Module 3A: Thinking in Distributions

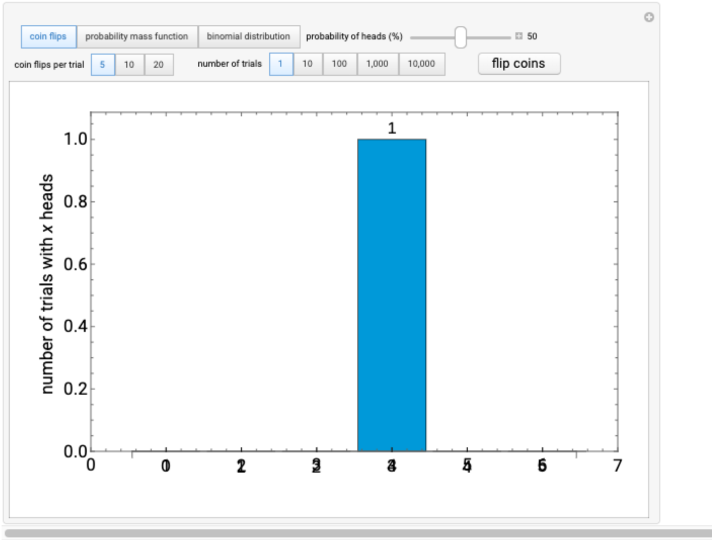
Building block for Hypothesis Testing

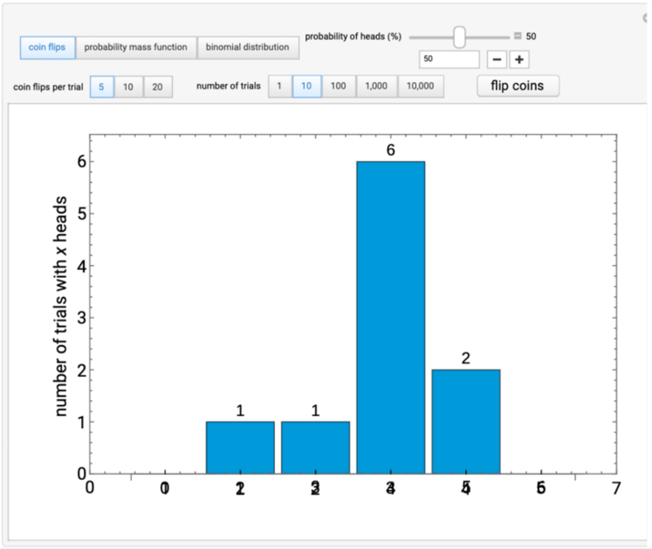
Agenda:

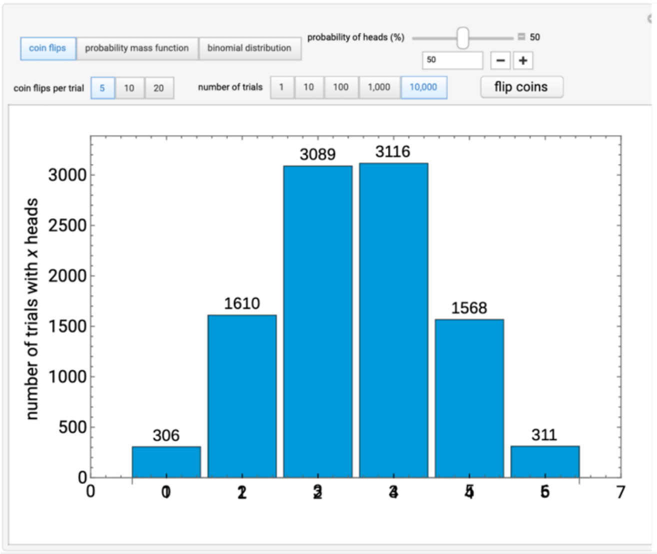
- Major distributions:
  - **Discrete Distributions**
    - Bernoulli
    - Binomial
    - Poisson
    - Hypergeometric
  - **Continuous Distributions**
    - Normal
    - Uniform
    - Exponential
    - Gamma
- Interactive simulations
- **Central Limit Theorem**
  - **Sampling Distribution of the mean**

