# Common Probability Distributions

Keegan Korthauer

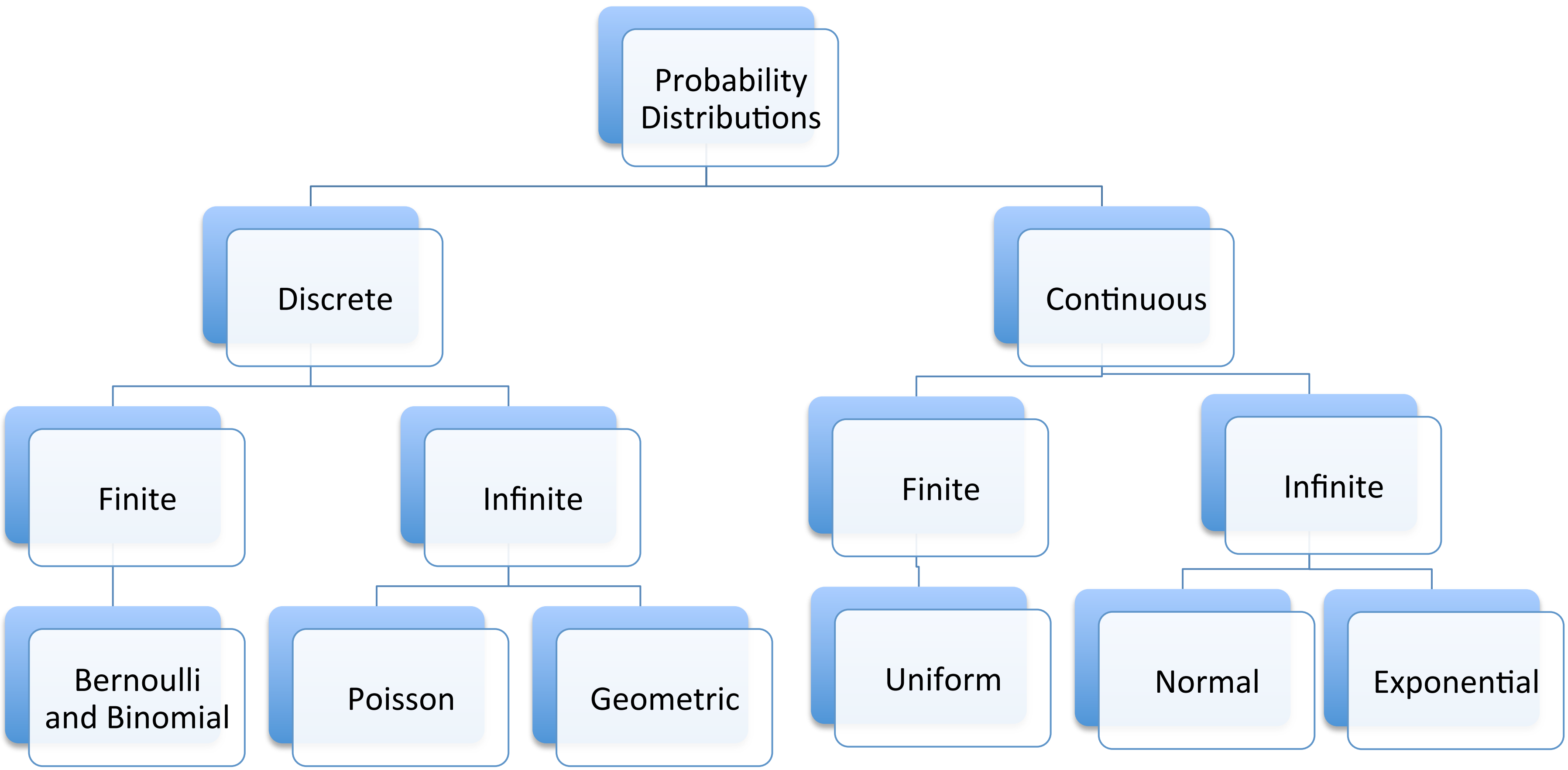
Department of Statistics

UW Madison

# Motivation

- Analyze samples drawn from the population
- Approximate knowledge of population distribution with a standard family of curves or functions
- We will learn about several common distributions for both discrete and continuous random variables

