Probability

3.1 Discrete Random Variables Basics

Alex Tsun

Agenda

- Intro to Discrete Random Variables
- Probability Mass Functions
- Expectation

Flipping two coins

Suppose you flip a fair coin twice. What is the sample space Ω?

Flipping two coins

Suppose you flip a fair coin twice. What is the sample space Ω?

$$\Omega = \{HH, HT, TH, TT\}$$

Flipping two coins

Suppose you flip a fair coin twice. What is the sample space Ω?

$$\Omega = \{HH, HT, TH, TT\}$$

Usually, we don't care about the exact outcome (HT vs TH), but just the fact we got exactly one head.

Flipping two coins

Suppose you flip a fair coin twice. What is the sample space Ω?

$$\Omega = \{HH, HT, TH, TT\}$$

Usually, we don't care about the exact outcome (HT vs TH), but just the fact we got exactly one head. So we define a **random variable** as a *numeric* function of the outcome. What does this mean?

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Let X be the number of heads in two independent flips of a fair coin.

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Let X be the number of heads in two independent flips of a fair coin.

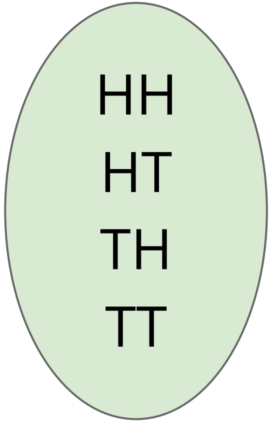
X is a function, $X: \Omega \to \mathbb{R}$ which takes outcomes $\omega \in \Omega$ and maps them to a number.

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Let X be the number of heads in two independent flips of a fair coin.

X is a function, $X: \Omega \to \mathbb{R}$ which takes outcomes $\omega \in \Omega$ and maps them to a number.



Ω

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Let X be the number of heads in two independent flips of a fair coin.

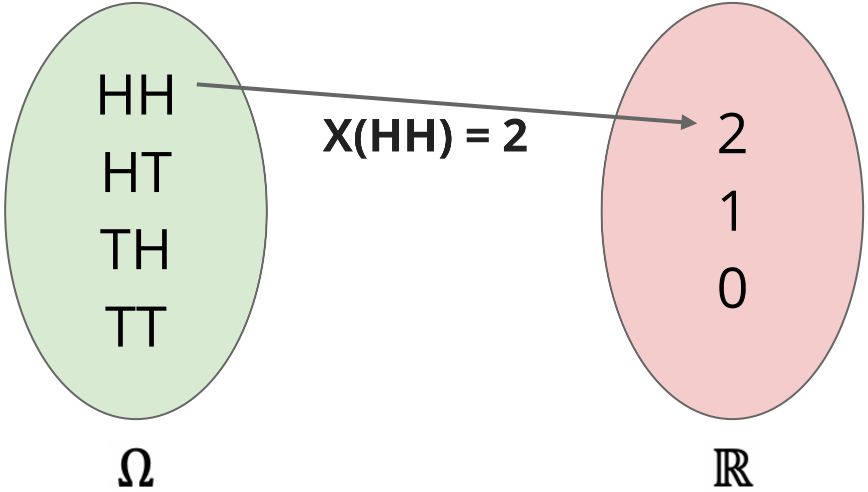
X is a function, $X: \Omega \to \mathbb{R}$ which takes outcomes $\omega \in \Omega$ and maps them to a number.

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Let X be the number of heads in two independent flips of a fair coin.

X is a function, $X : \Omega \to \mathbb{R}$ which takes outcomes $\omega \in \Omega$ and maps them to a number.

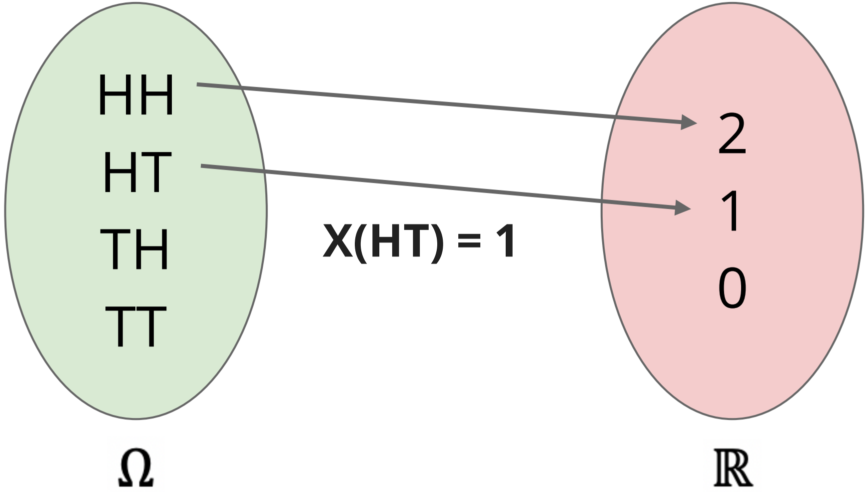


Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Let X be the number of heads in two independent flips of a fair coin.

X is a function, $X: \Omega \to \mathbb{R}$ which takes outcomes $\omega \in \Omega$ and maps them to a number.

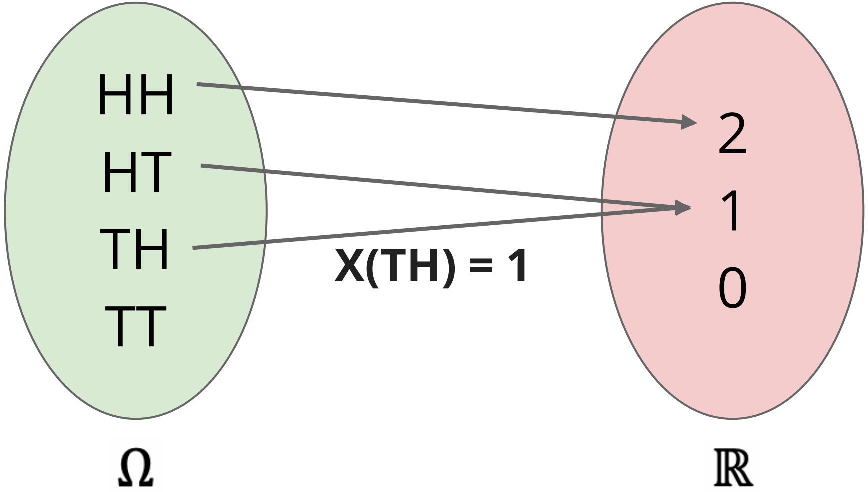


Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Let X be the number of heads in two independent flips of a fair coin.

X is a function, $X: \Omega \to \mathbb{R}$ which takes outcomes $\omega \in \Omega$ and maps them to a number.

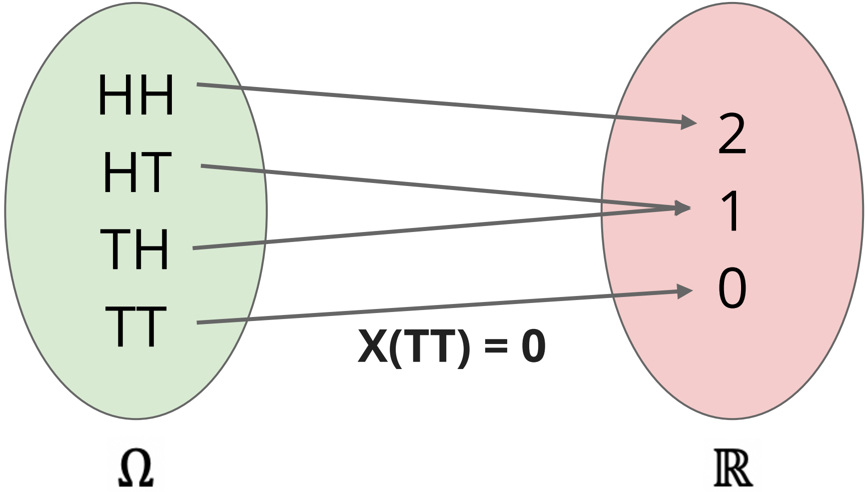


Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

Let X be the number of heads in two independent flips of a fair coin.

X is a function, $X: \Omega \to \mathbb{R}$ which takes outcomes $\omega \in \Omega$ and maps them to a number.



Random Variable

Suppose we conduct an experiment with sample space Ω. A **random variable (rv)** is a numeric function of the outcome, $X: \Omega \to \mathbb{R}$. That is, it maps outcomes $\omega \in \Omega$ to numbers, $\omega \mapsto X(\omega)$.

Random Variable

Suppose we conduct an experiment with sample space Ω. A **random variable (rv)** is a numeric function of the outcome, $X: \Omega \to \mathbb{R}$. That is, it maps outcomes $\omega \in \Omega$ to numbers, $\omega \mapsto X(\omega)$.

The set of possible values X can take on is its **range/support**, denoted Ω_X.

Random Variable

Suppose we conduct an experiment with sample space Ω. A **random variable (rv)** is a numeric function of the outcome, $X: \Omega \to \mathbb{R}$. That is, it maps outcomes $\omega \in \Omega$ to numbers, $\omega \mapsto X(\omega)$.

The set of possible values X can take on is its **range/support**, denoted Ω_X.

If Ω_X is finite or countably infinite (typically integers or a subset), X is a **discrete** random variable (drv). Else if Ω_X is uncountably large (the size of real numbers), X is **continuous** random variable.

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.		

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.		

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv
The number of flips of a fair coin up to and including my first head.		

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv
The number of flips of a fair coin up to and including my first head.	$\{1,2,\ldots\}$	

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\dots,n\}$	drv
The number of people born this year.	$\{0,1,2,\dots\}$	drv
The number of flips of a fair coin up to and including my first head.	$\{1,2,\dots\}$	drv

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv
The number of flips of a fair coin up to and including my first head.	$\{1,2,\ldots\}$	drv
The amount of time I wait for the next bus in seconds.		

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv
The number of flips of a fair coin up to and including my first head.	$\{1,2,\ldots\}$	drv
The amount of time I wait for the next bus in seconds.	$[0,\infty)$	

Identify those RVs

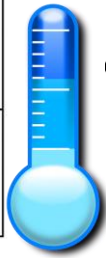
For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv
The number of flips of a fair coin up to and including my first head.	$\{1,2,\ldots\}$	drv
The amount of time I wait for the next bus in seconds.	$[0,\infty)$	crv

Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

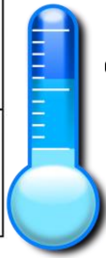
RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv
The number of flips of a fair coin up to and including my first head.	$\{1,2,\ldots\}$	drv
The amount of time I wait for the next bus in seconds.	$[0,\infty)$	crv
The temperature in Celsius of liquid water.		



Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

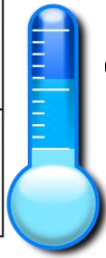
RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv
The number of flips of a fair coin up to and including my first head.	$\{1,2,\ldots\}$	drv
The amount of time I wait for the next bus in seconds.	$[0,\infty)$	crv
The temperature in Celsius of liquid water.	$(0,100)$	



Identify those RVs

For each of the following random variables, identify its range Ω_X and whether it is discrete or continuous.

RV Description	Range	drv or crv?
The number of heads in n flips of a fair coin.	$\{0,1,\ldots,n\}$	drv
The number of people born this year.	$\{0,1,2,\ldots\}$	drv
The number of flips of a fair coin up to and including my first head.	$\{1,2,\ldots\}$	drv
The amount of time I wait for the next bus in seconds.	$[0,\infty)$	crv
The temperature in Celsius of liquid water.	$(0,100)$	crv



Random Picture

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

$$X(HH) = 2 \qquad X(HT) = 1 \qquad X(TH) = 1 \qquad X(TT) = 0$$

Flipping two coins


$$\Omega = \{HH, HT, TH, TT\}$$

$$X(HH) = 2 \qquad X(HT) = 1 \qquad X(TH) = 1 \qquad X(TT) = 0$$

What is the support/range Ω_X?

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

$$X(HH) = 2 \qquad X(HT) = 1 \qquad X(TH) = 1 \qquad X(TT) = 0$$

What is the support/range Ω_X? $\boldsymbol{\Omega_X = \{0, 1, 2\}}$

Flipping two coins


$$\Omega = \{HH, HT, TH, TT\}$$

$$X(HH) = 2 \qquad X(HT) = 1 \qquad X(TH) = 1 \qquad X(TT) = 0$$

What is the support/range Ω_X? $\boldsymbol{\Omega_X = \{0, 1, 2\}}$

But what are the probabilities X takes on these values? For this, we define the **probability mass function (pmf)** of X, as $p_X \colon \Omega_X \to [0,1]$

$$p_X(k) = P(X = k)$$

Flipping two coins

$$\Omega = \{HH, HT, TH, TT\}$$

$$X(HH) = 2 \qquad X(HT) = 1 \qquad X(TH) = 1 \qquad X(TT) = 0$$

What is the support/range Ω_X? $\mathbf{\Omega_X = \{0, 1, 2\}}$

But what are the probabilities X takes on these values? For this, we define the **probability mass function (pmf)** of X, as $p_X: \Omega_X \to [0,1]$

$$p_X(k) = P(X = k)$$

$$p_X(k) = \begin{cases} & k = 0 \\ & k = 1 \\ & k = 2 \end{cases}$$

Flipping two coins


$$\Omega = \{HH, HT, TH, TT\}$$

$$X(HH) = 2 \qquad X(HT) = 1 \qquad X(TH) = 1 \qquad X(TT) = 0$$

What is the support/range Ω_X? $\mathbf{\Omega_X = \{0, 1, 2\}}$

But what are the probabilities X takes on these values? For this, we define the **probability mass function (pmf)** of X, as $p_X: \Omega_X \to [0,1]$

$$p_X(k) = P(X = k)$$

$$p_X(k) = \begin{cases} 1/4, & k = 0 \\ & k = 1 \\ & k = 2 \end{cases}$$

Flipping two coins


$$\Omega = \{HH, HT, TH, TT\}$$

$$X(HH) = 2 \qquad X(HT) = 1 \qquad X(TH) = 1 \qquad X(TT) = 0$$

What is the support/range Ω_X? $\mathbf{\Omega_X = \{0, 1, 2\}}$

But what are the probabilities X takes on these values? For this, we define the **probability mass function (pmf)** of X, as $p_X: \Omega_X \to [0,1]$

$$p_X(k) = P(X = k)$$

$$p_X(k) = \begin{cases} 1/4, & k = 0 \\ 1/2, & k = 1 \\ & k = 2 \end{cases}$$

Flipping two coins


$$\Omega = \{HH, HT, TH, TT\}$$

$$X(HH) = 2 \qquad X(HT) = 1 \qquad X(TH) = 1 \qquad X(TT) = 0$$

What is the support/range Ω_X? $\mathbf{\Omega_X = \{0, 1, 2\}}$

But what are the probabilities X takes on these values? For this, we define the **probability mass function (pmf)** of X, as $p_X\colon \Omega_X \to [0,1]$

$$p_X(k) = P(X = k)$$

$$p_X(k) = \begin{cases} 1/4, & k = 0 \\ 1/2, & k = 1 \\ 1/4, & k = 2 \end{cases}$$

Probability Mass Function (PMF)

The **probability mass function (pmf)** of a discrete random variable X assigns probabilities to the possible values of the random variable. That is, $p_X: \Omega_X \to [0,1]$ where

$$p_X(k) = P(X = k)$$

Probability Mass Function (PMF)

The <u>**probability mass function (pmf)**</u> of a discrete random variable X assigns probabilities to the possible values of the random variable. That is, $p_X\colon \Omega_X \to [0,1]$ where

$$p_X(k) = P(X = k)$$

Note that $\{X = a\}$ for $a \in \Omega_X$ form a partition of Ω, since each outcome $\omega \in \Omega$ is mapped to exactly one number.

Probability Mass Function (PMF)

The **probability mass function (pmf)** of a discrete random variable X assigns probabilities to the possible values of the random variable. That is, $p_X : \Omega_X \to [0,1]$ where

$$p_X(k) = P(X = k)$$

Note that $\{X = a\}$ for $a \in \Omega_X$ form a partition of Ω, since each outcome $\omega \in \Omega$ is mapped to exactly one number. Hence,

$$\sum_{z \in \Omega_X} p_X(z) = 1$$

Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin?

Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin? 1

Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin? **1**

Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.


Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin? **1**

Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once.

Flipping two coins


What is the *expected* number of heads in 2 flips of a fair coin? **1**
Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once. We would divide by the number of times (4), to get 1 on average.

Flipping two coins


What is the *expected* number of heads in 2 flips of a fair coin? **1**

Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once. We would divide by the number of times (4), to get 1 on average.

$$\frac{2+1+1+0}{4}$$

Flipping two coins


What is the *expected* number of heads in 2 flips of a fair coin? 1
Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once. We would divide by the number of times (4), to get 1 on average.

$$\frac{2+1+1+0}{4} = 2 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} = 1$$

Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin? 1
Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once. We would divide by the number of times (4), to get 1 on average.

$$\frac{2+1+1+0}{4} = \overset{X(HH)}{2} \cdot \frac{1}{4} + \overset{X(HT)}{1} \cdot \frac{1}{4} + \overset{X(TH)}{1} \cdot \frac{1}{4} + \overset{X(TT)}{0} \cdot \frac{1}{4} = 1$$

Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin? 1

Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once. We would divide by the number of times (4), to get 1 on average.

$$\frac{2+1+1+0}{4} = \overset{X(HH)}{2} \cdot \underset{P(HH)}{\frac{1}{4}} + \overset{X(HT)}{1} \cdot \underset{P(HT)}{\frac{1}{4}} + \overset{X(TH)}{1} \cdot \underset{P(TH)}{\frac{1}{4}} + \overset{X(TT)}{0} \cdot \underset{P(TT)}{\frac{1}{4}} = 1$$

Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin? 1

Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once. We would divide by the number of times (4), to get 1 on average.

$$\frac{2+1+1+0}{4} = \overset{X(HH)}{2} \cdot \underset{P(HH)}{\frac{1}{4}} + \overset{X(HT)}{1} \cdot \underset{P(HT)}{\frac{1}{4}} + \overset{X(TH)}{1} \cdot \underset{P(TH)}{\frac{1}{4}} + \overset{X(TT)}{0} \cdot \underset{P(TT)}{\frac{1}{4}} = 1 = \sum_{\omega\in\Omega} X(\omega)P(\omega)$$

Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin? **1**

Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once. We would divide by the number of times (4), to get 1 on average.

$$\frac{2+1+1+0}{4} = 2\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 0\cdot\frac{1}{4} = 1 = \sum_{\omega\in\Omega} X(\omega)P(\omega)$$

$$= 2\cdot\overset{p_X(2)}{\frac{1}{4}} + 1\cdot\overset{p_X(1)}{\frac{2}{4}} + 0\cdot\overset{p_X(0)}{\frac{1}{4}} = 1$$

Flipping two coins

What is the *expected* number of heads in 2 flips of a fair coin? 1

Since X is the # of heads in 2 flips of a fair coin, we denote this $E[X]$.

How did we get this number? Imagine we repeated this experiment 4 times. Then, we would expect to get HH, HT, TH, TT each once. We would divide by the number of times (4), to get 1 on average.

$$\frac{2+1+1+0}{4} = 2 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} = 1 = \sum_{\omega \in \Omega} X(\omega) P(\omega)$$

$p_X(2)$ $p_X(1)$ $p_X(0)$

$$= 2 \cdot \frac{1}{4} + 1 \cdot \frac{2}{4} + 0 \cdot \frac{1}{4} = 1 = \sum_{k \in \Omega_X} k \cdot p_X(k)$$

Expectation

The **<u>expectation/expected value/average</u>** of a discrete random variable X is

$$E[X] = \sum_{\omega \in \Omega} X(\omega) P(\omega)$$

Or equivalently,

$$E[X] = \sum_{k \in \Omega_X} k \cdot p_X(k)$$

The interpretation is that we take an average of the values in Ω_X, but weighted by their probabilities.

