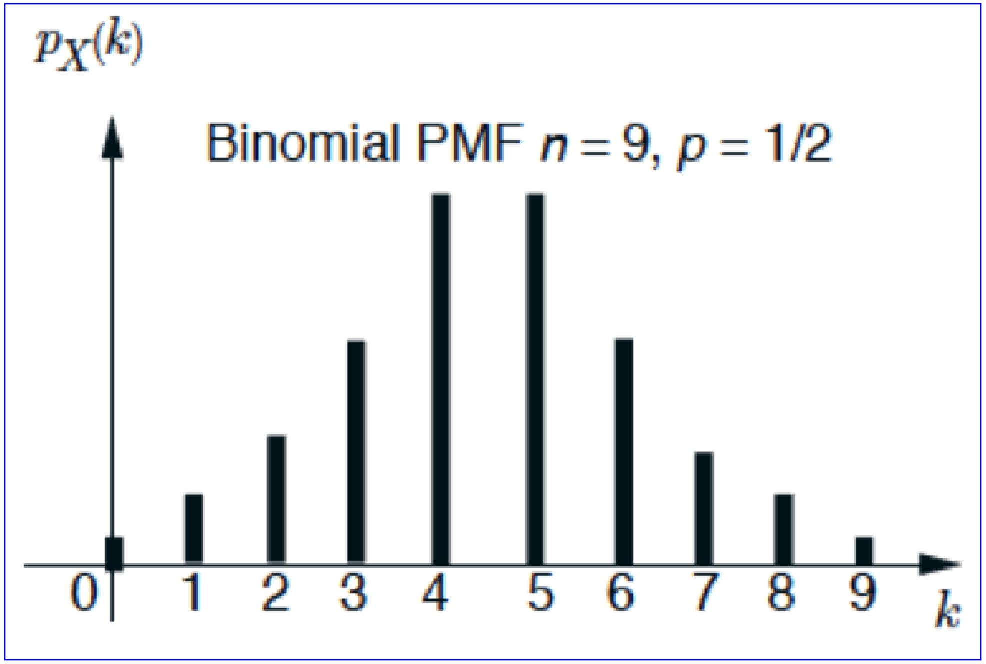
We can check to see if the probabilities over the range of Y values sum to 1.

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1 \text{ (Binomial expansion)!}$$

So Yes, this is a valid PMF.

The Binomial Distribution: Summary

- Sum of independent Bernoullis give a Binomial!
- We will denote the Binomial PMF by $Binomial(n, p)$ or $Bin(n, p)$.
- We will write $X \sim Binomial(n, p)$ to indicate that X is **distributed** as a Binomial random variable with parameters n and p.
- Quick look at the histogram of $X \sim Binomial(9, 1/2)$.



When p=1/2, the PMF is symmetric around n/2.

It's skewed right if p<1/2

It's skewed left if p>1/2

Binomial Distribution: Example

You have an urn with 10 blue and 20 red balls. You pick 9 balls at random with replacement. Let X be the number of blue balls.

(1) What is the probability that you picked exactly 3 blue balls?

(2) What is the probability that you picked at least 2 blue balls?

The Geometric Distribution

- The Bernoulli PMF describes probability of success/failure in a single trial.
- The Binomial PMF describes the probability of k successes out of n trials.
- Sometimes we may also be interested in doing trials until we see a success.
- Example: Adam resolves to keep buying lottery tickets until he wins a hundred million dollars. He is interested in the random variable “number of lottery tickets bought until he wins the 100M$ lottery”.
- Example: Annie is trying to catch a taxi. How many occupied taxis will drive past before she finds one that takes passengers?
- If X = the # of trials required to get the first success, with p = P(success)
- Then X is a Geometric Random Variable, with parameter p, or X ~ Geo(p)
- Note that all trials are independent, with identical Ber(p) distributions.