# Some Common Distributions



# COMMON DISCRETE DISTRIBUTIONS

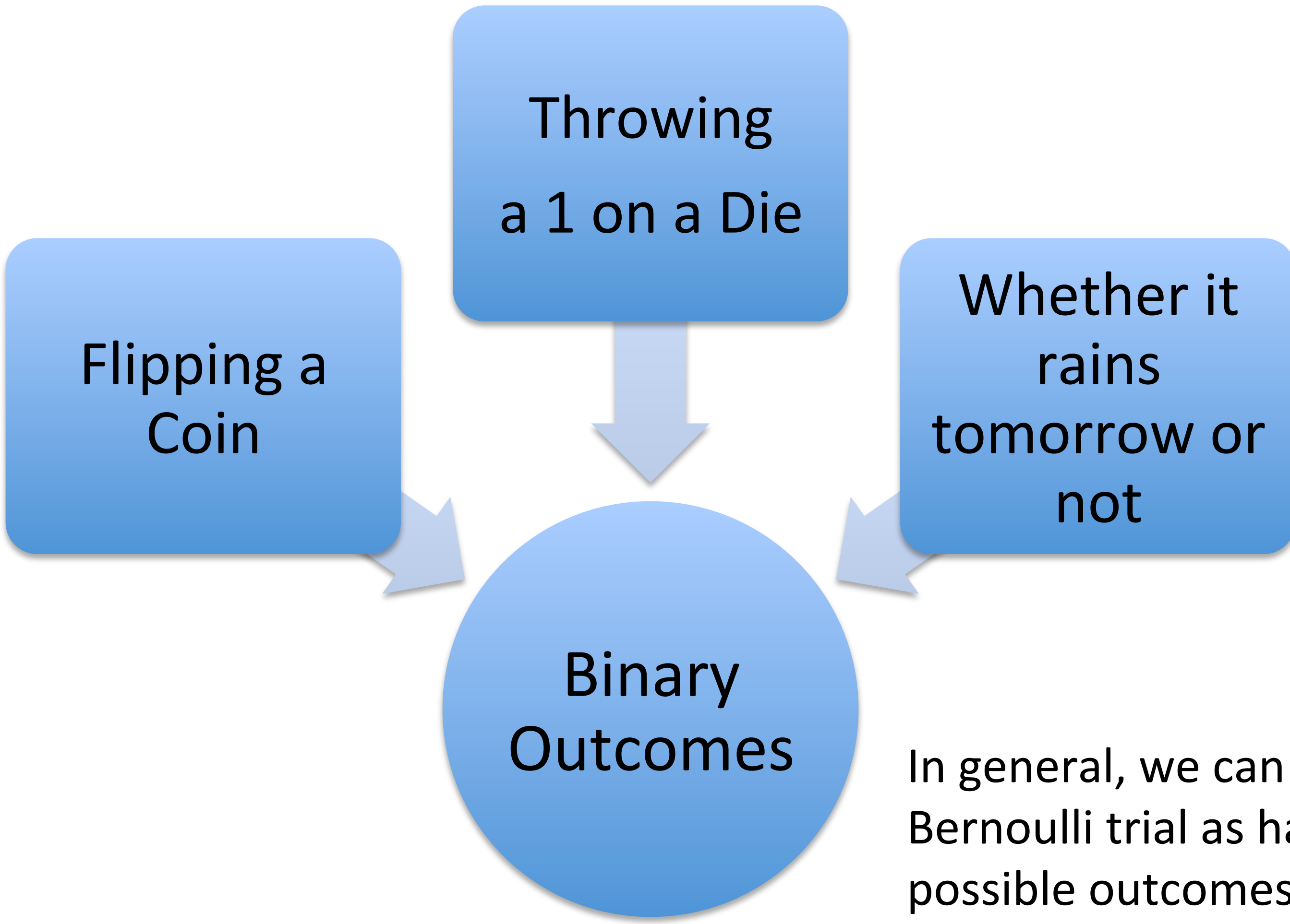
Bernoulli

Binomial

Poisson

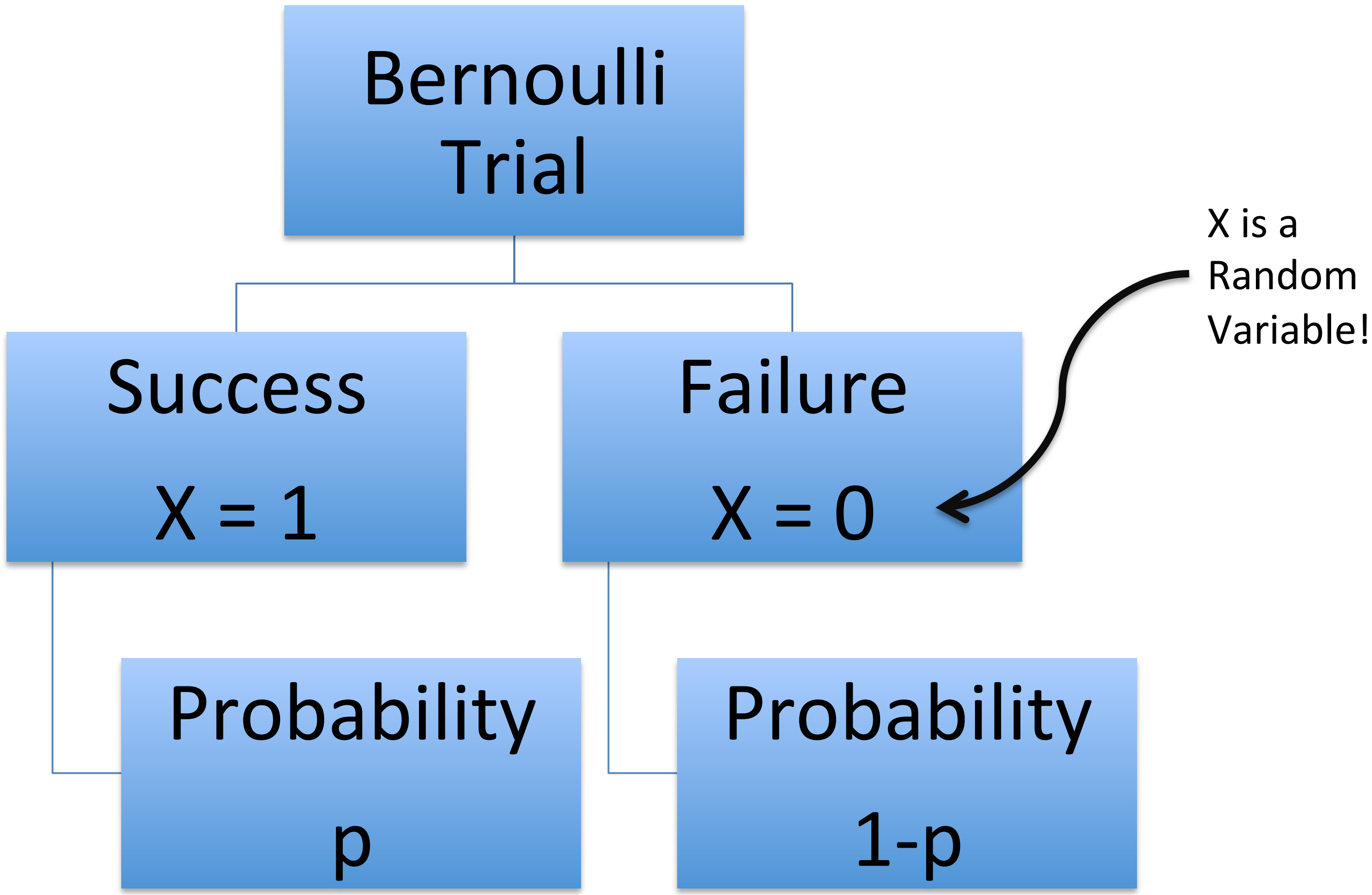
Geometric

# Bernoulli Trial




In general, we can describe a Bernoulli trial as having 2 possible outcomes – success or failure

# Bernoulli Distribution



# X ~ Bernoulli(p)

- ~ means "is distributed as"
- Probability mass function, where p is the success probability:

$$P(X=1) = p$$

$$P(X=0) = 1-p$$

- Mean and variance of X:

$$\mu_X = p$$

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

Where did that come from??

# Coin Toss Example

- We are going to toss a fair coin one time
- Let X=1 if the coin comes up heads and X=0 if the coin comes up tails

There are only two outcomes (success or failure) and the 'success' probability is $p = P(X=1) = 0.5$.

Thus, $X \sim \text{Bernoulli}(0.5)$.