Probability

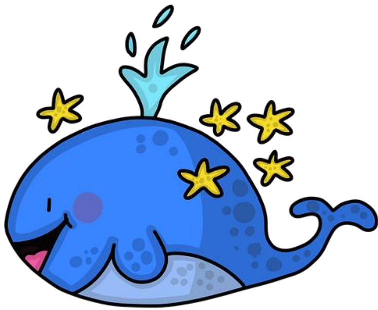
3.2 More on Expectation

Alex Tsun

Agenda

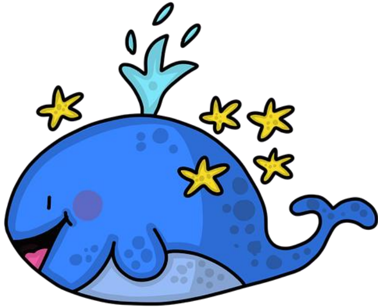
- Linearity of Expectation (LoE)
- Law of the Unconscious Statistician (Lotus)

Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

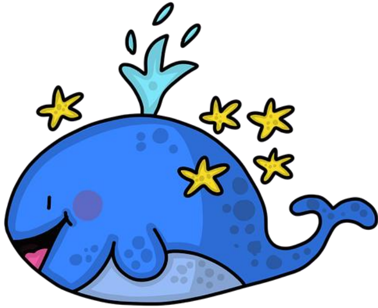
Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with E[X] = 3.

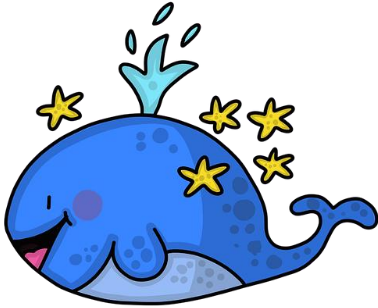
Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with $E[X] = 3$.
- Every day your friend catches Y fish, with $E[Y] = 7$.

Linearity of Expectation (Idea)



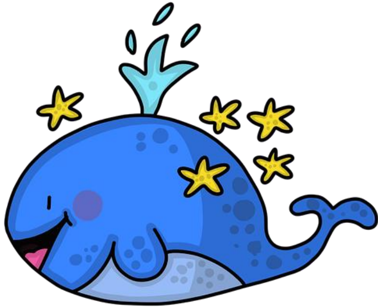
Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with E[X] = 3.
- Every day your friend catches Y fish, with E[Y] = 7.

How many fish do the two of you bring in (Z = X + Y) on an average day?

E[Z] = E[X + Y] =

Linearity of Expectation (Idea)



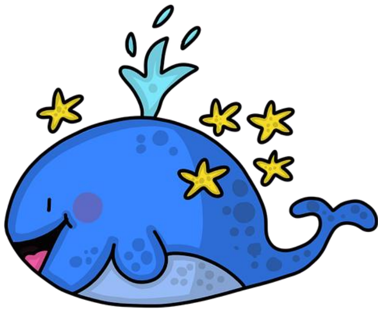
Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with $E[X] = 3$.
- Every day your friend catches Y fish, with $E[Y] = 7$.

How many fish do the two of you bring in ($Z = X + Y$) on an average day?

$E[Z] = E[X + Y] =$ $3 + 7 = 10$

Linearity of Expectation (Idea)



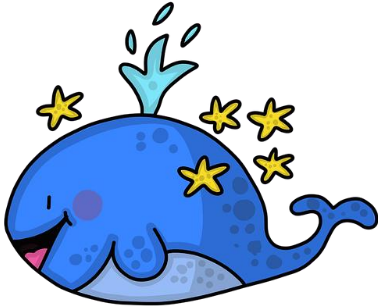
Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with $E[X] = 3$.
- Every day your friend catches Y fish, with $E[Y] = 7$.

How many fish do the two of you bring in ($Z = X + Y$) on an average day?

$$E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10$$

Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

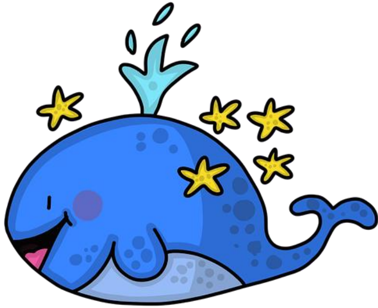
- Every day you catch X fish, with $E[X] = 3$.
- Every day your friend catches Y fish, with $E[Y] = 7$.

How many fish do the two of you bring in ($Z = X + Y$) on an average day?

$$E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10$$

You can sell each fish for \$5 at a store, but you need to pay \$20 in rent. How much profit do you expect to make? $E[5Z - 20] =$

Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

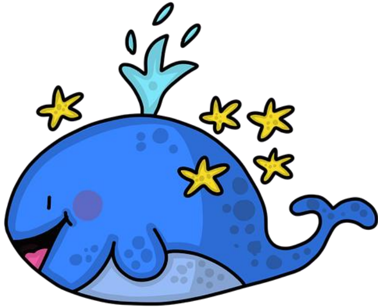
- Every day you catch X fish, with $E[X] = 3$.
- Every day your friend catches Y fish, with $E[Y] = 7$.

How many fish do the two of you bring in ($Z = X + Y$) on an average day?

$$E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10$$

You can sell each fish for \$5 at a store, but you need to pay \$20 in rent. How much profit do you expect to make? $E[5Z - 20] =$ $5 \times 10 - 20 = 30$

Linearity of Expectation (Idea)



Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with E[X] = 3.
- Every day your friend catches Y fish, with E[Y] = 7.

How many fish do the two of you bring in (Z = X + Y) on an average day?

$$E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10$$

You can sell each fish for \$5 at a store, but you need to pay \$20 in rent. How much profit do you expect to make? $E[5Z - 20] = 5E[Z] - 20 = 5 \times 10 - 20 = 30$

Linearity of Expectation (LoE)

Linearity of Expectation: Let Ω be the sample space of an experiment, $X, Y: \Omega \to \mathbb{R}$ be (possibly "dependent") random variables both defined on Ω, and $a, b, c \in \mathbb{R}$ be scalars. Then,

$$E[X + Y] = E[X] + E[Y]$$

and

$$E[aX + b] = aE[X] + b$$

Combining them gives,

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

Linearity of Expectation (Proof)

$$E[X] + E[Y]$$

Linearity of Expectation (Proof)

$$E[X] + E[Y] = \sum_{\omega \in \Omega} X(\omega)P(\omega) + \sum_{\omega \in \Omega} Y(\omega)P(\omega)$$

Linearity of Expectation (Proof)

$$E[X] + E[Y] = \sum_{\omega \in \Omega} X(\omega)P(\omega) + \sum_{\omega \in \Omega} Y(\omega)P(\omega)$$

$$= \sum_{\omega \in \Omega} \big(X(\omega) + Y(\omega)\big)P(\omega)$$

Linearity of Expectation (Proof)

$$E[X] + E[Y] = \sum_{\omega \in \Omega} X(\omega)P(\omega) + \sum_{\omega \in \Omega} Y(\omega)P(\omega)$$

$$= \sum_{\omega \in \Omega} \big(X(\omega) + Y(\omega)\big)P(\omega)$$

$$= \sum_{\omega \in \Omega} (X + Y)(\omega)\, P(\omega)$$

Linearity of Expectation (Proof)

$$E[X] + E[Y] = \sum_{\omega \in \Omega} X(\omega)P(\omega) + \sum_{\omega \in \Omega} Y(\omega)P(\omega)$$

$$= \sum_{\omega \in \Omega} \big(X(\omega) + Y(\omega)\big)P(\omega)$$

$$= \sum_{\omega \in \Omega} (X + Y)(\omega)\, P(\omega)$$

$$= E[X + Y]$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

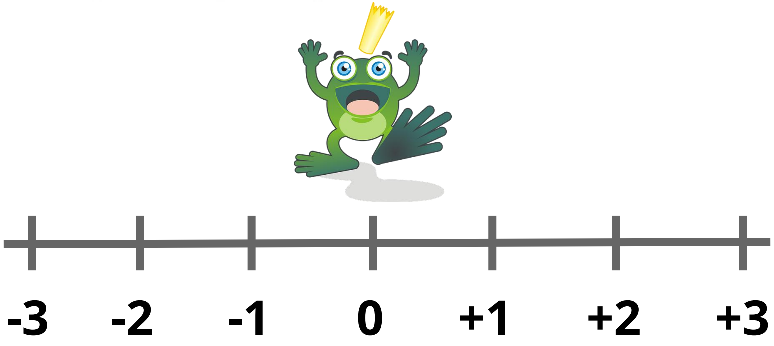
where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



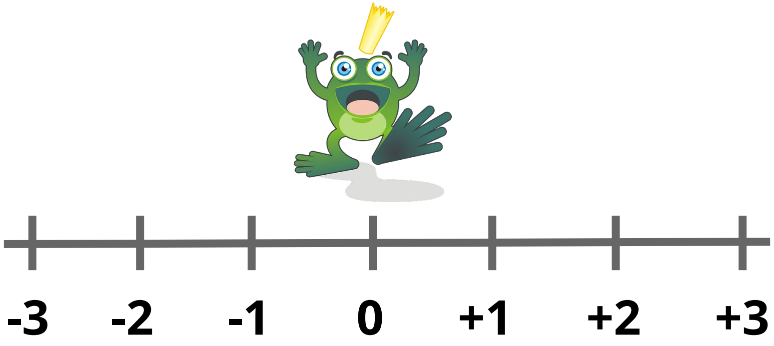
Frogger!

Brute Force

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?

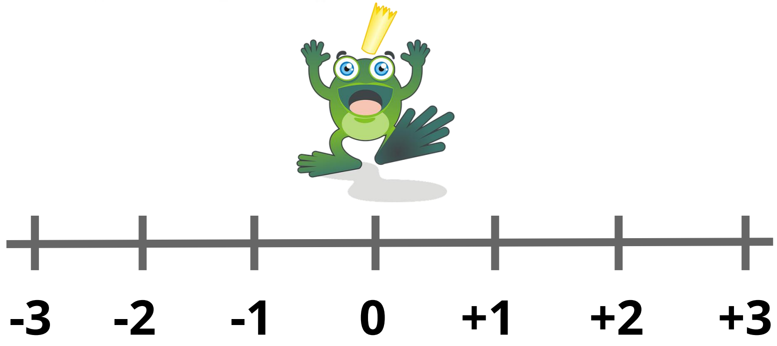


Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Brute Force

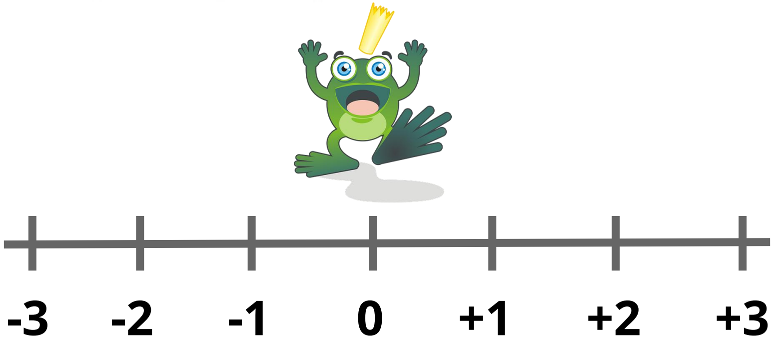
$$\Omega_X = \{-2, -1, 0, +1, +2\}$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Brute Force

$$\Omega_X = \{-2, -1, 0, +1, +2\}$$

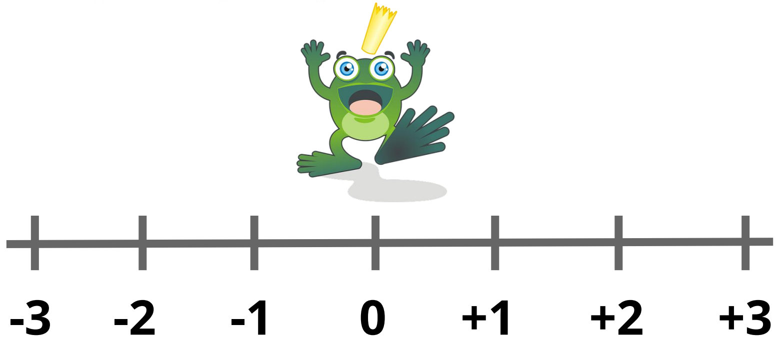
$$p_X(k) = \begin{cases} & k = -2 \\ & k = -1 \\ & k = 0 \\ & k = 1 \\ & k = 2 \end{cases}$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Brute Force

$$\Omega_X = \{-2, -1, 0, +1, +2\}$$

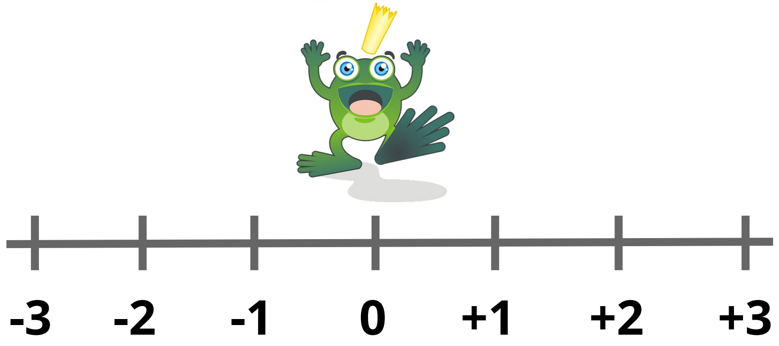
$$p_X(k) = \begin{cases} p_L^2, & k = -2 \\ & k = -1 \\ & k = 0 \\ & k = 1 \\ & k = 2 \end{cases}$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Brute Force

$$\Omega_X = \{-2, -1, 0, +1, +2\}$$

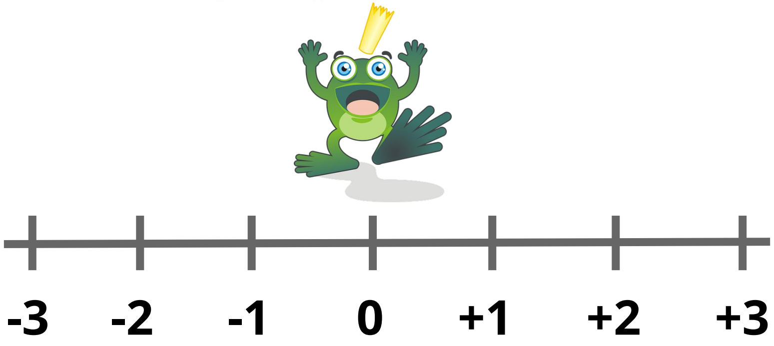
$$p_X(k) = \begin{cases} p_L^2, & k = -2 \\ & k = -1 \\ & k = 0 \\ & k = 1 \\ p_R^2, & k = 2 \end{cases}$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Brute Force

$$\Omega_X = \{-2, -1, 0, +1, +2\}$$

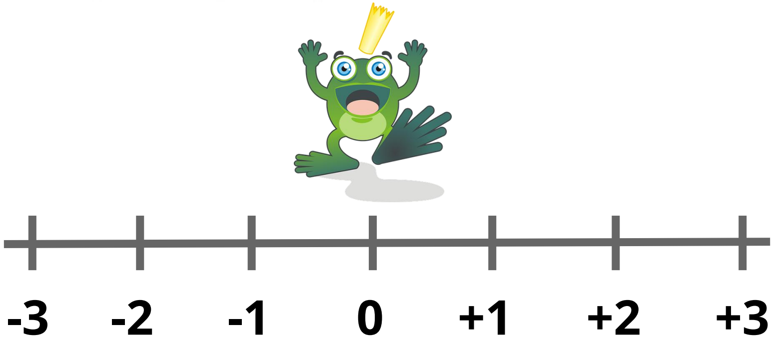
$$p_X(k) = \begin{cases} p_L^2, & k = -2 \\ 2p_L p_S, & k = -1 \\ & k = 0 \\ & k = 1 \\ p_R^2, & k = 2 \end{cases}$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Brute Force

$$\Omega_X = \{-2, -1, 0, +1, +2\}$$

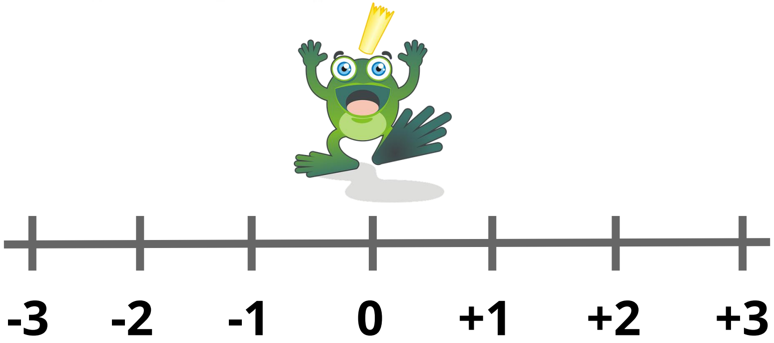
$$p_X(k) = \begin{cases} p_L^2, & k = -2 \\ 2p_L p_S, & k = -1 \\ & k = 0 \\ 2p_R p_S, & k = 1 \\ p_R^2, & k = 2 \end{cases}$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Brute Force

$$\Omega_X = \{-2, -1, 0, +1, +2\}$$

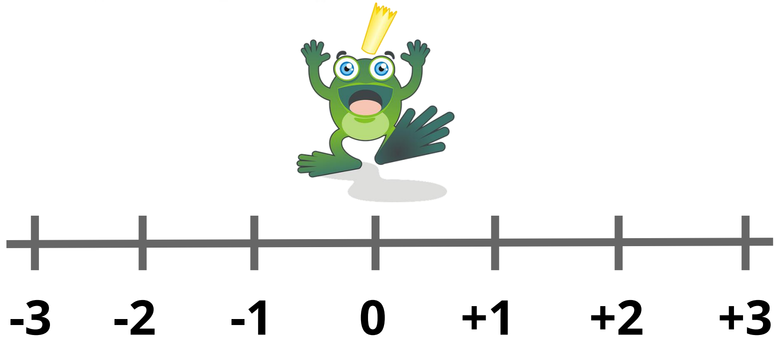
$$p_X(k) = \begin{cases} p_L^2, & k = -2 \\ 2p_L p_S, & k = -1 \\ 2p_L p_R + p_S^2, & k = 0 \\ 2p_R p_S, & k = 1 \\ p_R^2, & k = 2 \end{cases}$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