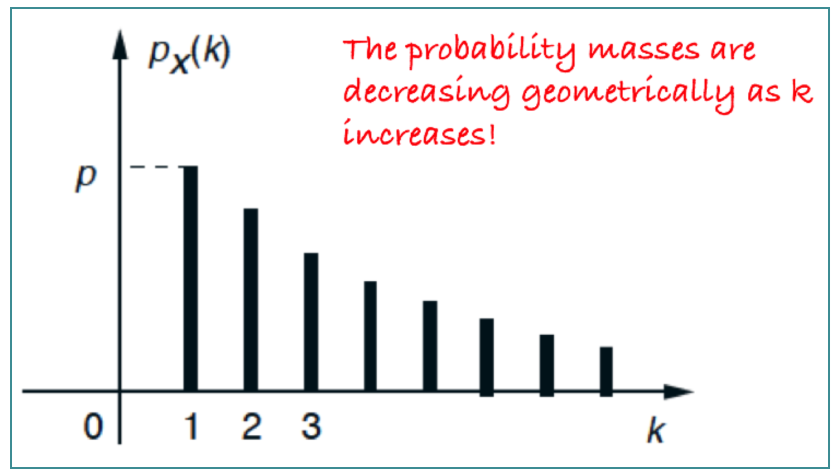
The Geometric Distribution

<u>Example</u>: We repeatedly toss a coin with P(Head) = p. The geometric random variable X is the number of tosses required to get the first head.

Its PMF is given by:

$$P(X = k) = P(\{\underbrace{TT\ldots T}_{k-1}H\}) = (1-p)^{k-1}p.$$

where k = 1, 2, …



This is a legitimate PMF because:

$$\sum_{k=1}^{\infty} P(X=k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p$$

$$= p\sum_{k=0}^{\infty} (1-p)^{k} = p\frac{1}{1-(1-p)} = 1$$

Recall: Geometric Series:

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$$

The Geometric Distribution

Example:

Annie is trying to catch a taxi. Suppose the probability that any taxi is free is 0.4. What is the probability that 20 occupied taxis will drive past before she finds one that takes passengers?

Write the probability statement and then calculate your answer.

The Geometric Distribution

- Suppose X is a Geometric random variable. What is the probability that the number of trials required to get a single success is at least k?
- i.e., What is $P(X \geq k)$?
- $P(X \geq k) =$ $$\sum_{i=k}^{\infty} p(1-p)^{i-1} = p\left[(1-p)^{k-1} + (1-p)^{k} + (1-p)^{k+1} + \ldots\right]$$
$$= p(1-p)^{k-1}\left[1 + (1-p) + (1-p)^{2} + (1-p)^{3} + \ldots\right]$$
$$= p(1-p)^{k-1}\left[\frac{1}{1-(1-p)}\right] = (1-p)^{k-1}$$
- Intuitively, this is the probability that the first (k-1) tosses are tails!
- Similarly, $P(X > k) = (1-p)^k$
- This leads us to the Memoryless Property of Geometric random variables.

The Geometric Random Variable

Memoryless Property

Suppose we are told that there are “a” failures initially. The chance of “b” additional trials until we see the first success is exactly the same as if we started the experiment for the first time after the string of “a” failures, and the information of “a” initial failures is not given to us at all.

Example: Gambling in a casino

What is P(X = a+b | X > a) ?

$$P(X = a + b | X > a) = \frac{P(X = a + b)}{P(X > a)}$$

$$= \frac{p(1-p)^{a+b-1}}{(1-p)^a}$$

$$= p(1-p)^{b-1} = P(X = b)$$

The Geometric Random Variable

Memoryless Property

So now we know that, by the memoryless property,

$P(X = a+b \mid X > a) = P(X = b)$

Now note that $P(X > a+b \mid X > a) = P(X > b)$

The *only* memoryless discrete probability distributions are the geometric distributions, which feature the number of independent Bernoulli trials needed to get the first “success,” with a fixed probability of success p on each trial.

These are the distributions of waiting time in a Bernoulli process.

The Geometric Random Variable

Example:

Annie is trying to catch a taxi. Suppose the probability that a taxi is free is 0.4. Suppose we know that the first 5 taxis that pass by are occupied. What is the probability that

(a) Exactly 20 occupied taxis will drive past before she finds one that takes passengers?