Increase number of trials

Binomial Distribution via Coin Flips





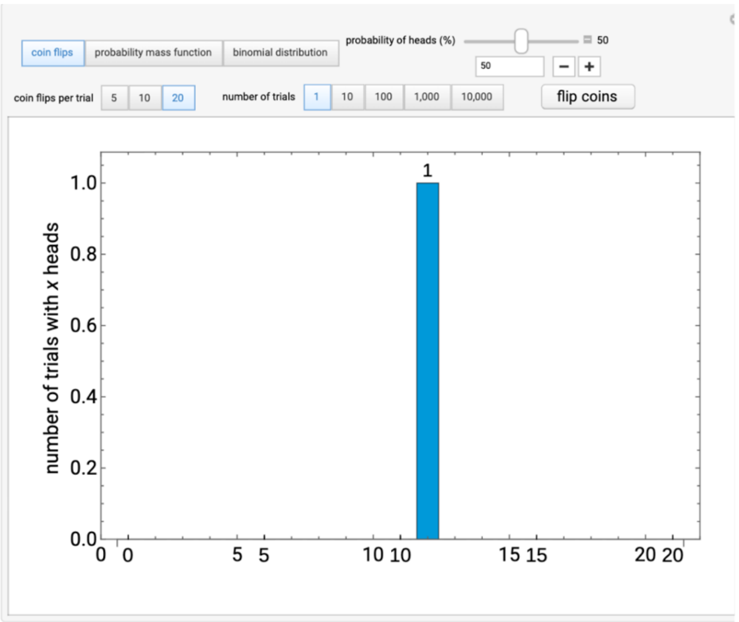


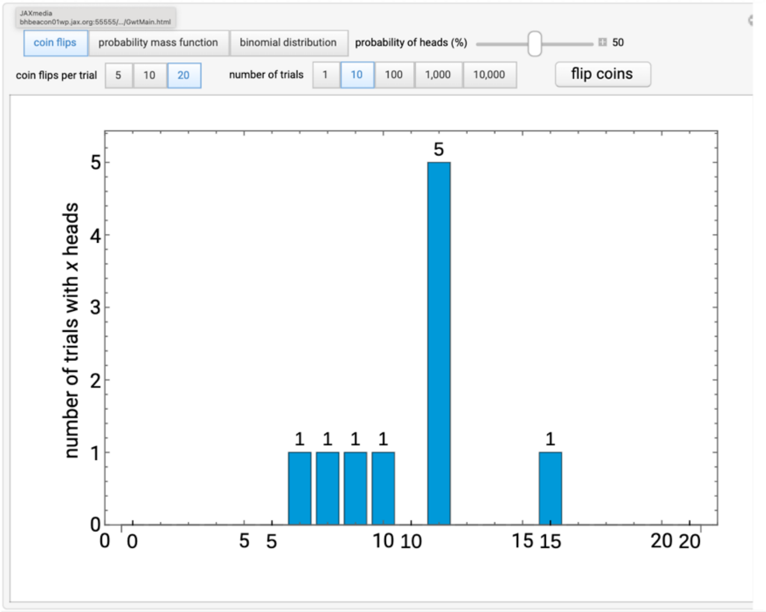
From Probability to Distributions: how randomness becomes predictable

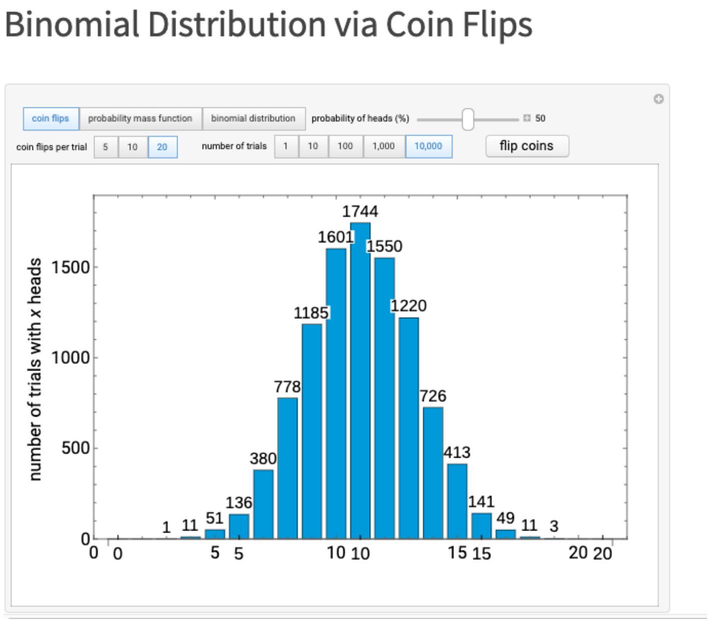
https://demonstrations.wolfram.com/BinomialDistributionViaCoinFlips/

Let’s try out simulations with p=0.5, p=0.75, p=0.90

Increase sample sizes







From Probability to Distributions: how randomness becomes predictable

https://demonstrations.wolfram.com/BinomialDistributionViaCoinFlips/

Let’s try out simulations with p=0.5, p=0.75, p=0.90

- **Probabilities are models that build up to create distributions**
  - Gives information about location and spread of the data
- There are standard distributions that explain common biological phenomenon

- **Discrete Distributions**
  - **Bernoulli**
  - **Binomial**
  - Poisson
  - Hypergeometric