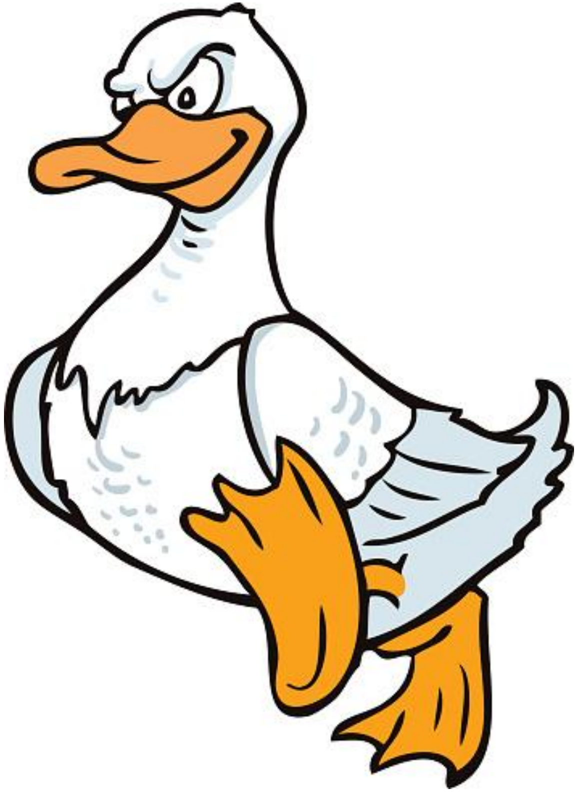
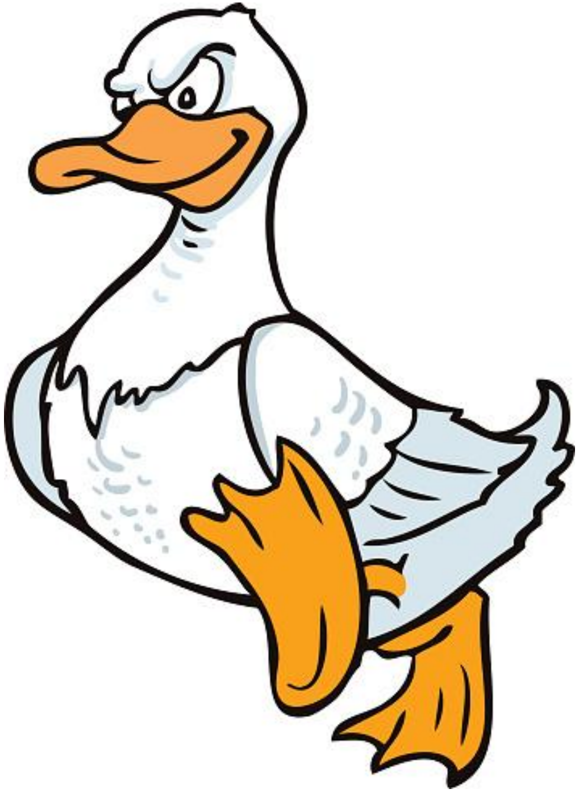
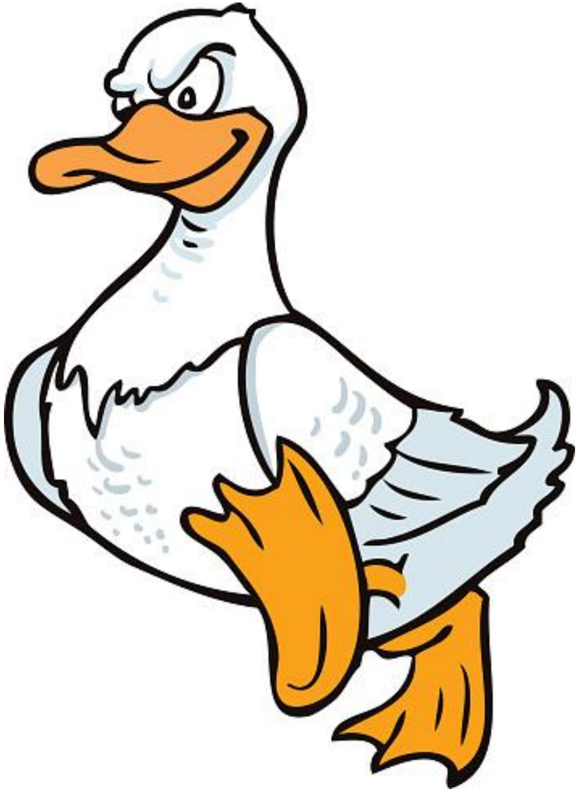
Brute Force

$$\Omega_X = \{-2, -1, 0, +1, +2\}$$

$$p_X(k) = \begin{cases} p_L^2, & k = -2 \\ 2p_L p_S, & k = -1 \\ 2p_L p_R + p_S^2, & k = 0 \\ 2p_R p_S, & k = 1 \\ p_R^2, & k = 2 \end{cases}$$

$$E[X] = (-2) \cdot p_L^2 + (-1) \cdot 2p_L p_S + \cdots = 2(p_R - p_L)$$

Random Picture

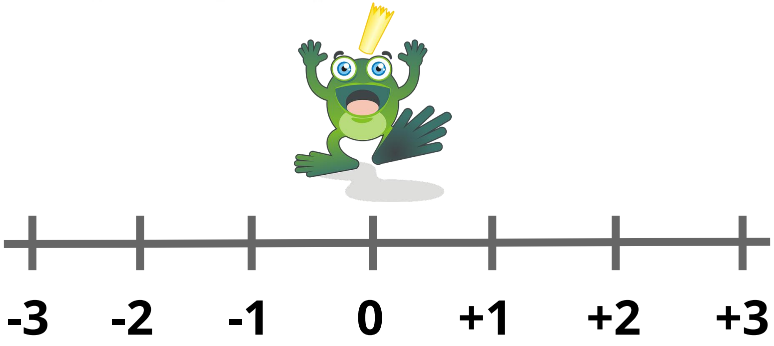
Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

LINEARITY

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Frogger!

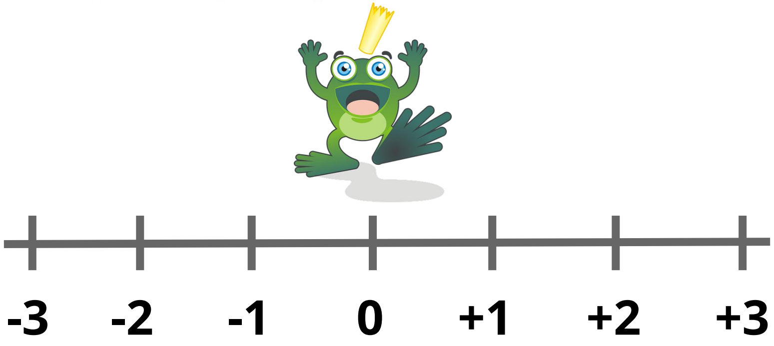
A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?

Linearity

Let X_1, X_2 be the distance the frog travels at time steps 1,2 respectively.

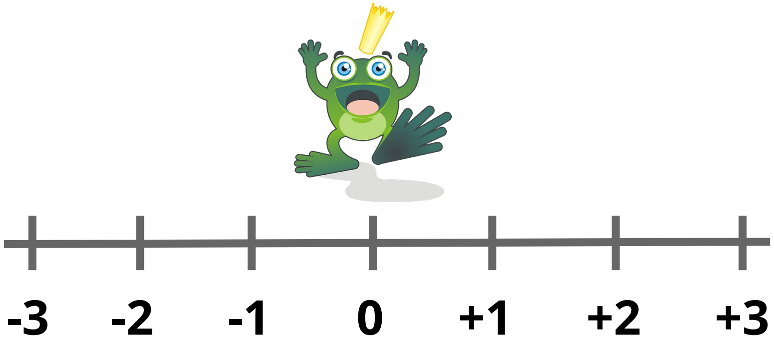


Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Linearity

Let X_1, X_2 be the distance the frog travels at time steps 1,2 respectively.

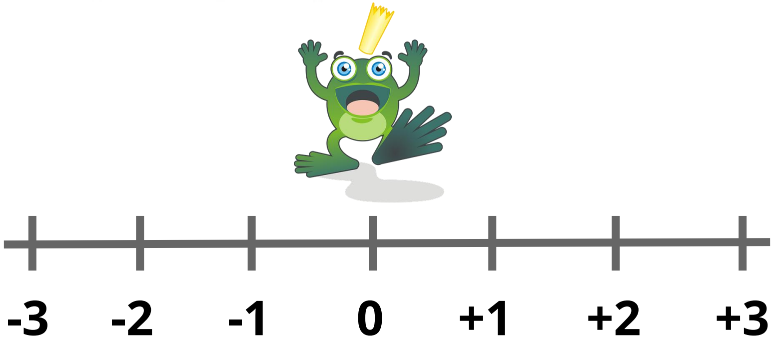
Important Observation: $X = X_1 + X_2$. Why?

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Linearity

Let X_1, X_2 be the distance the frog travels at time steps 1,2 respectively.

Important Observation: $X = X_1 + X_2$. Why?

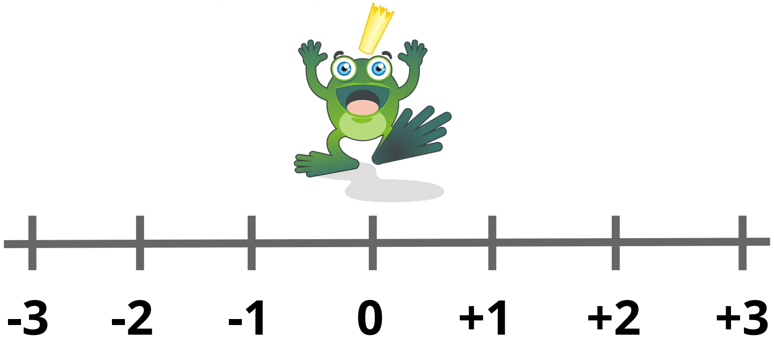
$$\Omega_{X_1} = \Omega_{X_2} = \{-1, 0, +1\}$$

FROGGER!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



LINEARITY

Let X_1, X_2 be the distance the frog travels at time steps 1,2 respectively.

Important Observation: $X = X_1 + X_2$. Why?

$$\Omega_{X_1} = \Omega_{X_2} = \{-1, 0, +1\}$$

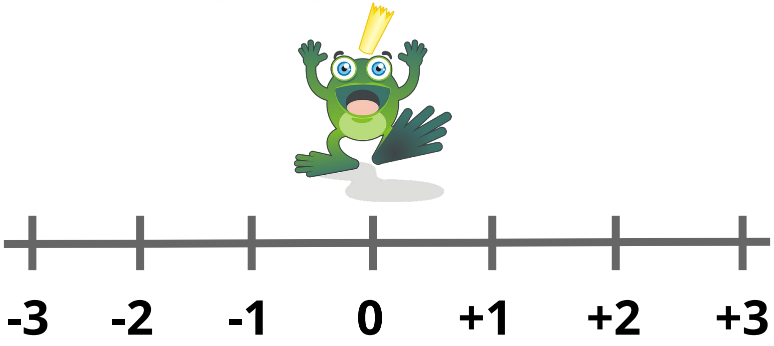
$$p_{X_i}(k) = \begin{cases} p_L, & k = -1 \\ p_S, & k = 0 \\ p_R, & k = 1 \end{cases}$$

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Linearity

Let X_1, X_2 be the distance the frog travels at time steps 1,2 respectively.

Important Observation: $X = X_1 + X_2$. Why?

$$\Omega_{X_1} = \Omega_{X_2} = \{-1,0,+1\}$$

$$p_{X_i}(k) = \begin{cases} p_L, & k = -1 \\ p_S, & k = 0 \\ p_R, & k = 1 \end{cases}$$

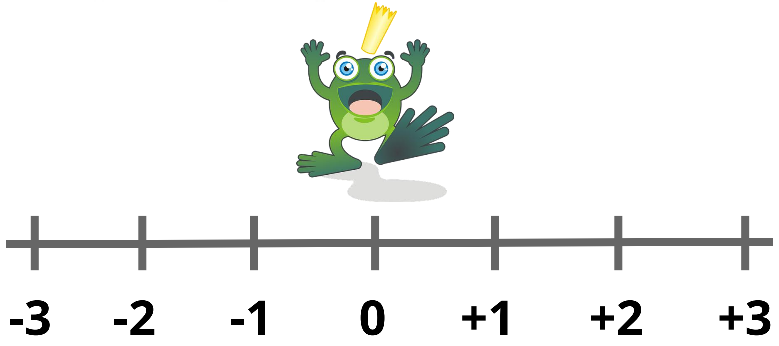
So $E[X_i] = -1 \cdot p_L + 0 \cdot p_S + 1 \cdot p_R = p_R - p_L$.

Frogger!

A frog starts on a 1-dimensional number line at 0. At each time step, it moves

- left with probability p_L
- right with probability p_R
- stays with probability p_S

where $p_L + p_R + p_S = 1$. Let X be the position of the frog after 2 (independent) time steps. What is $E[X]$?



Linearity

Let X_1, X_2 be the distance the frog travels at time steps 1,2 respectively.

Important Observation: $X = X_1 + X_2$. Why?

$$\Omega_{X_1} = \Omega_{X_2} = \{-1, 0, +1\}$$

$$p_{X_i}(k) = \begin{cases} p_L, & k = -1 \\ p_S, & k = 0 \\ p_R, & k = 1 \end{cases}$$

So $E[X_i] = -1 \cdot p_L + 0 \cdot p_S + 1 \cdot p_R = p_R - p_L$.

By linearity of expectation,

$E[X] = E[X_1 + X_2] = E[X_1] + E[X_2] = 2(p_R - p_L)$.

Frogger!

Which method was Easier?

Frogger!

Which method was Easier?

In this case, it might be debatable.

Frogger!

Which method was Easier?

In this case, it might be debatable.

But what if we changed the number of time steps to 100 or 1000?

Frogger!

Which method was Easier?

In this case, it might be debatable.

But what if we changed the number of time steps to 100 or 1000?

the first method would be completely infeasible, but the second (LoE) would be basically the same.

Flipping 2 coins again

Let X be the number of heads in two independent flips of a fair coin. Recall

$$\Omega_X = \{0,1,2\}$$

$$p_X(d) = \begin{cases} 1/4, & d = 0 \\ 1/2, & d = 1 \\ 1/4, & d = 2 \end{cases}$$

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Flipping 2 coins again

Let X be the number of heads in two independent flips of a fair coin. Recall

$$\Omega_X = \{0,1,2\}$$

$$p_X(d) = \begin{cases} 1/4, & d = 0 \\ 1/2, & d = 1 \\ 1/4, & d = 2 \end{cases}$$

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Let $Y = X^3$. What does that mean?

Flipping 2 coins again

Let X be the number of heads in two independent flips of a fair coin. Recall

$$\Omega_X = \{0,1,2\}$$

$$p_X(d) = \begin{cases} 1/4, & d = 0 \\ 1/2, & d = 1 \\ 1/4, & d = 2 \end{cases}$$

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Let $Y = X^3$. What does that mean? Literally the cubed number of heads.

Flipping 2 coins again

Let X be the number of heads in two independent flips of a fair coin. Recall

$$\Omega_X = \{0,1,2\}$$

$$p_X(d) = \begin{cases} 1/4, & d = 0 \\ 1/2, & d = 1 \\ 1/4, & d = 2 \end{cases}$$

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Let $Y = X^3$. What does that mean? Literally the cubed number of heads.

$$\Omega_Y = \{0,1,8\}$$

Flipping 2 coins again

Let X be the number of heads in two independent flips of a fair coin. Recall

$$\Omega_X = \{0,1,2\}$$

$$p_X(d) = \begin{cases} 1/4, & d = 0 \\ 1/2, & d = 1 \\ 1/4, & d = 2 \end{cases}$$

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Let $Y = X^3$. What does that mean? Literally the cubed number of heads.

$$\Omega_Y = \{0,1,8\}$$

$$p_Y(z) = \begin{cases} 1/4, & z = 0 \\ 1/2, & z = 1 \\ 1/4, & z = 8 \end{cases}$$

Flipping 2 coins again

Let X be the number of heads in two independent flips of a fair coin. Recall

$$\Omega_X = \{0,1,2\}$$

$$p_X(d) = \begin{cases} 1/4, & d = 0 \\ 1/2, & d = 1 \\ 1/4, & d = 2 \end{cases}$$

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Let $Y = X^3$. What does that mean? Literally the cubed number of heads.

$$\Omega_Y = \{0,1,8\}$$

$$p_Y(z) = \begin{cases} 1/4, & z = 0 \\ 1/2, & z = 1 \\ 1/4, & z = 8 \end{cases}$$

$$E[X^3] = E[Y] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 8 \cdot \frac{1}{4} = 2.5$$

Flipping 2 coins again

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1 \qquad E[X^3] = E[Y] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 8 \cdot \frac{1}{4} = 2.5$$

See an easier way to compute $E[X^3] = E[Y]$ without going through the trouble of writing out p_Y?

Flipping 2 coins again

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1 \qquad E[X^3] = E[Y] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 8 \cdot \frac{1}{4} = 2.5$$

See an easier way to compute $E[X^3] = E[Y]$ without going through the trouble of writing out p_Y?

$$E[X^3] = 0^3 \cdot \frac{1}{4} + 1^3 \cdot \frac{1}{2} + 2^3 \cdot \frac{1}{4} = 2.5$$

Flipping 2 coins again

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1 \qquad E[X^3] = E[Y] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 8 \cdot \frac{1}{4} = 2.5$$

See an easier way to compute $E[X^3] = E[Y]$ without going through the trouble of writing out p_Y?

$$E[X^3] = 0^3 \cdot \frac{1}{4} + 1^3 \cdot \frac{1}{2} + 2^3 \cdot \frac{1}{4} = 2.5$$

$$E[X^3] = \sum_{b \in \Omega_X} b^3 p_X(b)$$

Flipping 2 coins again

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

$$E[X^3] = E[Y] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 8 \cdot \frac{1}{4} = 2.5$$

See an easier way to compute $E[X^3] = E[Y]$ without going through the trouble of writing out p_Y?

$$E[X^3] = 0^3 \cdot \frac{1}{4} + 1^3 \cdot \frac{1}{2} + 2^3 \cdot \frac{1}{4} = 2.5$$

$$E[X^3] = \sum_{b \in \Omega_X} b^3 p_X(b)$$

$$E[g(X)] = \sum_{b \in \Omega_X} g(b) p_X(b)$$

Flipping 2 coins again

$$E[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1 \qquad E[X^3] = E[Y] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 8 \cdot \frac{1}{4} = 2.5$$

See an easier way to compute $E[X^3] = E[Y]$ without going through the trouble of writing out p_Y?

$$E[X^3] = 0^3 \cdot \frac{1}{4} + 1^3 \cdot \frac{1}{2} + 2^3 \cdot \frac{1}{4} = 2.5$$

$$E[X^3] = \sum_{b \in \Omega_X} b^3 p_X(b)$$

$$E[g(X)] = \sum_{b \in \Omega_X} g(b) p_X(b)$$

Notice $2.5 = E[X^3] \neq E[X]^3 = 1$.

Law of the Unconscious Statistician (LOTUS)

Law of the Unconscious Statistician (LOTUS): Let X be a discrete random variable with range Ω_X and $g: D \to \mathbb{R}$ be a function defined at least over Ω_X ($\Omega_X \subseteq D$). Then,

$$E[g(X)] = \sum_{b \in \Omega_X} g(b) p_X(b)$$

Note that in general, $E[g(X)] \neq g(E[X])$. For example, $E[X^2] \neq (E[X])^2$, or $E[\log(X)] \neq \log(E[X])$.

Probability

3.3 Variance

Alex Tsun

Agenda

- LoE with Indicator RVs
- Variance and Standard Deviation (SD)

Red Hair

There are 7 mermaids in the sea. Below is a table of each mermaid with their hair color. How many mermaids have red hair?

Red Hair

There are 7 mermaids in the sea. Below is a table of each mermaid with their hair color. How many mermaids have red hair?

Mermaid	1	2	3	4	5	6	7
Color	RED	BLUE	PURPLE	RED	BLACK	YELLOW	RED

Red Hair

There are 7 mermaids in the sea. Below is a table of each mermaid with their hair color. How many mermaids have red hair?

Mermaid	1	2	3	4	5	6	7
Color	RED	BLUE	PURPLE	RED	BLACK	YELLOW	RED
1 / 0	1	0	0	1	0	0	1

Red Hair

There are 7 mermaids in the sea. Below is a table of each mermaid with their hair color. How many mermaids have red hair?

Mermaid	1	2	3	4	5	6	7
Color	RED	BLUE	PURPLE	RED	BLACK	YELLOW	RED
1 / 0	1	0	0	1	0	0	1

Sum the bottom row of "indicator variables" to get 3.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

Brute Force: $\Omega_X = \{0,1,2, \ldots, n-2, n\}$.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

Brute Force: $\Omega_X = \{0,1,2,\ldots,n-2,n\}$.

$$p_X(n) = \frac{1}{n!}$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

Brute Force: $\Omega_X = \{0,1,2,\ldots,n-2,n\}$.

$$p_X(n) = \frac{1}{n!}$$

$$p_X(0) = ???$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

Brute Force: $\Omega_X = \{0,1,2,\ldots,n-2,n\}$.

$$p_X(n) = \frac{1}{n!}$$

$$p_X(0) = ???$$

Too hard → Use linearity!

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

Quick question: does it matter where you are in line?

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

Quick question: does it matter where you are in line?

If first in line, $P(\text{get hat back}) = \frac{1}{n}$, because there are n in total.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

Quick question: does it matter where you are in line?

If first in line, $P(\text{get hat back}) = \frac{1}{n}$, because there are n in total.
If last in line, $P(\text{get hat back}) = \frac{1}{n}$, because there is 1 left, and the chance it is yours is $\frac{1}{n}$.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) =$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) =$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) =$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \ldots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] =$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \ldots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] \overset{\text{LINEARITY}}{=} \sum_{i=1}^{n} E[X_i] =$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] \overset{\text{Linearity}}{=} \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{n} =$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] \overset{\text{Linearity}}{=} \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{n} = n \cdot \frac{1}{n} = 1$$

Hat Check

n people go to a party and leave their hat with a hat-check person. At the end of the party, they return hats randomly since they don't care about their job. Let X be the number of people who got their original hat back. What is $E[X]$?