(b) More than 20 occupied taxis will drive past before she finds one that takes passengers?

(c) At least 20 occupied taxis will drive past before she finds one that takes passengers?

Write the probability statement and then calculate your answer.

The Poisson Distribution

Example:

I have a book with 10000 words. Probability that a word has a typo is 1/1000. I am interested in how many misprints can be there on average? So a Poisson distribution often applies when you have a Binomial random variable with very large n and very small p but $n \times p$ is moderate. Here $np = 10$.

Our random variable might be:

- The number of car crashes in a given day.

- The number of yellow cars on the road in a given time period.

- The number of mutations on a strand of DNA.

We can describe such situations using a Poisson random variable.

The Poisson Distribution

A Poisson random variable takes non-negative integers as values.

It has a nonnegative parameter λ.

Its PMF is given by:

$$P(X=k)=\frac{e^{-\lambda}\lambda^{k}}{k!}$$

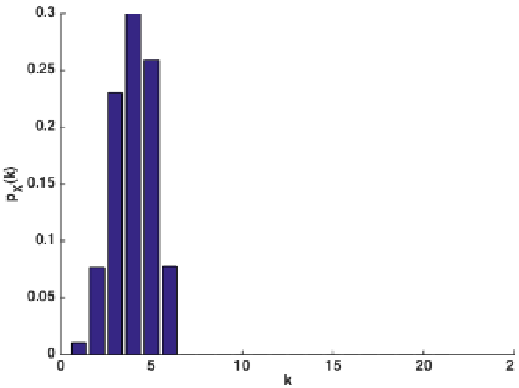
For k = 0, 1, 2, …

This is a legitimate PMF because:

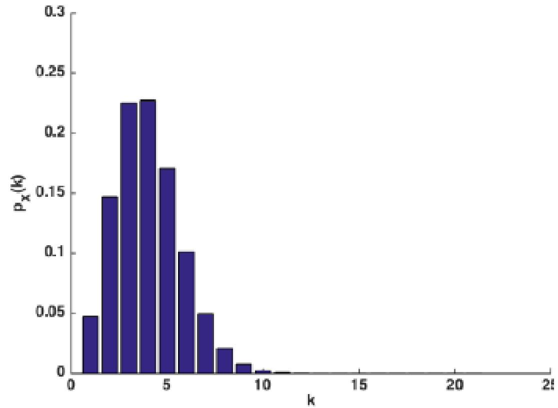
$$\sum_{k=0}^{\infty} P(X=k) = e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\ldots\right)=1$$

(The exponential series : e^{λ})

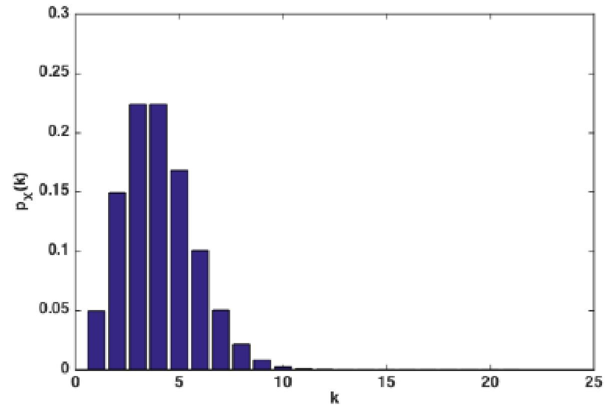
The Poisson Distribution



Binomial(5,0.6)



Binomial(100,0.03)



Poisson(3)

- When n is very large and p is very small, a binomial random variable can be well approximated by a Poisson with $\lambda = np$.
- In the above figure we increased n and decreased p so that $np = 3$.
- See how close the PMF's of the Binomial(100,0.03) and Poisson(3) are!
- More formally, we see that $\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{e^{-\lambda}\lambda^k}{k!}$ when n is large, k is fixed, and p is small and $\lambda = np$.

The Poisson Distribution

Example:

A baker put 500 raisins into dough, mixed well, and made 100 cookies. You take a random cookie. What is the probability of finding exactly 4 raisins in it?