- **Continuous Distributions**
  - **Normal**
  - Uniform
  - Exponential
  - Gamma

Websites for simulations:

1. https://seeing-theory.brown.edu/probability-distributions
2. https://probstats.org/

Bernoulli Distribution

Bernoulli Trials:

- ***Random** process with **only two mutually exclusive outcomes***
  - Coin toss: heads versus tails; Contest: Win or lose
  - General: one is called a success, one is called a failure

- *The probability, **p**, of success is the same in every trial*

- *The trials are **independent-** the outcome of any particular trial has no influence on the results of any other trial*

- *Visualization: **https://probstats.org/bernoulli.html***

Binomial Distribution

- **Binomial Random Variable**
  - Repeat a Bernoulli trial, with probability of success ***p***, to get a Binomial Random Variable
  - **X** is the number of successes in a fixed number, ***n***, of repeated Bernoulli trials
    - Example: P(X = k), where X represents the number of heads in two-coin flips so k = 0,1,2

k = # of successes	0	1	2
P(*X=k*)	0.25	0.50	0.25

P(X=0 heads) = **TT**(0.5*0.5)

P(X=1 head) = **HT**(0.5*0.5) + **TH**(0.5*0.5)

P(X=2 heads)= **HH**(0.5*0.5)

Visualization: https://probstats.org/binomial.html

- **<u>Binomial Distribution:</u>**
  - Describes the probability of a given number of '**<u>successes</u>**', which have a ***p*** probability, from a fixed number of independent trials, ***n***

$$\boxed{P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}}$$

  - **<u>The Binomial coefficient:</u>**

    It counts all the unique unordered sequences of getting *k* successes in *n* trials.
    ie. how many ways are there of getting k successes?

$$\binom{n}{k}=\frac{n!}{k!(n-k)!}$$

Where n! = n x (n-1) x (n-2) x ... x1
Also: **0! = 1 and 1! = 1**

• <u>The Binomial coefficient:</u>

$$\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

<u>Example:</u> How many ways are there of ordering 1 success (heads) and 1 failure (tails)? **TH, HT**

$\binom{2}{1} = \frac{2*1}{1!(2-1)!} = 2$ ways of ordering 1 success and 1 failure

<u>Example</u>: What is the probability of getting exactly the following pattern (2 successes and 3 failures): F F S F S

- The Binomial coefficient:

$$\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

Example: What is the probability of getting exactly the following pattern (2 successes and 3 failures): F F S F S

Answer: the hard way (drawing them all out...)

FFFSS FFSFS FFSSF FSFFS FSFSF FSSFF SFFFS SFFSF SFSFF SSFFF

The easy way: 5 choose 3 = 5 choose 2 = 5x4x3x2x1/ {(2x1)(3x2x1)} = 10

- <u>The Binomial coefficient:</u>

$$\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

<u>Example</u>: What is the probability of getting exactly the following pattern (2 successes and 3 failures): F F S F S

$\binom{5}{2} = \frac{5*4*3*2*1}{2*1(3*2*1)}$ = 10 ways of ordering 2 successes and 3 failures

$\binom{5}{3} = \frac{5*4*3*2*1}{3*2*1(2*1)}$ = 10 ways of ordering 3 successes and 2 failures