For $i = 1, \dots, n$, let $X_i = \begin{cases} 1, & \text{if } i^{th} \text{ person got hat back} \\ 0, & \text{otherwise} \end{cases}$. Then $X = \sum_{i=1}^{n} X_i$.

NOT "INDEPENDENT" RVS

We will use linearity of expectation.

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = P(i^{th} \text{ person got hat back}) = \frac{1}{n}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{n} = n \cdot \frac{1}{n} = 1$$

Linearity of Expectation with Indicators

<u>**Linearity of Expectation with Indicators:**</u> If asked only about the expectation of a RV X (and not its PMF), then you may be able to write X as a sum of possibly "dependent" indicator (1/0) random variables, and apply LoE.

For an indicator RV X_i,

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1)$$

Random Picture

Variance (Intuition)

Which game would you rather play? We flip a fair coin.

Variance (Intuition)

Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me $1.
- If Tails, I pay you $1.

Variance (Intuition)

Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

Game 2:

- If Heads, you pay me \$1000.
- If Tails, I pay you \$1000.

Variance (Intuition)

Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

Game 2:

- If Heads, you pay me \$1000.
- If Tails, I pay you \$1000.

Both games are fair.

$$E[G_1] = -1\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = 0$$

$$E[G_2] = -1000\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = 0$$

Variance (Intuition)

How far is a random variable from its mean, on average?

$$X - E[X]$$

Variance (Intuition)

HOW FAR IS A RANDOM VARIABLE FROM ITS MEAN, ON AVERAGE?

$$\lvert X - E[X] \rvert$$

Variance (Intuition)

how far is a random variable from its mean, on average?

$$|X - E[X]|$$

$$(X - E[X])^2$$

VARIANCE (INTUITION)

HOW FAR IS A RANDOM VARIABLE FROM ITS MEAN, ON AVERAGE?

$$E[\,|X - E[X]|\,]$$

$$E[\,(X - E[X])^2\,]$$

Variance and Standard Deviation (SD)

Variance: The variance of a random variable X is

$$Var(X) = E[(X - E[X])^2]$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2\ Var(X)$$

Variance and Standard Deviation (SD)

<u>Variance:</u> The variance of a random variable X is

More Useful

$$Var(X) = E[(X - E[X])^2] \boxed{= E[X^2] - (E[X])^2}$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2\ Var(X)$$

Variance and Standard Deviation (SD)

<u>**Variance:**</u> The variance of a random variable X is

More Useful

$$Var(X) = E[(X - E[X])^2] \boxed{= E[X^2] - (E[X])^2}$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2\ Var(X)$$

<u>**Standard Deviation (SD):**</u> The standard deviation of a random variable X is

$$\sigma_X = \sqrt{Var(X)}$$

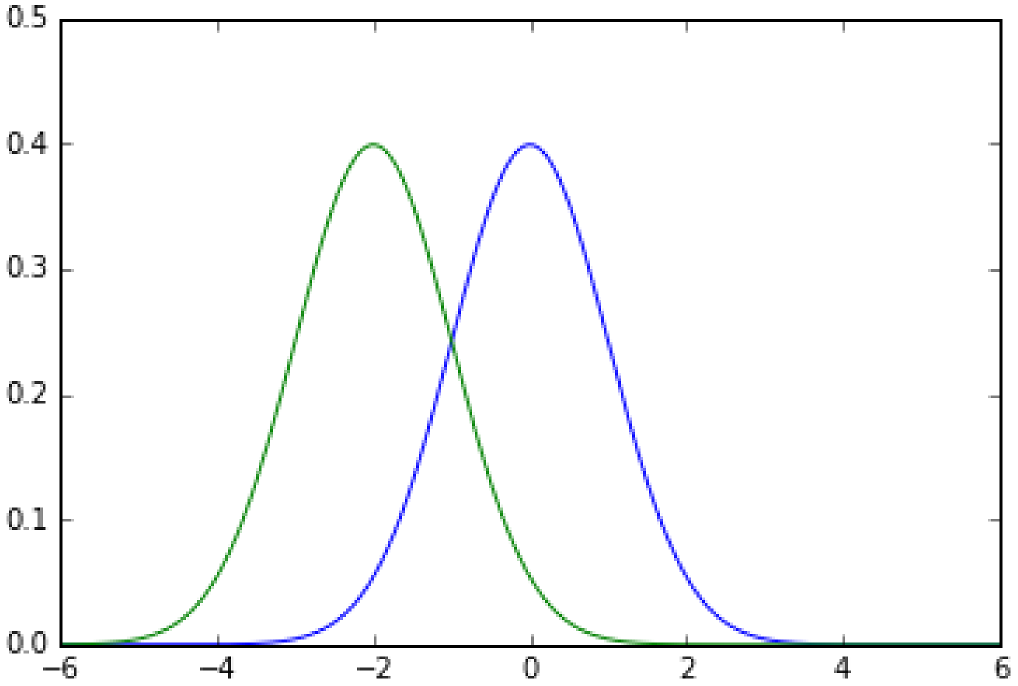
We want this because the units of variance are squared in terms of the original variable X, and this "undo's" our squaring, returning the units to the same as X.

Variance (Property)

$$Var(aX + b) = a^2Var(X)$$

Variance (Property)

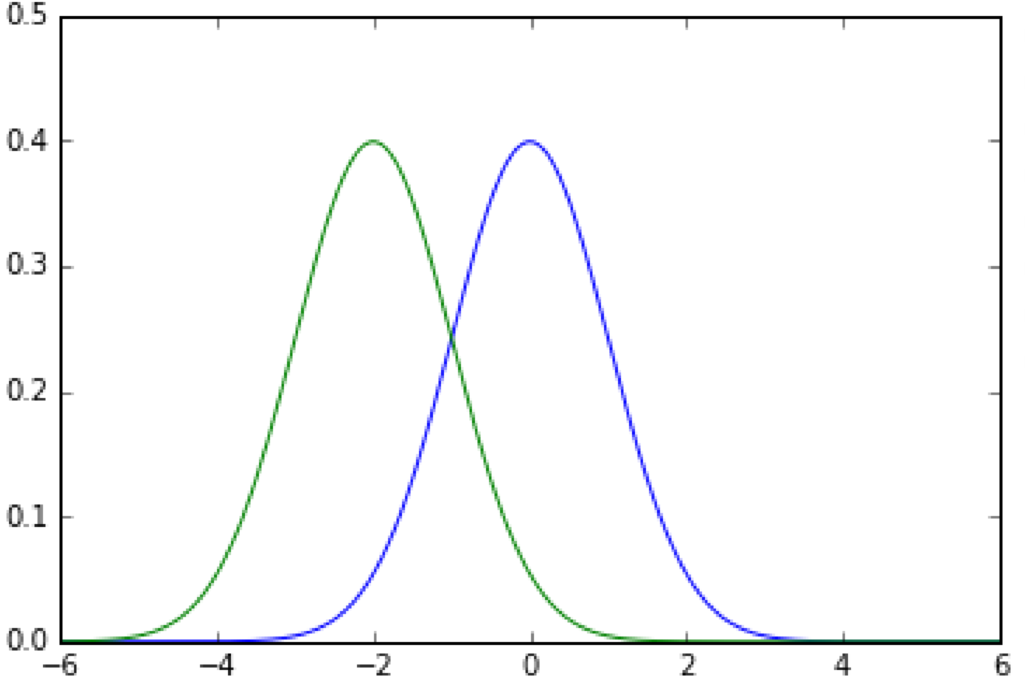
$$Var(aX + b) = a^2 Var(X)$$



Variance (Property)

$$Var(aX + b) = a^2Var(X)$$

$$Var(X + b) = Var(X)$$



Variance (Property)

$$Var(aX + b) = a^2 Var(X)$$

$Var(aX)$

Variance (Property)

$$Var(aX + b) = a^2 Var(X)$$

$$Var(aX) = E[(aX)^2] - (E[aX])^2$$

Variance (Property)

$$Var(aX + b) = a^2 Var(X)$$

$$Var(aX) = E[(aX)^2] - (E[aX])^2 = E[a^2X^2] - (aE[X])^2$$

Variance (Property)

$$Var(aX + b) = a^2 Var(X)$$

$$Var(aX) = E[(aX)^2] - (E[aX])^2 = E[a^2X^2] - (aE[X])^2$$

$$= a^2E[X^2] - a^2(E[X])^2$$

Variance (Property)

$$Var(aX + b) = a^2 Var(X)$$

$$Var(aX) = E[(aX)^2] - (E[aX])^2 = E[a^2X^2] - (aE[X])^2$$

$$= a^2E[X^2] - a^2(E[X])^2 = a^2(E[X^2] - E[X]^2)$$

Variance (Property)

$$Var(aX + b) = a^2 Var(X)$$

$$Var(aX) = E[(aX)^2] - (E[aX])^2 = E[a^2X^2] - (aE[X])^2$$

$$= a^2E[X^2] - a^2(E[X])^2 = a^2(E[X^2] - E[X]^2) = a^2Var(X)$$

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

$$E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + \cdots + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

LOTUS

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

$$E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + \cdots + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{91}{6} - (3.5)^2 = \frac{35}{12}$$

Bài 3.1. Cho biến ngẫu nhiên rời rạc X có bảng phân phối xác suất sau:

X	-2	-1	0	1	2
$\mathbb{P}$	$1/8$	$2/8$	$2/8$	$2/8$	$1/8$

(a) Tìm hàm phân phối xác suất $F(x)$.

(b) Tính $\mathbb{P}(-1 \leq X \leq 1)$ và $\mathbb{P}\big(X \leq -1$ hoặc $X = 2\big)$.

(c) Lập bảng phân phối xác suất của biến ngẫu nhiên $Y = X^2$.

Đáp án. (b) 6/8, 4/8. (c)

Y	0	1	4
$\mathbb{P}$	$2/8$	$4/8$	$2/8$

Bài 3.2. Biến ngẫu nhiên rời rạc X có hàm xác suất cho bởi

$$f(x) = \frac{2x+1}{25}, \quad x = 0, 1, 2, 3, 4$$

(a) Lập bảng phân phối xác suất của X.

(b) Tính $\mathbb{P}(2 \leq X < 4)$ và $\mathbb{P}(X > -10)$.

Đáp án. (b) 12/25, 1.

Probability

3.4 Zoo of Discrete RV's Part I

Alex Tsun

Agenda

- Independence
- The Bernoulli Process
- The Bernoulli/indicator RV
- The Binomial RV

Independence

Independence: Random variables X and Y are independent, denoted $X \perp Y$, if for all $x \in \Omega_X$ and $y \in \Omega_Y$, any of the three equivalent properties holds:

1. $P(X = x|Y = y) = P(X = x)$
2. $P(Y = y|X = x) = P(Y = y)$
3. $P(X = x, Y = y) = P(X = x)P(Y = y)$

Independence

<u>Independence:</u> Random variables X and Y are independent, denoted $X \perp Y$, if for all $x \in \Omega_X$ and $y \in \Omega_Y$, any of the three equivalent properties holds:

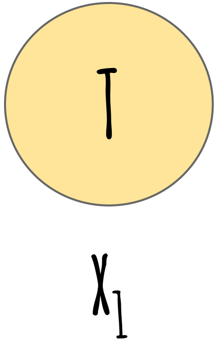
1. $P(X = x | Y = y) = P(X = x)$
2. $P(Y = y | X = x) = P(Y = y)$
3. $P(X = x, Y = y) = P(X = x)P(Y = y)$

This is the same as the event definition, but it must hold for all events $\{X = x\}$ and $\{Y = y\}$. We'll discuss this further, and prove this useful fact later as well: If $X \perp Y$, then

$$Var(X + Y) = Var(X) + Var(Y)$$

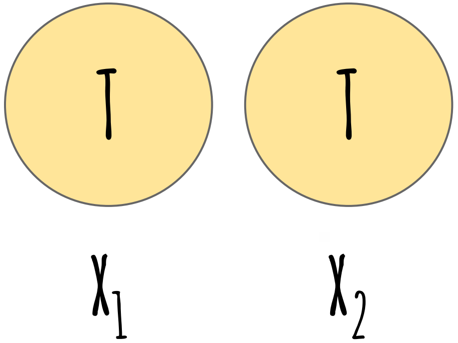
The Bernoulli Process

<u>Bernoulli Process:</u> A Bernoulli process with parameter p is a sequence of independent coin flips $X_1, X_2, X_3, \ldots$ where $P(\text{head}) = p$. If flip i is heads, then we encode $X_i = 1$; otherwise if flip i is tails, then $X_i = 0$.



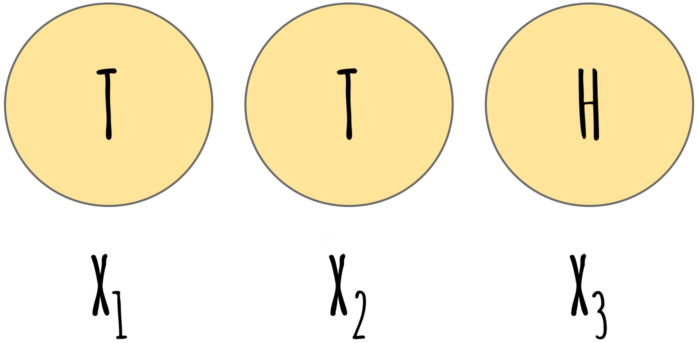
The Bernoulli Process

<u>Bernoulli Process:</u> A Bernoulli process with parameter p is a sequence of independent coin flips $X_1, X_2, X_3, \ldots$ where $P(\text{head}) = p$. If flip i is heads, then we encode $X_i = 1$; otherwise if flip i is tails, then $X_i = 0$.



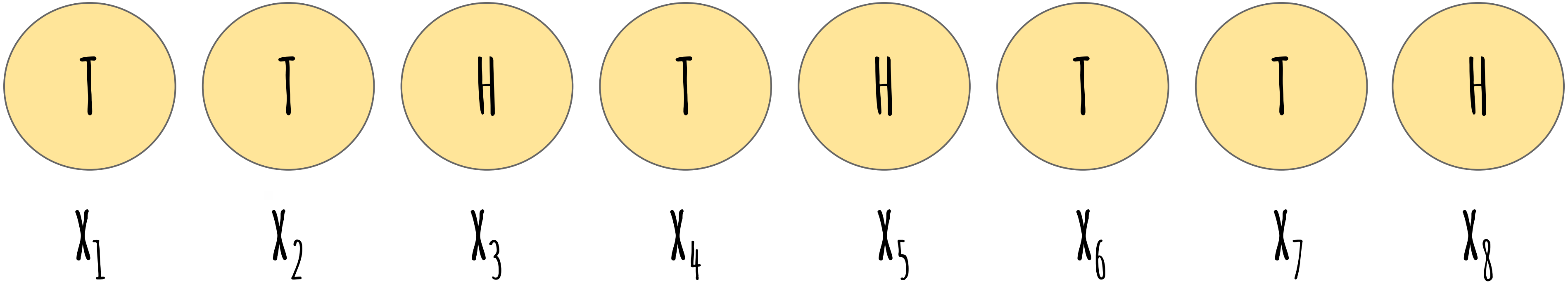
The Bernoulli Process

<u>**Bernoulli Process:**</u> A Bernoulli process with parameter p is a sequence of independent coin flips $X_1, X_2, X_3, \ldots$ where $P(\text{head}) = p$. If flip i is heads, then we encode $X_i = 1$; otherwise if flip i is tails, then $X_i = 0$.



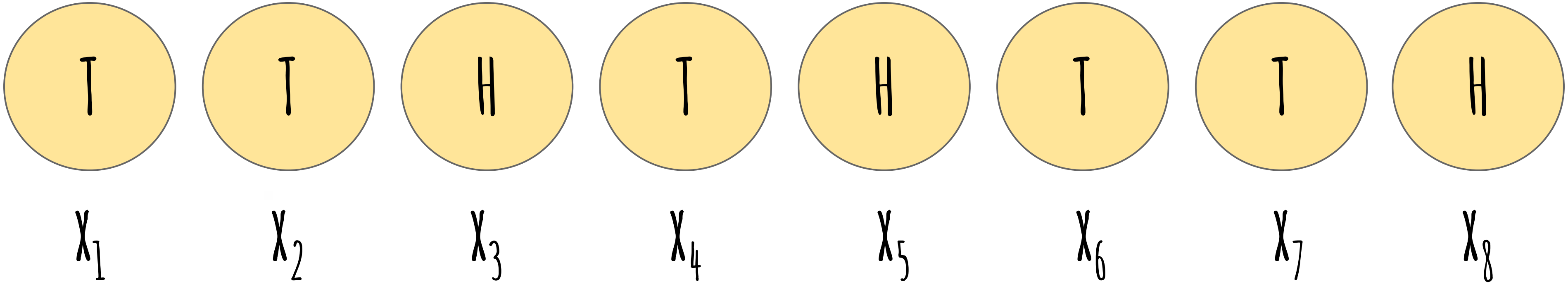
The Bernoulli Process

Bernoulli Process: A Bernoulli process with parameter p is a sequence of independent coin flips $X_1, X_2, X_3, \ldots$ where $P(\text{head}) = p$. If flip i is heads, then we encode $X_i = 1$; otherwise if flip i is tails, then $X_i = 0$.



The Bernoulli Process

Bernoulli Process: A Bernoulli process with parameter p is a sequence of independent coin flips $X_1, X_2, X_3, \ldots$ where $P(\text{head}) = p$. If flip i is heads, then we encode $X_i = 1$; otherwise if flip i is tails, then $X_i = 0$.



From this process, we can measure several interesting things.

The Bernoulli RV

Ber(p)

T

X_1

The Bernoulli RV

Ber(p) Ber(p)

T T

X_1 X_2

The Bernoulli RV

Ber(p)	Ber(p)	Ber(p)
T	T	H
X_1	X_2	X_3

The Bernoulli RV

Ber(p) Ber(p) Ber(p) Ber(p) ...

T	T	H	T	H	T	T	H
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8

Bernoulli RV Properties

$$E[X] =$$

Bernoulli RV Properties

$$E[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

Bernoulli RV Properties

$$E[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$E[X^2] =$$

Bernoulli RV Properties

$$E[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$E[X^2] = 1^2 \cdot p + 0^2 \cdot (1 - p) = p$$ LOTUS

Bernoulli RV Properties

$$E[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$E[X^2] = 1^2 \cdot p + 0^2 \cdot (1 - p) = p$$ LOTUS

$$Var(X) = E[X^2] - (E[X])^2 =$$

Bernoulli RV Properties

$$E[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$E[X^2] = 1^2 \cdot p + 0^2 \cdot (1 - p) = p$$ LOTUS

$$Var(X) = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p)$$

The Bernoulli/indicator RV

Bernoulli RV: $X \sim Ber(p)$ if and only if X has the following pmf:

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

The Bernoulli/indicator RV

<u>Bernoulli RV:</u> $X \sim Ber(p)$ if and only if X has the following pmf:

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

Each X_i in the Bernoulli process with parameter p is a Bernoulli/indicator rv with parameter p. It simply represents a binary outcome, like a coin flip.

$$E[X] = p \qquad Var(X) = p(1-p)$$

The Bernoulli/indicator RV

Bernoulli RV: $X \sim Ber(p)$ if and only if X has the following pmf:

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

Each X_i in the Bernoulli process with parameter p is a Bernoulli/indicator rv with parameter p. It simply represents a binary outcome, like a coin flip.

$$E[X] = p \qquad Var(X) = p(1 - p)$$

By a clever trick, we can write the PMF as

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

Random Picture

The Binomial RV

$\text{Bin}(3, p) = 1$

T	T	H
X_1	X_2	X_3

The Binomial RV

$\text{Bin}(3, p) = 1$ $\quad$ $\text{Bin}(1, p) = 0$

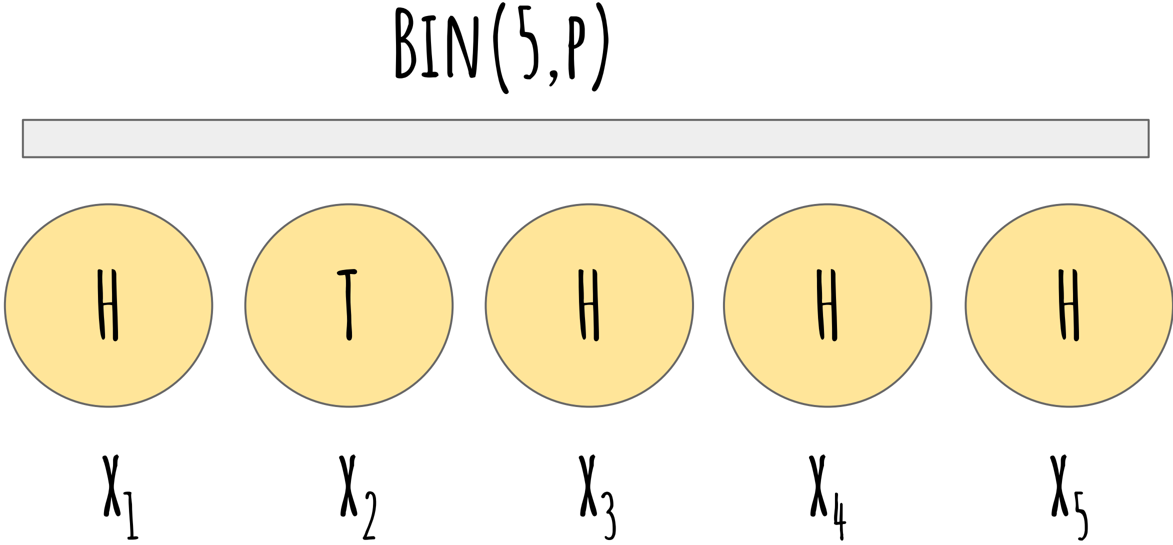
T	T	H	T
X_1	X_2	X_3	X_4

The Binomial RV

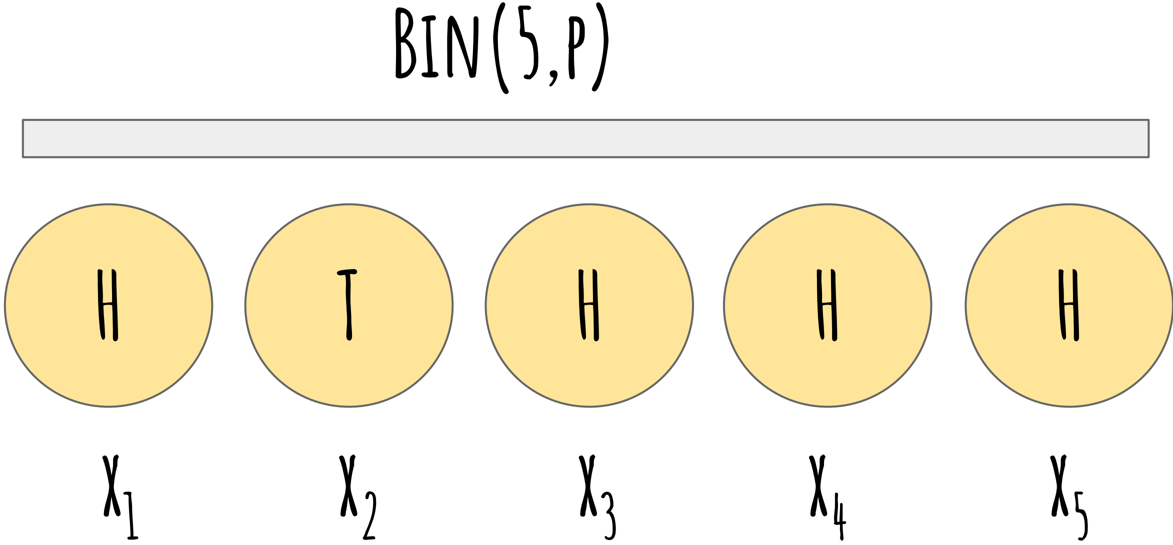
$\text{Bin}(3,p) = 1$ $\text{Bin}(1,p) = 0$ $\text{Bin}(4,p) = 2$

T	T	H	T	H	T	T	H
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8

Binomial RV PMF

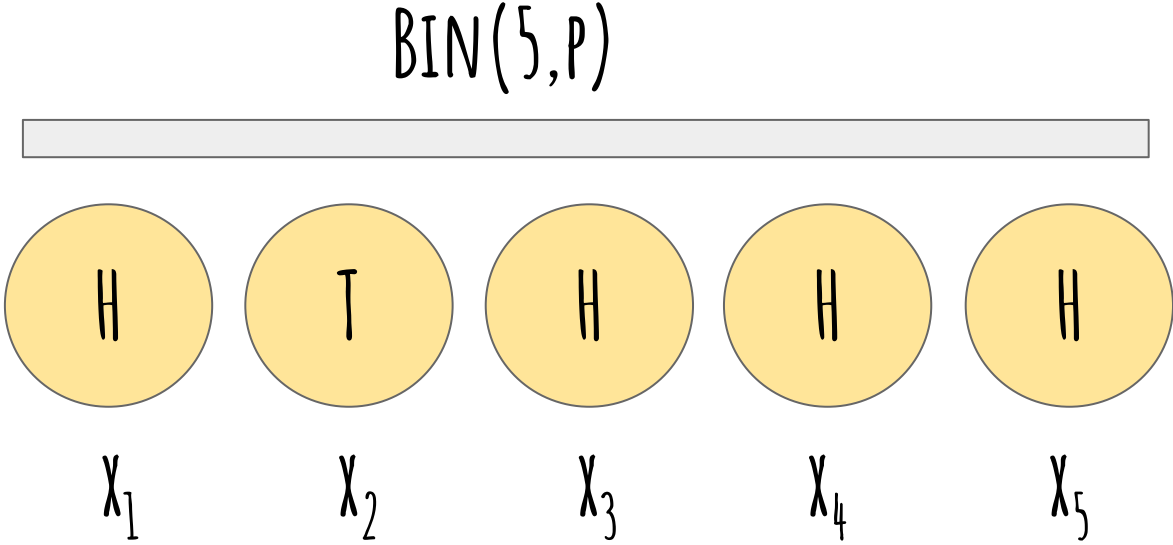


Binomial RV PMF



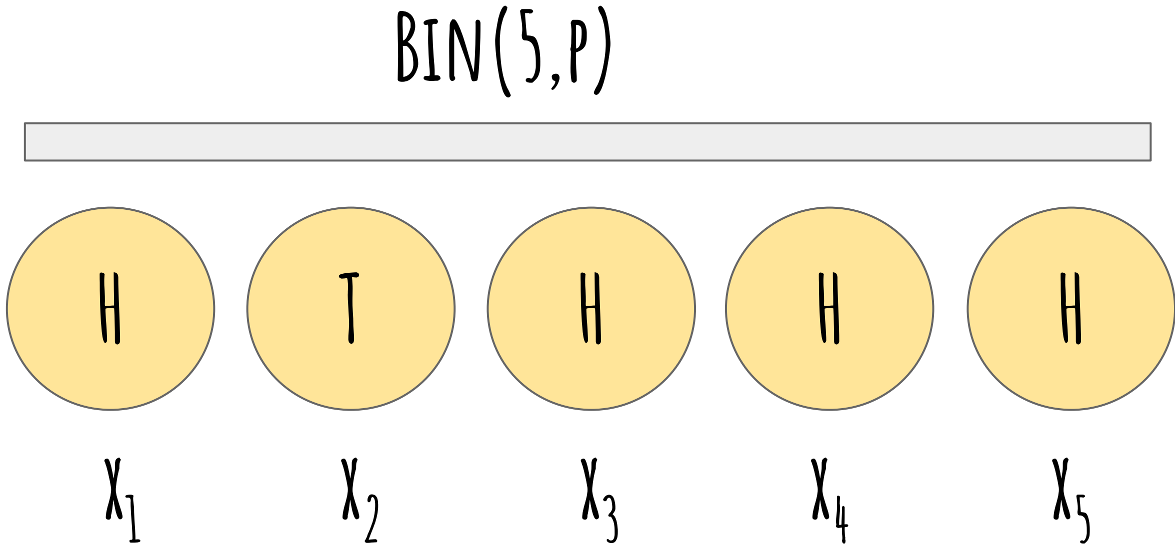
$$P(HTHHH) =$$

Binomial RV PMF



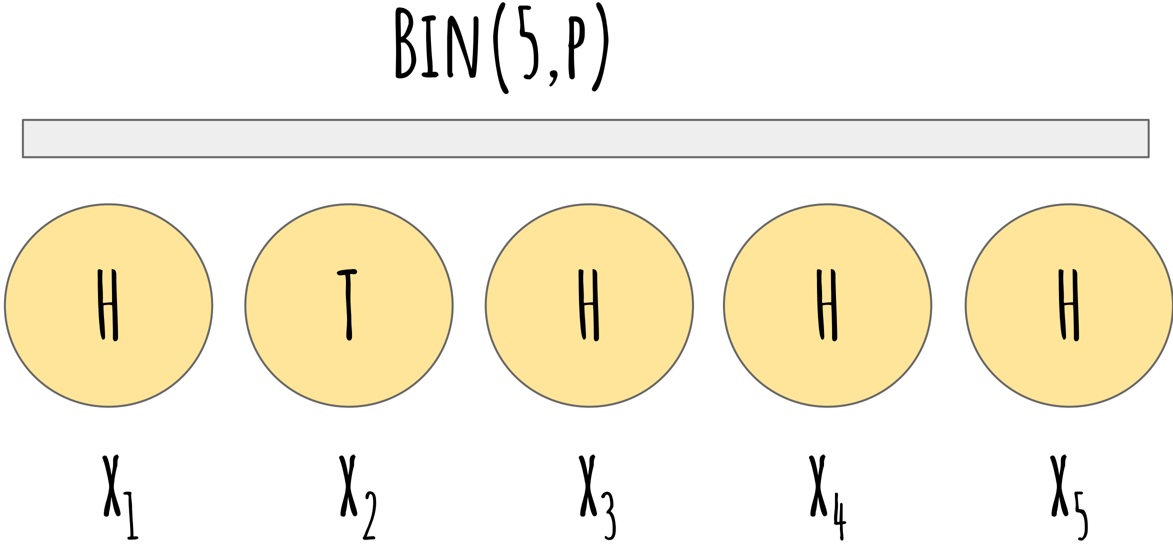
$$P(HTHHH) = p$$

Binomial RV PMF



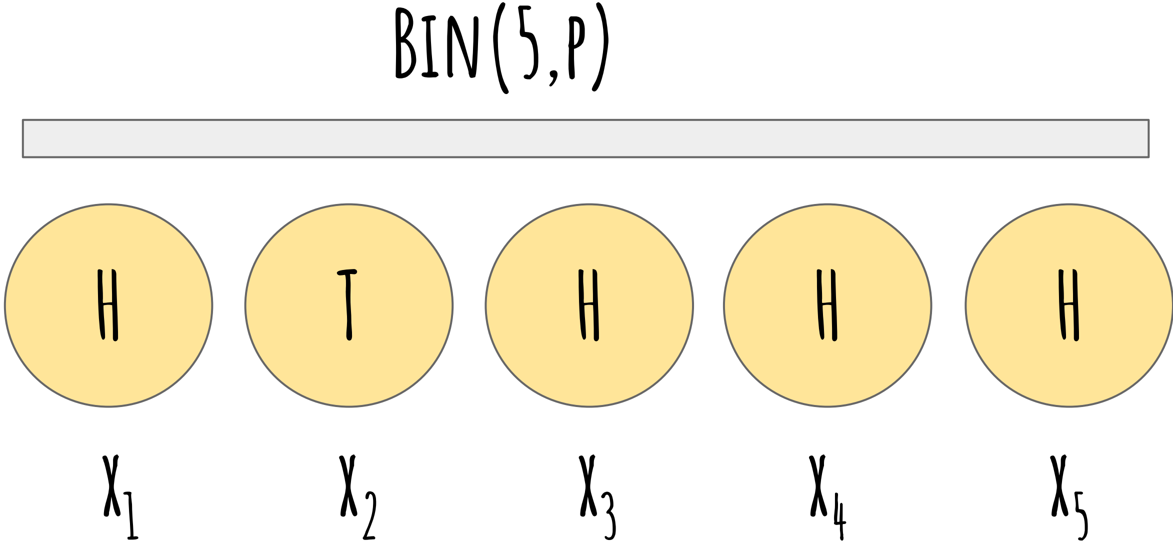
$$P(HTHHH) = p(1-p)$$

Binomial RV PMF



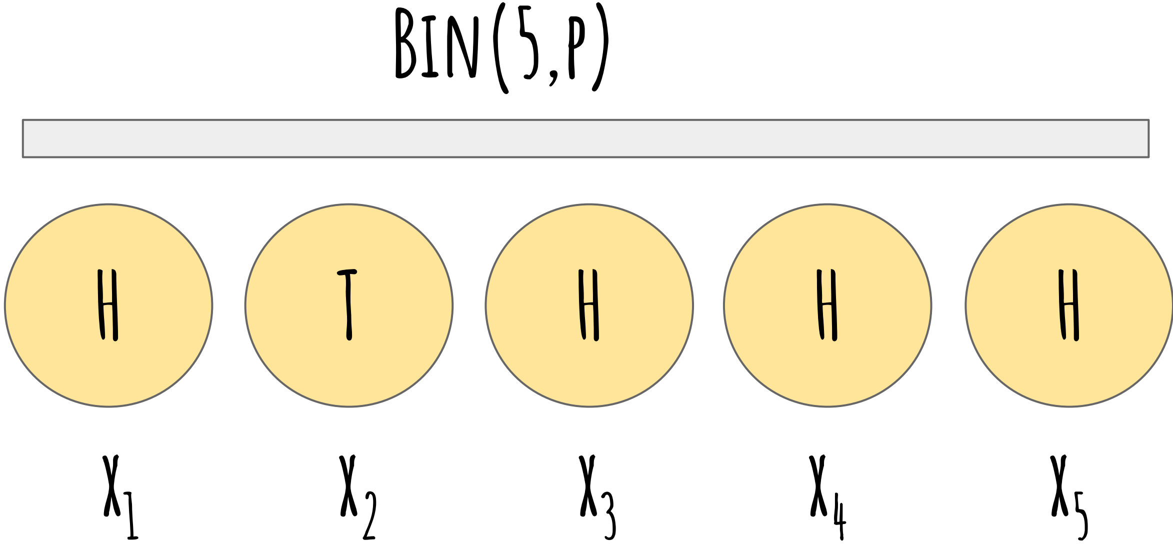
$$P(HTHHH) = p(1-p)ppp =$$

Binomial RV PMF



$$P(HTHHH) = p(1-p)ppp = p^4(1-p)^1$$

Binomial RV PMF



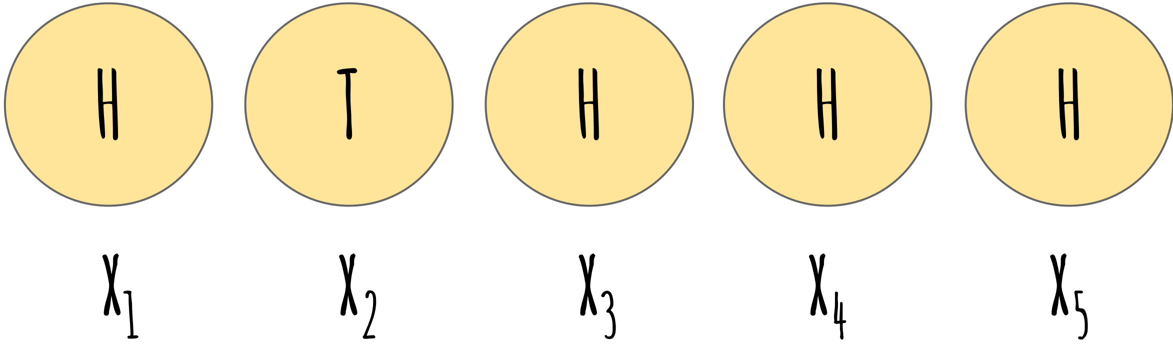
$$P(HTHHH) = p(1-p)ppp = p^4(1-p)^1$$

$$P(X = 4) =$$

Binomial RV PMF

HHHHT

Bin(5,p)



$$P(HTHHH) = p(1-p)ppp = p^4(1-p)^1$$

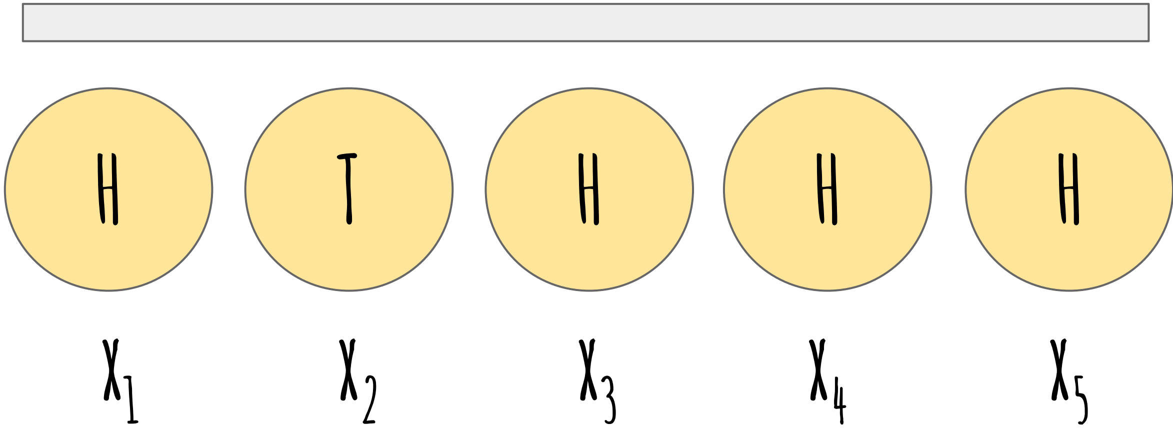
$$P(X = 4) =$$

Binomial RV PMF

HHHHT

HHHTH

$\mathrm{Bin}(5, p)$



$$P(HTHHH) = p(1-p)ppp = p^4(1-p)^1$$

$$P(X = 4) =$$

Binomial RV PMF

$\text{Bin}(5, p)$

HHHHT
HHHTH
HHTHH
HTHHH
THHHH

H	T	H	H	H
X_1	X_2	X_3	X_4	X_5

$$P(HTHHH) = p(1-p)ppp = p^4(1-p)^1$$

$$P(X = 4) =$$

Binomial RV PMF

HHHHT
HHHTH
HHTHH
HTHHH
THHHH

$\binom{5}{4}$ ways to get 4 heads out of 5

Bin(5,p)

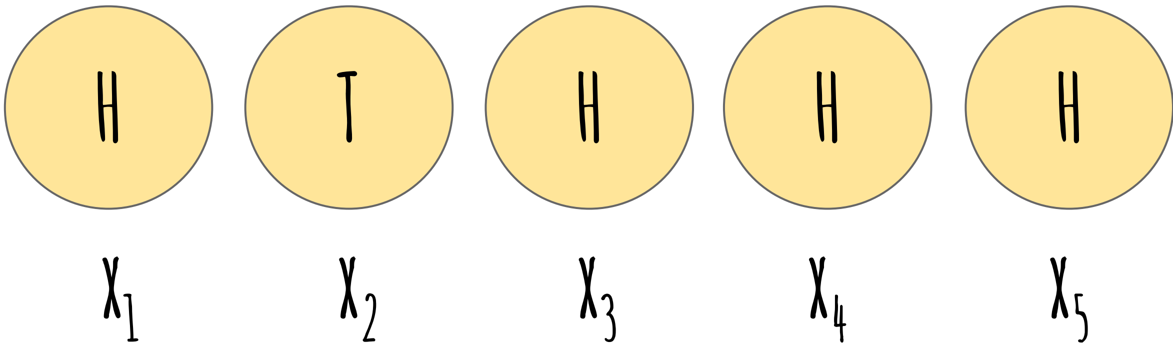
H	T	H	H	H
X_1	X_2	X_3	X_4	X_5

$$P(HTHHH) = p(1-p)ppp = p^4(1-p)^1$$

$$P(X=4) =$$

Binomial RV PMF

BIN(5,p)



HHHHT
HHHTH
HHTHH
HTHHH
THHHH

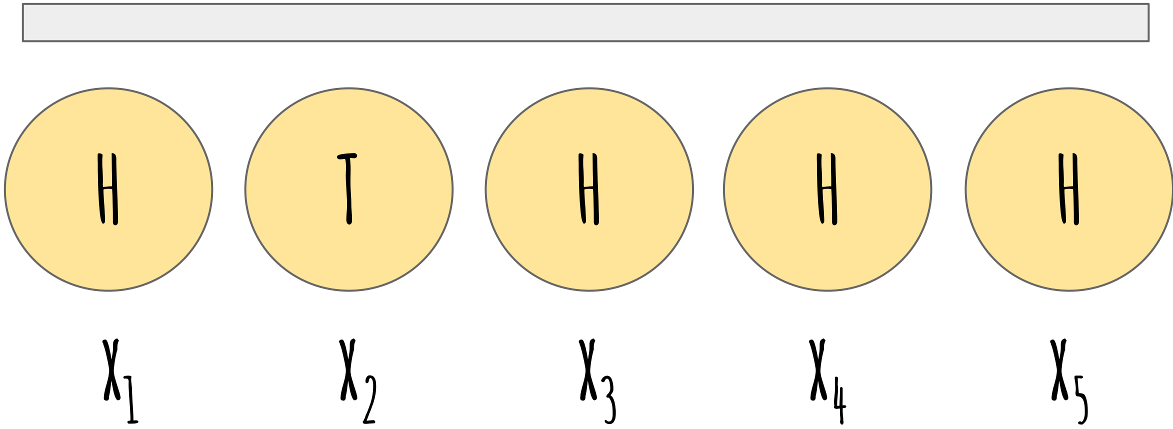
$$P(HTHHH) = p(1-p)ppp = p^4(1-p)^1$$

$\binom{5}{4}$ ways to get 4 heads out of 5

$$P(X=4) = \binom{5}{4} p^4 (1-p)^{5-4}$$

Binomial RV PMF

BIN(5,p)



HHHHT
HHHTH
HHTHH
HTHHH
THHHH

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$\binom{5}{4}$ ways to get 4 heads out of 5

$$P(X=4) = \binom{5}{4} p^4 (1-p)^{5-4}$$

Binomial RV PMF sums to 1?