P(getting one particular pattern, F F S F S) = 1/10 ways

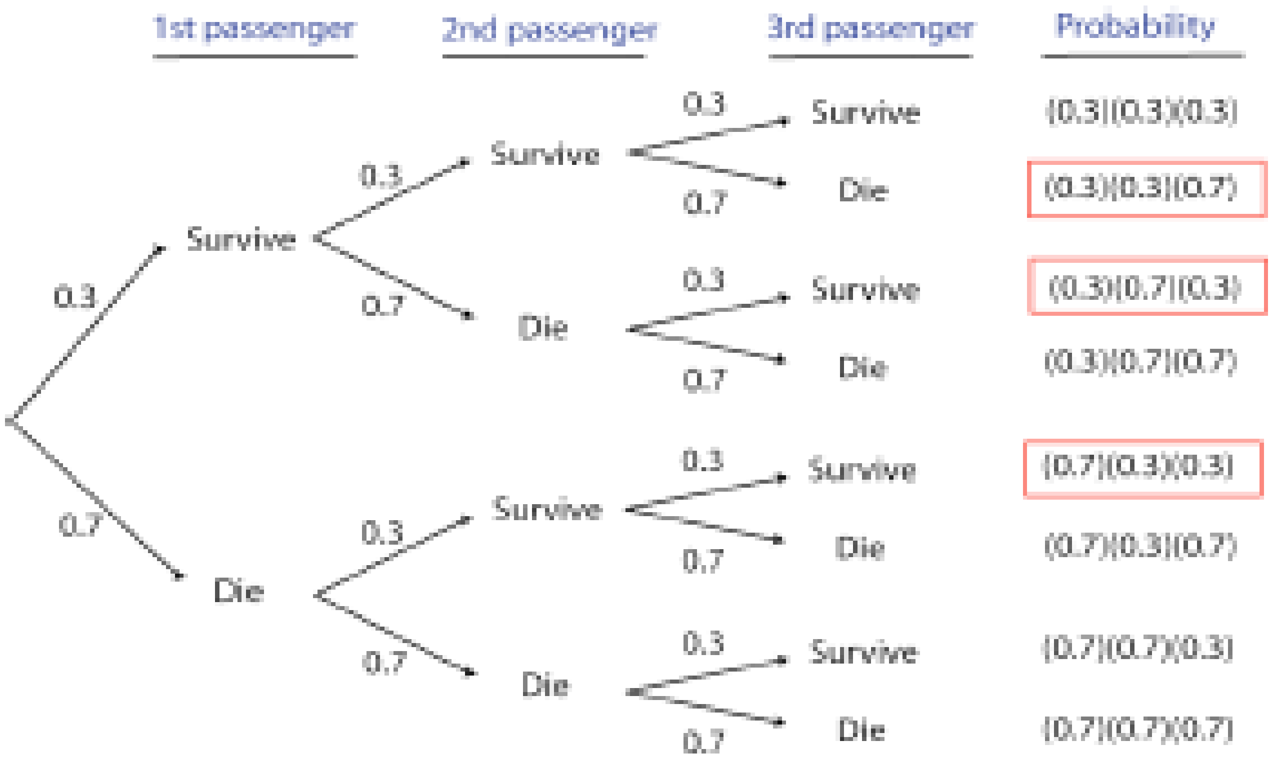
Binomial Distribution

Example: **2092** passengers on the titanic; **654** survived

P(surviving) = 654/2092 = **0.3**

Question: *What is the probability that 2 out of 3 randomly chosen passengers survived?*

Answer: The hard way…

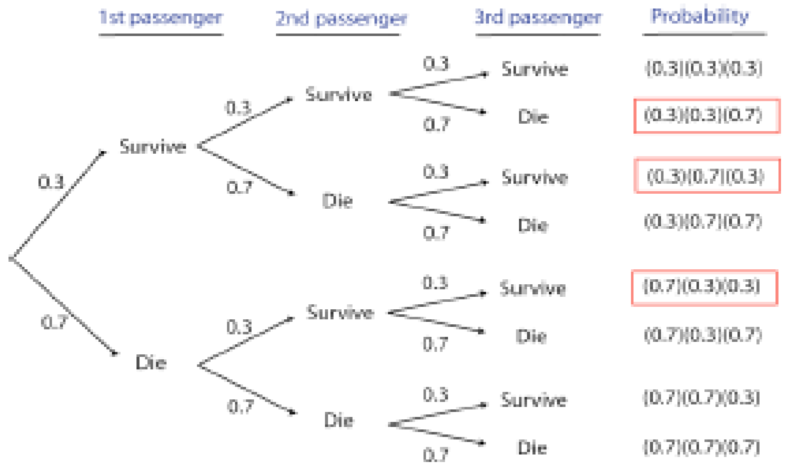


Binomial Distribution

<u>Example:</u> **2092** passengers on the titanic; **654** survived

P(surviving) = 654/2092 = **0.3**

<u>Question :</u> *What is the probability that 2 out of 3 randomly chosen passengers survived?*



<u>Answer:</u> The hard way…

<u>Answer:</u> The easy way… the Binomial

$$\binom{3}{2}(0.3)^2(0.7)$$

$$\boxed{P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}}$$

<u>Properties of the Binomial Distribution:</u>

A binomial random variable has values that are the number of successes

The long way of demonstrating the mean and standard deviation:

If 40% of brand A widgets have a particular defect, and I buy 5 of these widgets, what is the expected number of defective widgets that I now own?

Use https://probstats.org/binomial.html with n=5, p=0.40 to visualize the distribution…..

Properties of the Binomial Distribution:

A binomial random variable has values that are the number of successes

ANSWER (hard way):

P(k=0)=(5 choose 0)$0.4^0 0.6^5$

Outcome:	0 widgets	1 widget	2 widget	3 widget	4 widget	5 widget
Probability:	0.07776	0.25920	0.34560	0.23040	0.07680	0.01024
Random Variable:	0	1	2	3	4	5

$$\boxed{\bar{X} = 0(0.07776) + 1(0.25920) + 2(0.34560) + 3(0.23040) + 4(0.07680) + 5(0.01024) = 2}$$

This is the same answer as would be obtained by simply multiplying the probability of "success" times the number of cases….

$$\mu = np$$

$$\sigma^2 = np(1-p)$$