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Binomial RV PMF sums to 1?

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

$$\sum_{k \in \Omega_X} p_X(k) =$$

Binomial RV PMF sums to 1?

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

$$\sum_{k \in \Omega_X} p_X(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} =$$

Binomial RV PMF sums to 1?

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

$$\sum_{k \in \Omega_X} p_X(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = \big(p + (1-p)\big)^n =$$

Binomial RV PMF sums to 1?

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

$$\sum_{k \in \Omega_X} p_X(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = \left(p + (1-p)\right)^n = 1^n = 1$$

Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) =$$

Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} =$$

Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = ???$$

Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = ???$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] =$$

Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = ???$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] =$$

LoE regardless of independence

Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = ???$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np$$

LoE regardless of independence

Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = ???$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np$$

LoE regardless of independence

Accept variance for now... But if X, Y are "independent",
$Var(X+Y) = Var(X) + Var(Y)$. Recall if $X_i \sim Ber(p)$, $Var(X_i) = p(1-p)$.

Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = ???$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np$$

LoE regardless of independence

Accept variance for now... But if X, Y are "independent",
$Var(X+Y) = Var(X) + Var(Y)$. Recall if $X_i \sim Ber(p)$, $Var(X_i) = p(1-p)$.

$$Var(X) = np(1-p)$$

The Binomial RV

Binomial RV: $X \sim Bin(n, p)$ if and only if X has the following pmf:

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

X is the sum of n independent $Ber(p)$ random variables.

The Binomial RV

<u>Binomial RV:</u> $X \sim Bin(n, p)$ if and only if X has the following pmf:

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

X is the sum of n independent $Ber(p)$ random variables.

$$E[X] = np \qquad Var(X) = np(1-p)$$

For example, X is the number of heads in n independent coin flips with $P(\text{head}) = p$. Notice that $Bin(1, p) \equiv Ber(p)$. By definition, if X, Y are independent with $X \sim Bin(n, p)$ and $Y \sim Bin(m, p)$, then $X + Y \sim Bin(n + m, p)$.

Binomial RV Example

A factory produces 100 cars per day, but a car is defective with probability 0.02. What's the probability the factory produces 2 or more defective cars?

Binomial RV Example

A factory produces 100 cars per day, but a car is defective with probability 0.02. What's the probability the factory produces 2 or more defective cars?

Let X be the number of defective cars it produces. $X \sim Bin(100, 0.02)$

Binomial RV Example

A factory produces 100 cars per day, but a car is defective with probability 0.02. What's the probability the factory produces 2 or more defective cars?

Let X be the number of defective cars it produces. $X \sim Bin(100, 0.02)$, so

$$P(X \geq 2) =$$

Binomial RV Example

A factory produces 100 cars per day, but a car is defective with probability 0.02. What's the probability the factory produces 2 or more defective cars?

Let X be the number of defective cars it produces. $X \sim Bin(100, 0.02)$, so

$$P(X \geq 2) = 1 - P(X < 2) =$$

Binomial RV Example

A factory produces 100 cars per day, but a car is defective with probability 0.02. What's the probability the factory produces 2 or more defective cars?

Let X be the number of defective cars it produces. $X \sim Bin(100, 0.02)$, so

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1)$$

Binomial RV Example

A factory produces 100 cars per day, but a car is defective with probability 0.02. What's the probability the factory produces 2 or more defective cars?

Let X be the number of defective cars it produces. $X \sim Bin(100, 0.02)$, so

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1)$$
$$= 1 - \binom{100}{0}(0.02)^0(1 - 0.02)^{100} - \binom{100}{1}(0.02)^1(1 - 0.02)^{99}$$

Binomial RV Example

A factory produces 100 cars per day, but a car is defective with probability 0.02. What's the probability the factory produces 2 or more defective cars?

Let X be the number of defective cars it produces. $X \sim Bin(100, 0.02)$, so

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1)$$
$$= 1 - \binom{100}{0}(0.02)^0(1-0.02)^{100} - \binom{100}{1}(0.02)^1(1-0.02)^{99}$$
$$\approx 0.5967$$

Probability

3.5 Zoo of Discrete RV's Part II

Alex Tsun

Agenda

- The Uniform RV
- The Geometric RV
- The Negative Binomial RV

The Uniform RV

Uniform (Discrete) RV: $X \sim Unif(a, b)$ where $a < b$ are integers, if and only if X has the following pmf:

$$p_X(k) = \begin{cases} \dfrac{1}{b - a + 1}, & k \in \{a, a + 1, \dots, b\} \\ 0, & \text{otherwise} \end{cases}$$

The Uniform RV

Uniform (Discrete) RV: $X \sim Unif(a, b)$ where $a < b$ are integers, if and only if X has the following pmf:

$$p_X(k) = \begin{cases} \dfrac{1}{b - a + 1}, & k \in \{a, a + 1, \dots, b\} \\ 0, & \text{otherwise} \end{cases}$$

X is equally likely to take on any value in $\Omega_X = \{a, a + 1, \dots, b\}$ of size $b - a + 1$.

$$E[X] = \frac{a + b}{2} \qquad Var(X) = \frac{(b - a)(b - a + 2)}{12}$$

For example, a roll of a fair 6-sided die is $Unif(1,6)$.

The Geometric RV

$\text{Geo}(p) = 3$

T T H

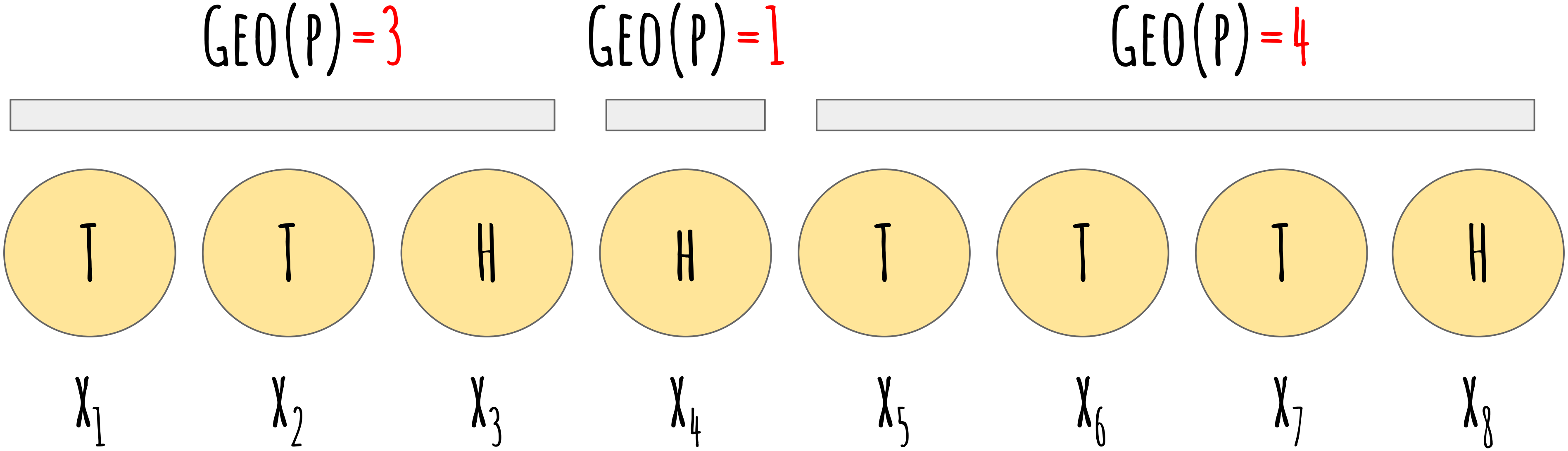
X_1 X_2 X_3

The Geometric RV

$Geo(p) = 3$ $Geo(p) = 1$

T	T	H	H
X_1	X_2	X_3	X_4

The Geometric RV



Geometric RV PMF

Geo(p) = 4

T T T H

$P(X = 4) =$

Geometric RV PMF

Geo(p) = 4

T T T H

$$P(X = 4) = P(TTTH) =$$

Geometric RV PMF

Geo(p) = 4

T T T H

$$P(X = 4) = P(TTTH) = (1 - p)$$

Geometric RV PMF

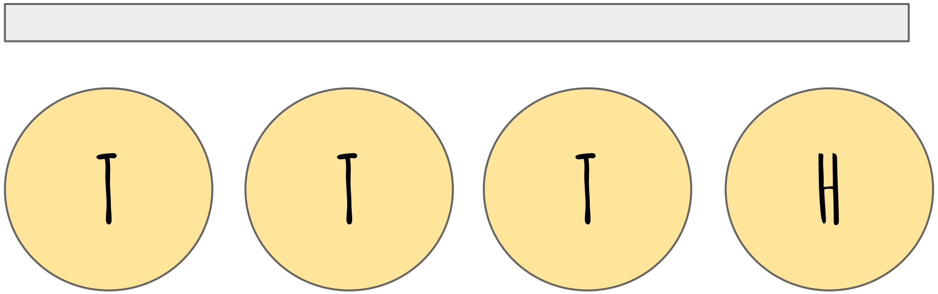
Geo(p) = 4

T T T H

$$P(X = 4) = P(TTTH) = (1 - p)(1 - p)$$

Geometric RV PMF

Geo(p) = 4



$$P(X = 4) = P(TTTH) = (1 - p)(1 - p)(1 - p)p =$$

Geometric RV PMF

Geo(p) = 4

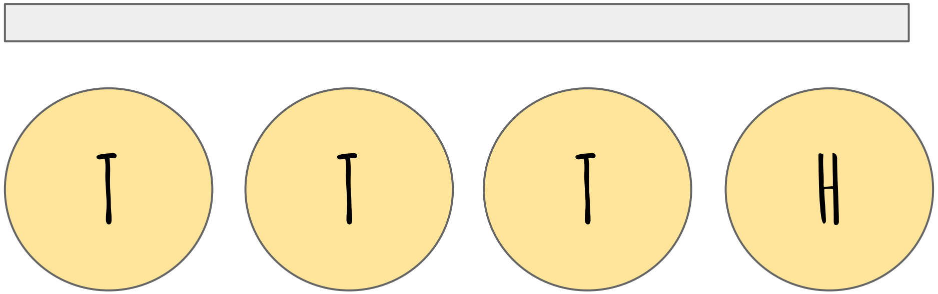
T T T H

$$P(X=4) = P(TTTH) = (1-p)(1-p)(1-p)p = (1-p)^{4-1}p$$

Geometric RV PMF

$$p_X(k) = (1-p)^{k-1}p$$

Geo(p) = 4



$$P(X=4) = P(TTTH) = (1-p)(1-p)(1-p)p = (1-p)^{4-1}p$$

Geometric RV PMF sums to 1?

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

Geometric RV PMF sums to 1?

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

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

Geometric RV PMF sums to 1?

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

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

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

Geometric RV PMF sums to 1?

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

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

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

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \qquad \text{for } -1 < r < 1$$

Geometric RV PMF sums to 1?

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

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

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

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \qquad \text{for } -1 < r < 1$$

$$= p \cdot \frac{1}{1-(1-p)}$$

Geometric RV PMF sums to 1?

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

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

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

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \qquad \text{for } -1 < r < 1$$

$$= p \cdot \frac{1}{1-(1-p)}$$

$$= \frac{p}{p} = 1$$

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

- 1/2?

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

- 1/2? $\rightarrow E[X] = 2$

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

- $1/2$? $\rightarrow E[X] = 2$
- $1/10$?

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

- 1/2? $\rightarrow E[X] = 2$
- 1/10? $\rightarrow E[X] = 10$

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

- $1/2$? $\rightarrow E[X] = 2$
- $1/10$? $\rightarrow E[X] = 10$
- $1/7$?

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

- $1/2$? $\rightarrow E[X] = 2$
- $1/10$? $\rightarrow E[X] = 10$
- $1/7$? $\rightarrow E[X] = 7$

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

- $1/2$? $\rightarrow E[X] = 2$
- $1/10$? $\rightarrow E[X] = 10$
- $1/7$? $\rightarrow E[X] = 7$

Intuitively, $E[X] = \frac{1}{p}$. We'll prove this later in an elegant way. See the next slide for brute force (requires calculus).

Geometric RV Properties

Suppose $X \sim Geo(p)$. What do you think $E[X]$ is if $P(head) =$

- 1/2? $\rightarrow E[X] = 2$
- 1/10? $\rightarrow E[X] = 10$
- 1/7? $\rightarrow E[X] = 7$

Intuitively, $E[X] = \frac{1}{p}$. We'll prove this later in an elegant way. See the next slide for brute force (requires calculus).

$Var(X) = \frac{1-p}{p^2}$ by messy algebra.

Geometric RV Expectation (ugly way)

$$E[X] = \sum_{k=1}^{\infty} k p_X(k) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p = \sum_{k=0}^{\infty} k(1-p)^{k-1}p$$

$$= p\sum_{k=0}^{\infty} k(1-p)^{k-1} = p\sum_{k=0}^{\infty} \frac{d}{dp}[-(1-p)^k] = -p \cdot \frac{d}{dp}\left[\sum_{k=0}^{\infty}(1-p)^k\right]$$

$$= -p \cdot \frac{d}{dp}\left[\sum_{k=0}^{\infty}(1-p)^k\right] = -p \cdot \frac{d}{dp}\left[\frac{1}{1-(1-p)}\right] = -p \cdot \frac{d}{dp}\left[\frac{1}{p}\right]$$

$$= -p\left[-\frac{1}{p^2}\right] = \frac{1}{p}$$

The Geometric RV

<u>Geometric RV:</u> $X \sim Geo(p)$ if and only if X has the following pmf:

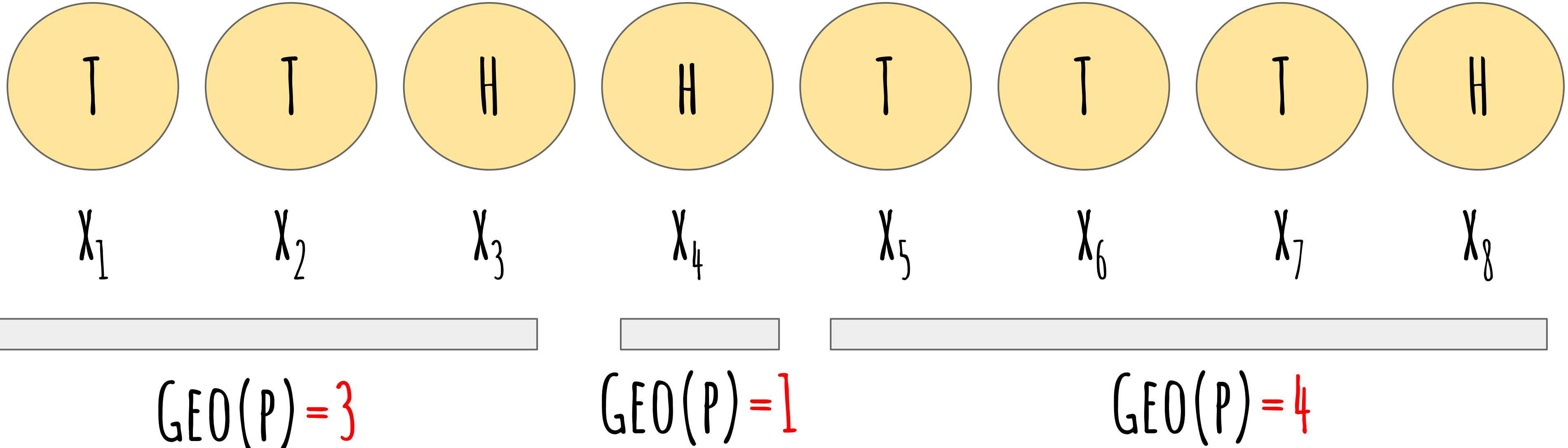
$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1,2,3,\ldots$$

$$E[X] = \frac{1}{p} \qquad Var(X) = \frac{1-p}{p^2}$$

For example, X is the number of independent coin flips with $P(\text{head}) = p$ up to and including the first head.

Random Picture

The Negative binomial RV



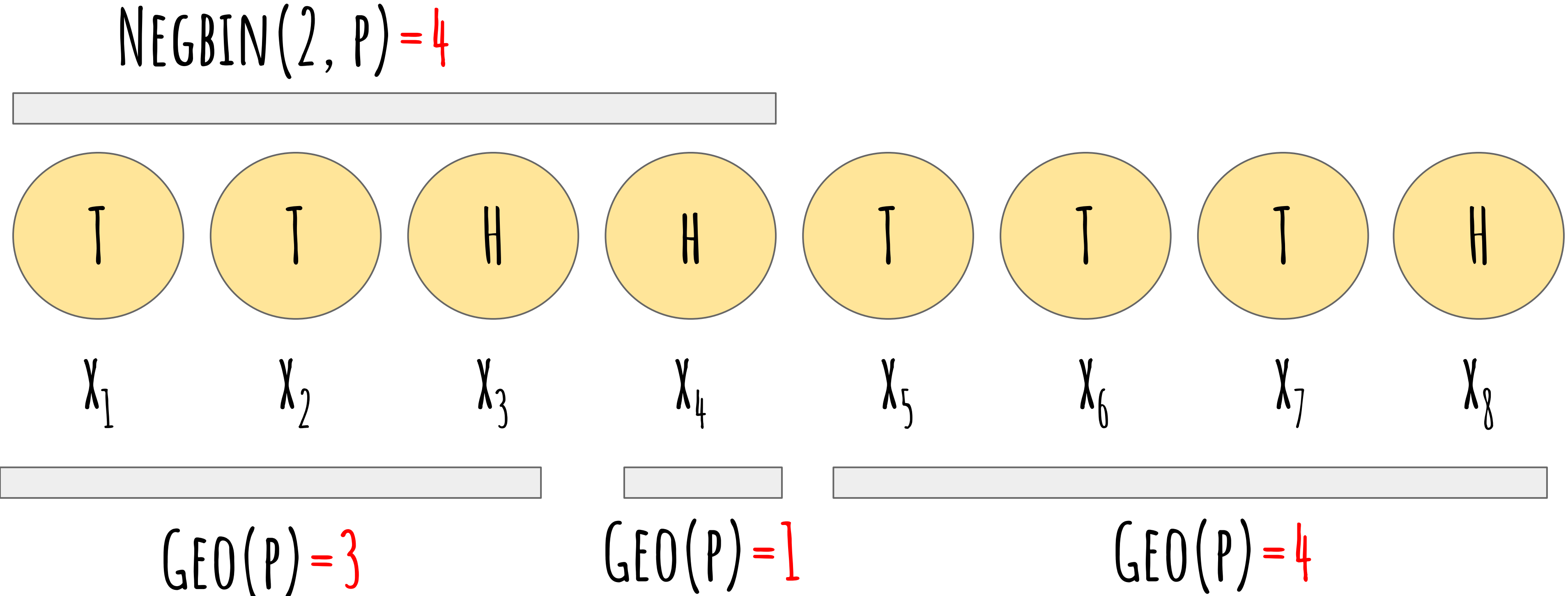
The Negative binomial RV

$\text{NegBin}(1, p) = 3$

T	T	H	H	T	T	T	H
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8

$\text{Geo}(p) = 3$ $\quad$ $\text{Geo}(p) = 1$ $\quad$ $\text{Geo}(p) = 4$

The Negative binomial RV



The Negative Binomial RV

$\text{NegBin}(3, p) = 8$

T	T	H	H	T	T	T	H
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8

$\text{Geo}(p) = 3$ $\text{Geo}(p) = 1$ $\text{Geo}(p) = 4$

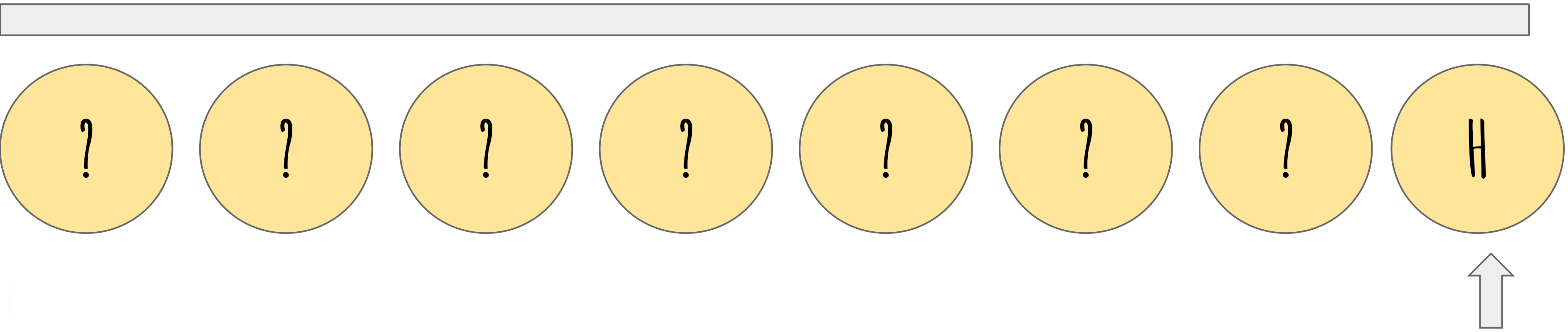
Negative Binomial RV PMF

$$X \sim NegBin(r = 3, p)$$

? ? ? ? ? ? ? ?

Negative Binomial RV PMF

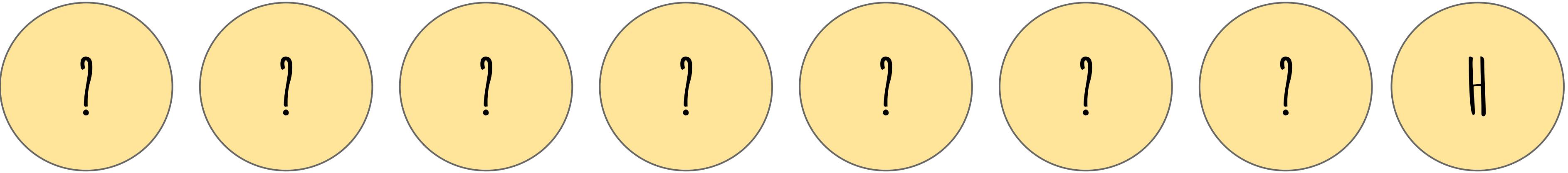
$$X \sim NegBin(r = 3, p)$$



Must be H

Negative Binomial RV PMF

$$X \sim NegBin(r = 3, p)$$



Exactly 2 $(r - 1)$ of the first 7 $(k - 1)$ must be heads.

Negative Binomial RV PMF

$X \sim NegBin(r = 3, p)$

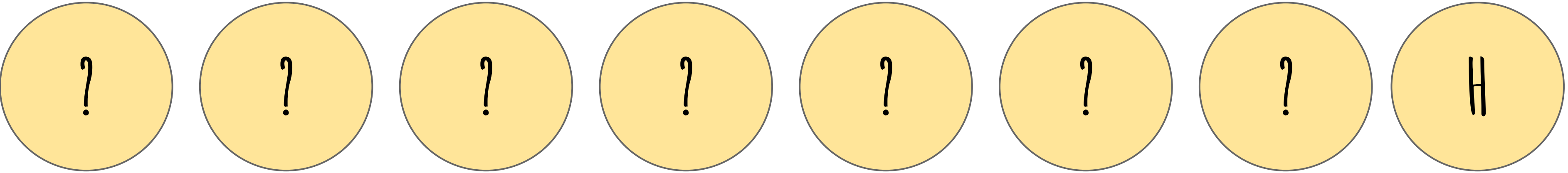
? ? ? ? ? ? ? H

Exactly 2 ($r - 1$) of the first 7 ($k - 1$) must be heads.

$$P(X = 8) = \binom{8-1}{3-1}$$

Negative Binomial RV PMF

$$X \sim NegBin(r = 3, p)$$

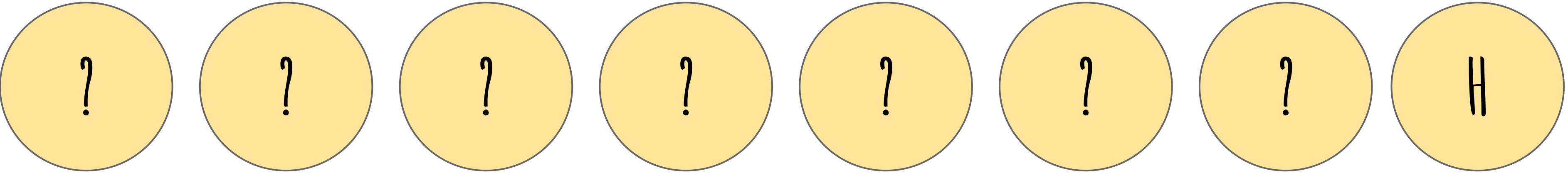


Exactly 2 ($r - 1$) of the first 7 ($k - 1$) must be heads.

$$P(X = 8) = \binom{8-1}{3-1} p^{3-1} (1-p)^5$$

Negative Binomial RV PMF

$$X \sim NegBin(r = 3, p)$$



Exactly 2 ($r - 1$) of the first 7 ($k - 1$) must be heads.

$$P(X = 8) = \binom{8-1}{3-1} p^{3-1} (1-p)^5 \cdot p$$

Negative Binomial RV PMF

$$X \sim NegBin(r = 3, p)$$

? ? ? ? ? ? ? H

Exactly 2 $(r-1)$ of the first 7 $(k-1)$ must be heads.

$$P(X = 8) = \binom{8-1}{3-1} p^{3-1} (1-p)^5 \cdot p$$

$$p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

Negative Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=r}^{\infty} k \binom{k-1}{r-1} p^k (1-p)^{k-r} = ???$$

Negative Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=r}^{\infty} k \binom{k-1}{r-1} p^k (1-p)^{k-r} = ???$$

Suppose $X_1, \ldots, X_r \sim Geo(p)$ are independent, so $X = \sum_{i=1}^{r} X_i$.

Negative Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=r}^{\infty} k \binom{k-1}{r-1} p^k (1-p)^{k-r} = ???$$

Suppose $X_1, \ldots, X_r \sim Geo(p)$ are independent, so $X = \sum_{i=1}^{r} X_i$.

$$E[X] = E\left[\sum_{i=1}^{r} X_i\right] = \sum_{i=1}^{r} E[X_i] = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}$$

Negative Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=r}^{\infty} k \binom{k-1}{r-1} p^k (1-p)^{k-r} = ???$$

Suppose $X_1, \ldots, X_r \sim Geo(p)$ are independent, so $X = \sum_{i=1}^{r} X_i$.

$$E[X] = E\left[\sum_{i=1}^{r} X_i\right] = \sum_{i=1}^{r} E[X_i] = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}$$

Accept variance for now… But if X, Y are "independent",
$Var(X+Y) = Var(X) + Var(Y)$. Recall if $X_i \sim Geo(p)$, $Var(X_i) = \frac{1-p}{p^2}$.

Negative Binomial RV Properties

$$E[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k=r}^{\infty} k \binom{k-1}{r-1} p^k (1-p)^{k-r} = ???$$

Suppose $X_1, \dots, X_r \sim Geo(p)$ are independent, so $X = \sum_{i=1}^{r} X_i$.

$$E[X] = E\left[\sum_{i=1}^{r} X_i\right] = \sum_{i=1}^{r} E[X_i] = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}$$

Accept variance for now… But if X, Y are "independent",
$Var(X+Y) = Var(X) + Var(Y)$. Recall if $X_i \sim Geo(p)$, $Var(X_i) = \frac{1-p}{p^2}$.

$$Var(X) = \frac{r(1-p)}{p^2}$$

The Negative Binomial RV

Negative Binomial RV: $X \sim NegBin(r, p)$ if and only if X has the following pmf:

$$p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \qquad k = r, r+1, \ldots$$

X is the sum of r independent $Geo(p)$ random variables.

The Negative Binomial RV

Negative Binomial RV: $X \sim NegBin(r,p)$ if and only if X has the following pmf:

$$p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \qquad k = r, r+1, \ldots$$

X is the sum of r independent $Geo(p)$ random variables.

$$E[X] = \frac{r}{p} \qquad Var(X) = \frac{r(1-p)}{p^2}$$

For example, X is the number of independent coin flips with $P(\text{head}) = p$ up to and including the r^{th} head. Notice that $NegBin(1,p) \equiv Geo(p)$. By definition, if X, Y are independent with $X \sim NegBin(r,p)$ and $Y \sim NegBin(s,p)$, then $X + Y \sim NegBin(r+s,p)$.

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn $\$2^k$ (i.e., if your first head was the third flip, you would earn $\$8$). How much would you pay to play this game?

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn $\$2^k$ (i.e., if your first head was the third flip, you would earn $\$8$). How much would you pay to play this game?

$$X \sim Geo\left(\frac{1}{2}\right)$$

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn \$$2^k$ (i.e., if your first head was the third flip, you would earn \$8). How much would you pay to play this game?

$X \sim Geo\left(\frac{1}{2}\right)$ $\quad p_X(k) = \left(1 - \frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \frac{1}{2^k},$ $\quad k = 1,2,3,\ldots$

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn $\$2^k$ (i.e., if your first head was the third flip, you would earn $\$8$). How much would you pay to play this game?

$$X \sim Geo\left(\frac{1}{2}\right) \qquad p_X(k) = \left(1 - \frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \frac{1}{2^k}, \qquad k = 1,2,3,\ldots$$

$$E[2^X] \neq 2^{E[X]} = 2^2 = 4$$

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn $\$2^k$ (i.e., if your first head was the third flip, you would earn $8). How much would you pay to play this game?

$$X \sim Geo\left(\frac{1}{2}\right) \qquad p_X(k) = \left(1 - \frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \frac{1}{2^k}, \qquad k = 1,2,3,\ldots$$

$$E[2^X] \neq 2^{E[X]} = 2^2 = 4$$

$$E[2^X] =$$

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn $\$2^k$ (i.e., if your first head was the third flip, you would earn $\$8$). How much would you pay to play this game?

$$X \sim Geo\left(\frac{1}{2}\right) \qquad p_X(k) = \left(1 - \frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \frac{1}{2^k}, \qquad k = 1,2,3,\ldots$$

$$E[2^X] \neq 2^{E[X]} = 2^2 = 4$$

LOTUS $$E[2^X] = \sum_{k=1}^{\infty} 2^k p_X(k) =$$

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn $\$2^k$ (i.e., if your first head was the third flip, you would earn $8). How much would you pay to play this game?

$$X \sim Geo\left(\frac{1}{2}\right) \qquad p_X(k) = \left(1 - \frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \frac{1}{2^k}, \qquad k = 1,2,3,\ldots$$

$$E[2^X] \neq 2^{E[X]} = 2^2 = 4$$

LOTUS

$$E[2^X] = \sum_{k=1}^{\infty} 2^k p_X(k) = \sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k} =$$

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn \$$2^k$ (i.e., if your first head was the third flip, you would earn \$8). How much would you pay to play this game?

$$X \sim Geo\left(\frac{1}{2}\right) \qquad p_X(k) = \left(1 - \frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \frac{1}{2^k}, \qquad k = 1,2,3,\ldots$$

$$E[2^X] \neq 2^{E[X]} = 2^2 = 4$$

LOTUS

$$E[2^X] = \sum_{k=1}^{\infty} 2^k p_X(k) = \sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k} = \sum_{k=1}^{\infty} 1 =$$

Geometric RV Example

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn $\$2^k$ (i.e., if your first head was the third flip, you would earn \$8). How much would you pay to play this game?

$$X \sim Geo\left(\frac{1}{2}\right) \qquad p_X(k) = \left(1 - \frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \frac{1}{2^k}, \qquad k = 1,2,3,\ldots$$

$$E[2^X] \neq 2^{E[X]} = 2^2 = 4$$

LOTUS

$$E[2^X] = \sum_{k=1}^{\infty} 2^k p_X(k) = \sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k} = \sum_{k=1}^{\infty} 1 = \infty$$

So theoretically, any finite amount since your expected winnings is infinite!

Probability

3.6 Zoo of Discrete RV's Part III

Alex Tsun

Agenda

- The Poisson RV
- The Poisson Process
- The Hypergeometric RV

The Poisson RV (Motivation)

No random variable so far for events in a unit time:

The Poisson RV (Motivation)

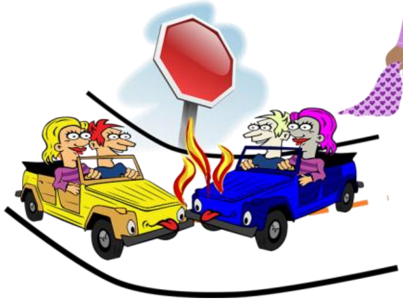
No random variable so far for events in a unit time:

- How many babies born in the next minute?

The Poisson RV (Motivation)

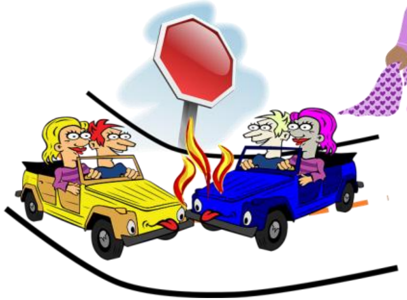
No random variable so far for events in a unit time:

- How many babies born in the next minute?
- How many car crashes happen per hour?



The Poisson RV (Motivation)

No random variable so far for events in a unit time:



- How many babies born in the next minute?
- How many car crashes happen per hour?

If use binomial, what to choose for n? There is no upper bound!

The Poisson RV (Idea)

Let's say we want to model babies born in the next minute, if the historical average is 2 babies/min.

One Unit of Time

The Poisson RV (Idea)

Let's say we want to model babies born in the next minute, if the historical average is 2 babies/min.

____ One Unit of Time

_______ ________ _______ ________ _______ $Bin(n = \;\; 5, \;\; p = \qquad)$

The Poisson RV (Idea)

Let's say we want to model babies born in the next minute, if the historical average is 2 babies/min.

One Unit of Time

$Bin(n = 5, \; p = 2/5)$

The Poisson RV (Idea)

Let's say we want to model babies born in the next minute, if the historical average is 2 babies/min.

One Unit of Time

$Bin(n = \;\; 5, \;\; p = 2/5)$

$Bin(n = 10, \;\; p = \qquad)$

The Poisson RV (Idea)

Let's say we want to model babies born in the next minute, if the historical average is 2 babies/min.

One Unit of Time

$Bin(n = 5, \ p = 2/5)$

$Bin(n = 10, \ p = 2/10)$

The Poisson RV (Idea)

Let's say we want to model babies born in the next minute, if the historical average is 2 babies/min.

One Unit of Time

$Bin(n = \;\; 5, \;\; p = 2/5)$

$Bin(n = 10, \;\; p = 2/10)$

$Bin(n = 70, \;\; p = \qquad)$

The Poisson RV (Idea)

Let's say we want to model babies born in the next minute, if the historical average is 2 babies/min.

One Unit of Time

$Bin(n = 5, \; p = 2/5)$

$Bin(n = 10, \; p = 2/10)$

$Bin(n = 70, \; p = 2/70)$

The Poisson RV PMF

Let λ be the historical average number of events per unit of time.

Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

The Poisson RV PMF

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n \to \infty} X_n = Poi(\lambda)$.

The Poisson RV PMF

Let λ be the historical average number of events per unit of time.

Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$p_Y(k) =$

The Poisson RV PMF

Let λ be the historical average number of events per unit of time.

Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$p_Y(k) = \lim_{n\to\infty} p_{X_n}(k) = \lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} = \lim_{n\to\infty} \frac{n!}{k!\,(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \lim_{n\to\infty} \frac{n!}{k!\,(n-k)!} \frac{\lambda^k}{n^k} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k} = \frac{\lambda^k}{k!} \cdot \lim_{n\to\infty} \frac{n!}{(n-k)!} \frac{1}{n^k} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

$$= \frac{\lambda^k}{k!} \cdot \lim_{n\to\infty} \left(\frac{n}{n} \cdot \frac{n-1}{n} \cdot \ldots \cdot \frac{n-k+1}{n}\right) \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k} = \frac{\lambda^k}{k!} \cdot \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = \frac{\lambda^k}{k!} e^{-\lambda}$$

Poisson RV PMF sums to 1?

Recall Taylor series for e^x:

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

Poisson RV PMF sums to 1?

Recall Taylor series for e^x:

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

$$\sum_{k=0}^{\infty} p_Y(k) =$$

Poisson RV PMF sums to 1?

Recall Taylor series for e^x:

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

$$\sum_{k=0}^{\infty} p_Y(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} =$$

Poisson RV PMF sums to 1?

Recall Taylor series for e^x:

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

$$\sum_{k=0}^{\infty} p_Y(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} =$$

Poisson RV PMF sums to 1?

Recall Taylor series for e^x:

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

$$\sum_{k=0}^{\infty} p_Y(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} =$$

Taylor Series

Poisson RV PMF sums to 1?

Recall Taylor series for e^x:

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

$$\sum_{k=0}^{\infty} p_Y(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

Taylor Series

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$

$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$

$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] =$$

Poisson Rv Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$
$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] = E\left[\lim_{n\to\infty} X_n\right] =$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$
$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] = E\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} E[X_n] =$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$
$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] = E\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \lambda = \lambda$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$
$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] = E\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \lambda = \lambda$$

$$Var(Y) =$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$

$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] = E\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \lambda = \lambda$$

$$Var(Y) = Var\left(\lim_{n\to\infty} X_n\right) =$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$

$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] = E\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \lambda = \lambda$$

$$Var(Y) = Var\left(\lim_{n\to\infty} X_n\right) = \lim_{n\to\infty} Var(X_n) =$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$
$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] = E\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \lambda = \lambda$$
$$Var(Y) = Var\left(\lim_{n\to\infty} X_n\right) = \lim_{n\to\infty} Var(X_n) = \lim_{n\to\infty} \lambda\left(1 - \frac{\lambda}{n}\right) =$$

Poisson RV Properties

Let λ be the historical average number of events per unit of time.
Send $n \to \infty$ in such a way that $np = \lambda$ is fixed (i.e., $p = \lambda/n$).

Let $X_n \sim Bin\left(n, \frac{\lambda}{n}\right)$ and $Y \sim \lim_{n\to\infty} X_n = Poi(\lambda)$.

$$E[X_n] = np = \lambda$$
$$Var(X_n) = np(1-p) = \lambda\left(1 - \frac{\lambda}{n}\right)$$

$$E[Y] = E\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \lambda = \lambda$$
$$Var(Y) = Var\left(\lim_{n\to\infty} X_n\right) = \lim_{n\to\infty} Var(X_n) = \lim_{n\to\infty} \lambda\left(1 - \frac{\lambda}{n}\right) = \lambda$$

The Poisson RV

<u>**Poisson RV:**</u> $X \sim Poi(\lambda)$ if and only if X has the following pmf:

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0,1,2,\ldots$$

If λ is the historical average of events per unit of time, then X is the number of events that occur in a unit of time.

The Poisson RV

Poisson RV: $X \sim Poi(\lambda)$ if and only if X has the following pmf:

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0,1,2,\ldots$$

If λ is the historical average of events per unit of time, then X is the number of events that occur in a unit of time.

$$E[X] = \lambda \qquad Var(X) = \lambda$$

For example, X is the number of babies born in a minute if on average λ babies are born per minute. By definition, if X, Y are independent with $X \sim Poi(\lambda)$ and $Y \sim Poi(\mu)$, then $X + Y \sim Poi(\lambda + \mu)$.

Random Picture

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

Convert λ's to the same unit of interest. For us, it's a minute.

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

Convert λ's to the same unit of interest. For us, it's a minute.

$X \sim Poi(2 \text{ users/min})$ $\qquad$ $Y \sim Poi(3 \text{ users/min})$

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

Convert λ's to the same unit of interest. For us, it's a minute.

$$X \sim Poi(2 \text{ users/min}) \qquad Y \sim Poi(3 \text{ users/min})$$

$$Z = X + Y \sim Poi(2+3) = Poi(5)$$

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

Convert λ's to the same unit of interest. For us, it's a minute.

$$X \sim Poi(2 \text{ users/min}) \qquad Y \sim Poi(3 \text{ users/min})$$

$$Z = X + Y \sim Poi(2 + 3) = Poi(5)$$

$$P(Z < 2) =$$

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

Convert λ's to the same unit of interest. For us, it's a minute.

$X \sim Poi(2 \text{ users/min})$ $\qquad$ $Y \sim Poi(3 \text{ users/min})$

$$Z = X + Y \sim Poi(2+3) = Poi(5)$$

$$P(Z < 2) = p_Z(0) + p_Z(1) =$$

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

Convert λ's to the same unit of interest. For us, it's a minute.

$$X \sim Poi(2 \text{ users/min}) \qquad Y \sim Poi(3 \text{ users/min})$$

$$Z = X + Y \sim Poi(2+3) = Poi(5)$$

$$P(Z < 2) = p_Z(0) + p_Z(1) = e^{-5}\frac{5^0}{0!} + e^{-5}\frac{5^1}{1!} =$$

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

Convert λ's to the same unit of interest. For us, it's a minute.

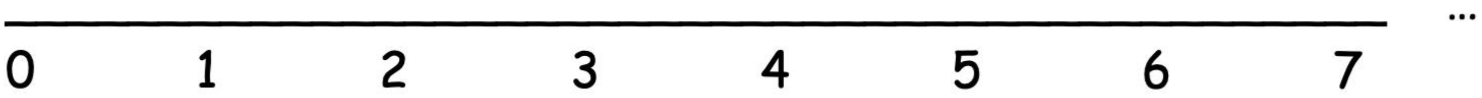
$$X \sim Poi(2 \text{ users/min}) \qquad Y \sim Poi(3 \text{ users/min})$$

$$Z = X + Y \sim Poi(2+3) = Poi(5)$$

$$P(Z < 2) = p_Z(0) + p_Z(1) = e^{-5}\frac{5^0}{0!} + e^{-5}\frac{5^1}{1!} = 6e^{-5} \approx 0.04$$

The Poisson Process

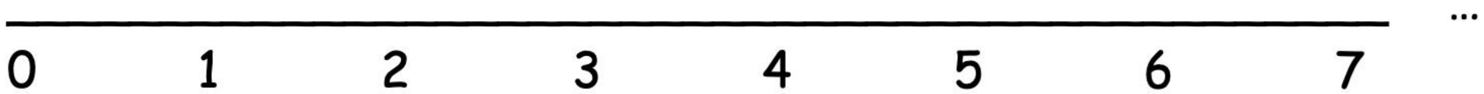
Poisson Process: A Poisson process with rate $\lambda > 0$ per unit of time, is a continuous-time process indexed by $t \in [0, \infty)$, so that $X(t)$ is the number of events that happened in the interval $[0, t]$. Notice that if $t_1 < t_2$, then $X(t_2) - X(t_1)$ is the number of events in $(t_1, t_2]$. The process has three properties:



The Poisson Process

Poisson Process: A Poisson process with rate $\lambda > 0$ per unit of time, is a continuous-time process indexed by $t \in [0, \infty)$, so that $X(t)$ is the number of events that happened in the interval $[0, t]$. Notice that if $t_1 < t_2$, then $X(t_2) - X(t_1)$ is the number of events in $(t_1, t_2]$. The process has three properties:

- $X(0) = 0$



The Poisson Process

<u>Poisson Process:</u> A Poisson process with rate $\lambda > 0$ per unit of time, is a continuous-time process indexed by $t \in [0, \infty)$, so that $X(t)$ is the number of events that happened in the interval $[0, t]$. Notice that if $t_1 < t_2$, then $X(t_2) - X(t_1)$ is the number of events in $(t_1, t_2]$. The process has three properties:

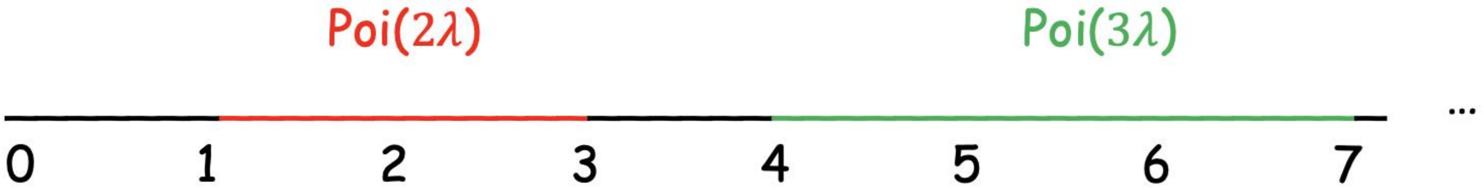
- $X(0) = 0$

- The number of events happening in any two disjoint intervals $[a, b]$ and $[c, d]$ are independent.

The Poisson Process

Poisson Process: A Poisson process with rate $\lambda > 0$ per unit of time, is a continuous-time process indexed by $t \in [0, \infty)$, so that $X(t)$ is the number of events that happened in the interval $[0, t]$. Notice that if $t_1 < t_2$, then $X(t_2) - X(t_1)$ is the number of events in $(t_1, t_2]$. The process has three properties:

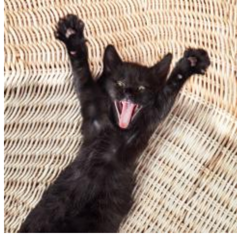
- $X(0) = 0$



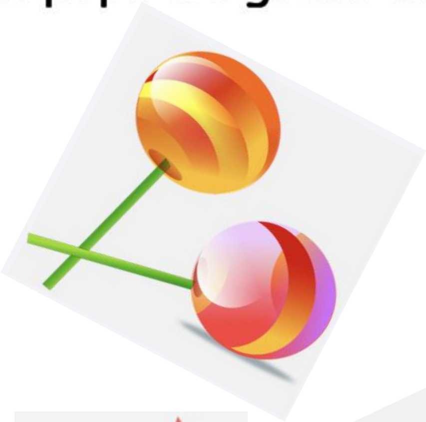
- The number of events happening in any two disjoint intervals $[a, b]$ and $[c, d]$ are independent.

- The number of events in any time interval $[t_1, t_2]$ is $Poi(\lambda(t_2 - t_1))$.

Hypergeometric RV PMF

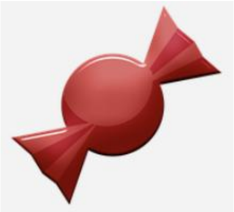


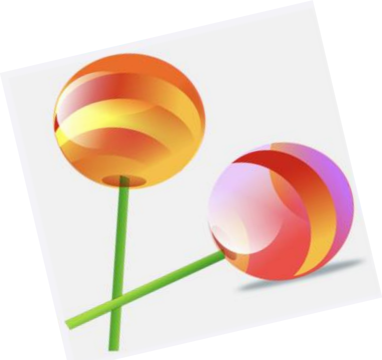
Suppose there is a candy bag of $N = 9$ total candies, $K = 4$ of which are lollipops. Our parents allow us grab $n = 3$ of them. Let X be the number of lollipops we grab. What is the probability that we get exactly 2 lollipops?

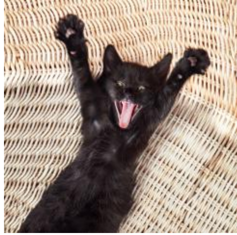


$$P(X = 2) =$$

$$p_X(k) =$$

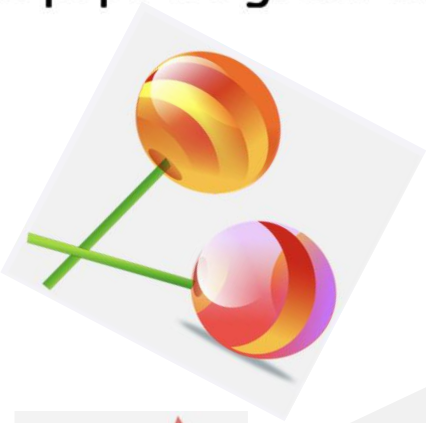






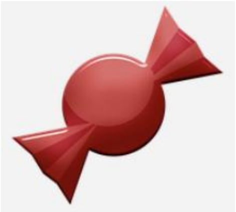
Hypergeometric RV PMF

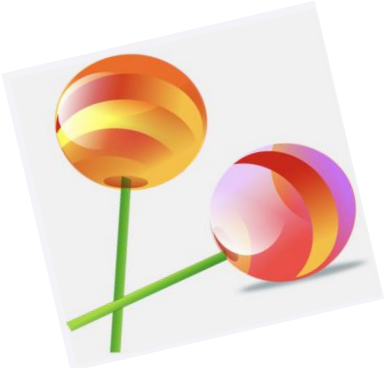
Suppose there is a candy bag of $N = 9$ total candies, $K = 4$ of which are lollipops. Our parents allow us grab $n = 3$ of them. Let X be the number of lollipops we grab. What is the probability that we get exactly 2 lollipops?



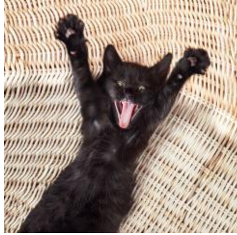
$$P(X = 2) = \frac{\quad}{\binom{9}{3}}$$

$$p_X(k) = \frac{\quad}{\binom{N}{n}}$$

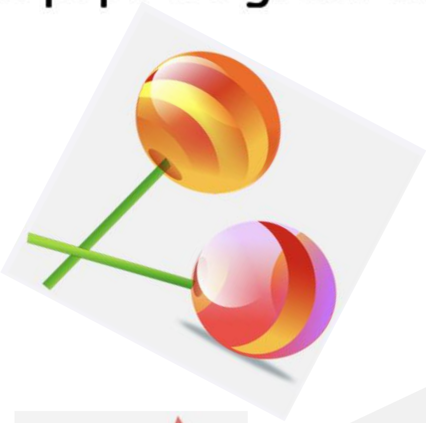




Hypergeometric RV PMF

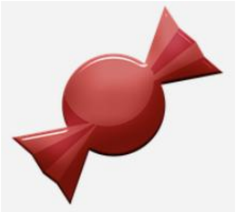


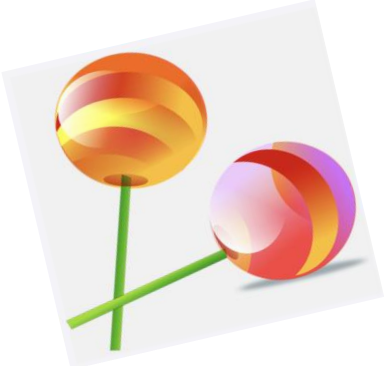
Suppose there is a candy bag of $N = 9$ total candies, $K = 4$ of which are lollipops. Our parents allow us grab $n = 3$ of them. Let X be the number of lollipops we grab. What is the probability that we get exactly 2 lollipops?



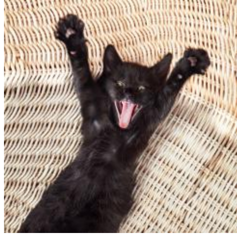
$$P(X = 2) = \frac{\binom{4}{2}}{\binom{9}{3}}$$

$$p_X(k) = \frac{\binom{K}{k}}{\binom{N}{n}}$$

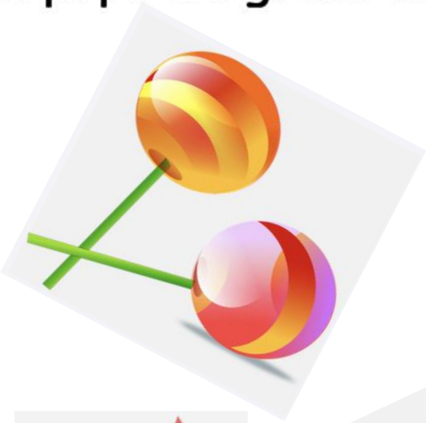




Hypergeometric RV PMF

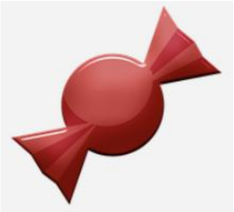


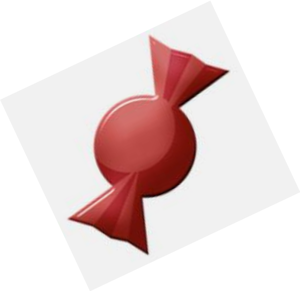
Suppose there is a candy bag of $N = 9$ total candies, $K = 4$ of which are lollipops. Our parents allow us grab $n = 3$ of them. Let X be the number of lollipops we grab. What is the probability that we get exactly 2 lollipops?

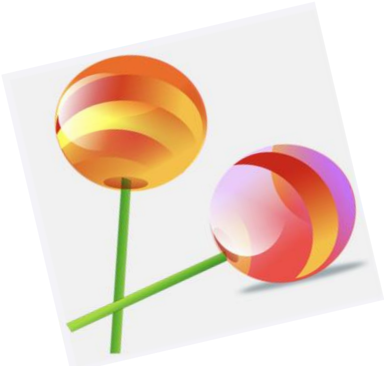


$$P(X = 2) = \frac{\binom{4}{2}\binom{5}{1}}{\binom{9}{3}}$$

$$p_X(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$







The Hypergeometric RV

Hypergeometric RV: $X \sim HypGeo(N, K, n)$ if and only if X has the following pmf:

$$p_X(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \qquad k = \max\{0, n+K-N\}, \ldots, \min\{K, n\}$$

For example, X is the number of successes when drawing n items without replacement from a bag containing N items, K of which are successes (hence $N - K$ failures).

The Hypergeometric RV

Hypergeometric RV: $X \sim HypGeo(N, K, n)$ if and only if X has the following pmf:

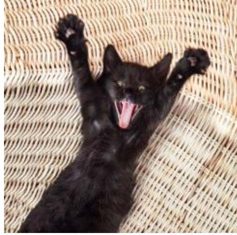
$$p_X(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \qquad k = \max\{0, n + K - N\}, \dots, \min\{K, n\}$$

For example, X is the number of successes when drawing n items without replacement from a bag containing N items, K of which are successes (hence $N - K$ failures).

$$E[X] = n\frac{K}{N} \qquad Var(X) = \text{un-important}$$

If we drew with replacement, then we would model this situation using $Bin\left(n, \frac{K}{N}\right)$.

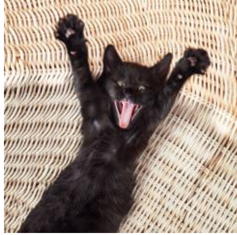
Hypergeometric RV Properties



Suppose $X \sim HypGeo(N, K, n)$.

Let $X_1, \ldots, X_n$ be indicator RV's (not independent) so that $X_i = 1$ if we got a lollipop on the i^{th} draw, and 0 otherwise. So $X = \sum_{i=1}^{n} X_i$.

Hypergeometric RV Properties

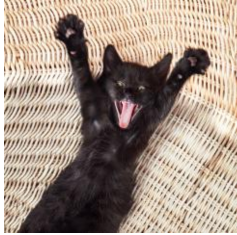


Suppose $X \sim HypGeo(N, K, n)$.

Let $X_1, \ldots, X_n$ be indicator RV's (not independent) so that $X_i = 1$ if we got a lollipop on the i^{th} draw, and 0 otherwise. So $X = \sum_{i=1}^{n} X_i$.

$$P(X_1 = 1) = \frac{K}{N}$$

Hypergeometric RV Properties



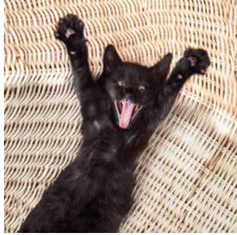
Suppose $X \sim HypGeo(N, K, n)$.

Let $X_1, \ldots, X_n$ be indicator RV's (not independent) so that $X_i = 1$ if we got a lollipop on the i^{th} draw, and 0 otherwise. So $X = \sum_{i=1}^{n} X_i$.

$$P(X_1 = 1) = \frac{K}{N}$$

$$P(X_2 = 1) = P(X_2 = 1 | X_1 = 1) P(X_1 = 1) + P(X_2 = 1 | X_1 = 0) P(X_1 = 0)$$
$$= \frac{K-1}{N-1} \cdot \frac{K}{N} + \frac{K}{N-1} \cdot \frac{N-K}{N} = \frac{K(N-1)}{N(N-1)} = \frac{K}{N}$$

Hypergeometric RV Properties



Suppose $X \sim HypGeo(N, K, n)$.

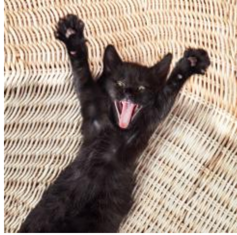
Let $X_1, \dots, X_n$ be indicator RV's (not independent) so that $X_i = 1$ if we got a lollipop on the i^{th} draw, and 0 otherwise. So $X = \sum_{i=1}^{n} X_i$.

$$P(X_1 = 1) = \frac{K}{N}$$

$$P(X_2 = 1) = P(X_2 = 1 | X_1 = 1) P(X_1 = 1) + P(X_2 = 1 | X_1 = 0) P(X_1 = 0)$$
$$= \frac{K-1}{N-1} \cdot \frac{K}{N} + \frac{K}{N-1} \cdot \frac{N-K}{N} = \frac{K(N-1)}{N(N-1)} = \frac{K}{N}$$

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = \frac{K}{N}$$

Hypergeometric RV Properties



Suppose $X \sim HypGeo(N, K, n)$.

Let $X_1, \dots, X_n$ be indicator RV's (not independent) so that $X_i = 1$ if we got a lollipop on the i^{th} draw, and 0 otherwise. So $X = \sum_{i=1}^{n} X_i$.

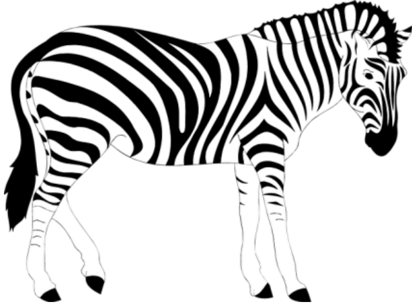
$$P(X_1 = 1) = \frac{K}{N}$$

$$P(X_2 = 1) = P(X_2 = 1 | X_1 = 1)P(X_1 = 1) + P(X_2 = 1 | X_1 = 0)P(X_1 = 0)$$
$$= \frac{K-1}{N-1} \cdot \frac{K}{N} + \frac{K}{N-1} \cdot \frac{N-K}{N} = \frac{K(N-1)}{N(N-1)} = \frac{K}{N}$$

$$E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = \frac{K}{N}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{K}{N} = n\frac{K}{N}$$

The Zoo of Discrete RV's



- The Bernoulli RV
- The Binomial RV
- The Geometric RV
- The Negative Binomial RV
- The Uniform RV
- The Poisson RV
- The Hypergeometric RV