# Light Bulb Example

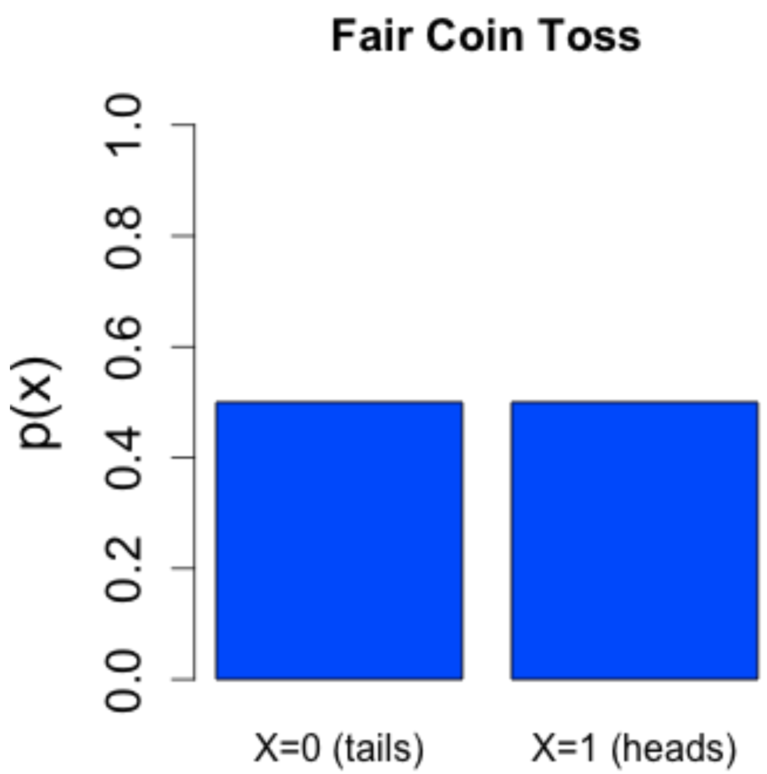
- Suppose 20% of light bulbs in a certain batch are defective
- Choose one bulb at random and let X=1 if the bulb is defective and X=0 if it is functioning
- What is the distribution of X?

There are only two outcomes (success or failure) and the ‘success’ probability is $p = P(X=1) = 0.2$.

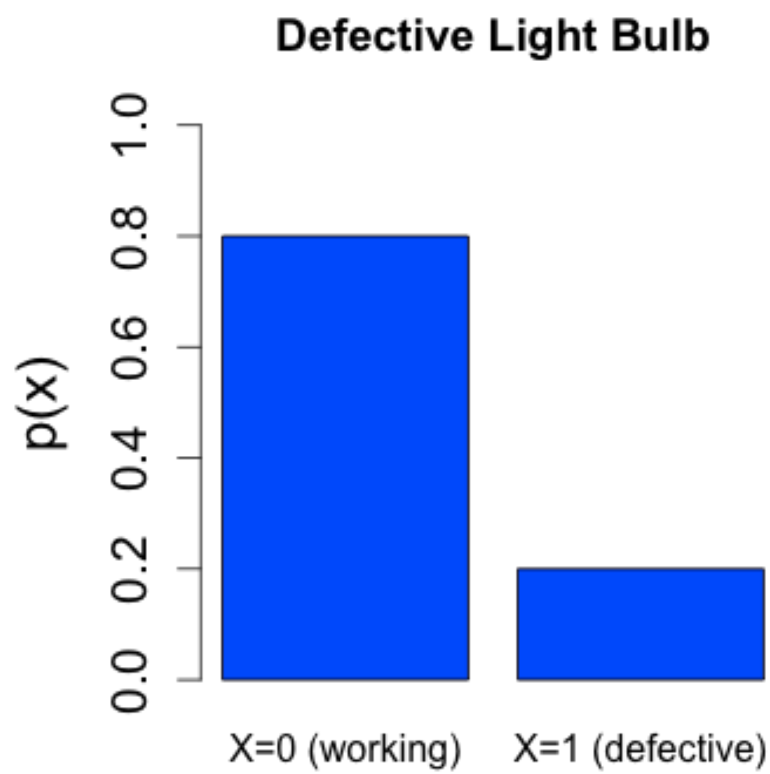
Thus, $X \sim \text{Bernoulli}(0.2)$.

*Note that ‘success’ means X=1, not that the light bulb is working*

# Probability Mass Functions

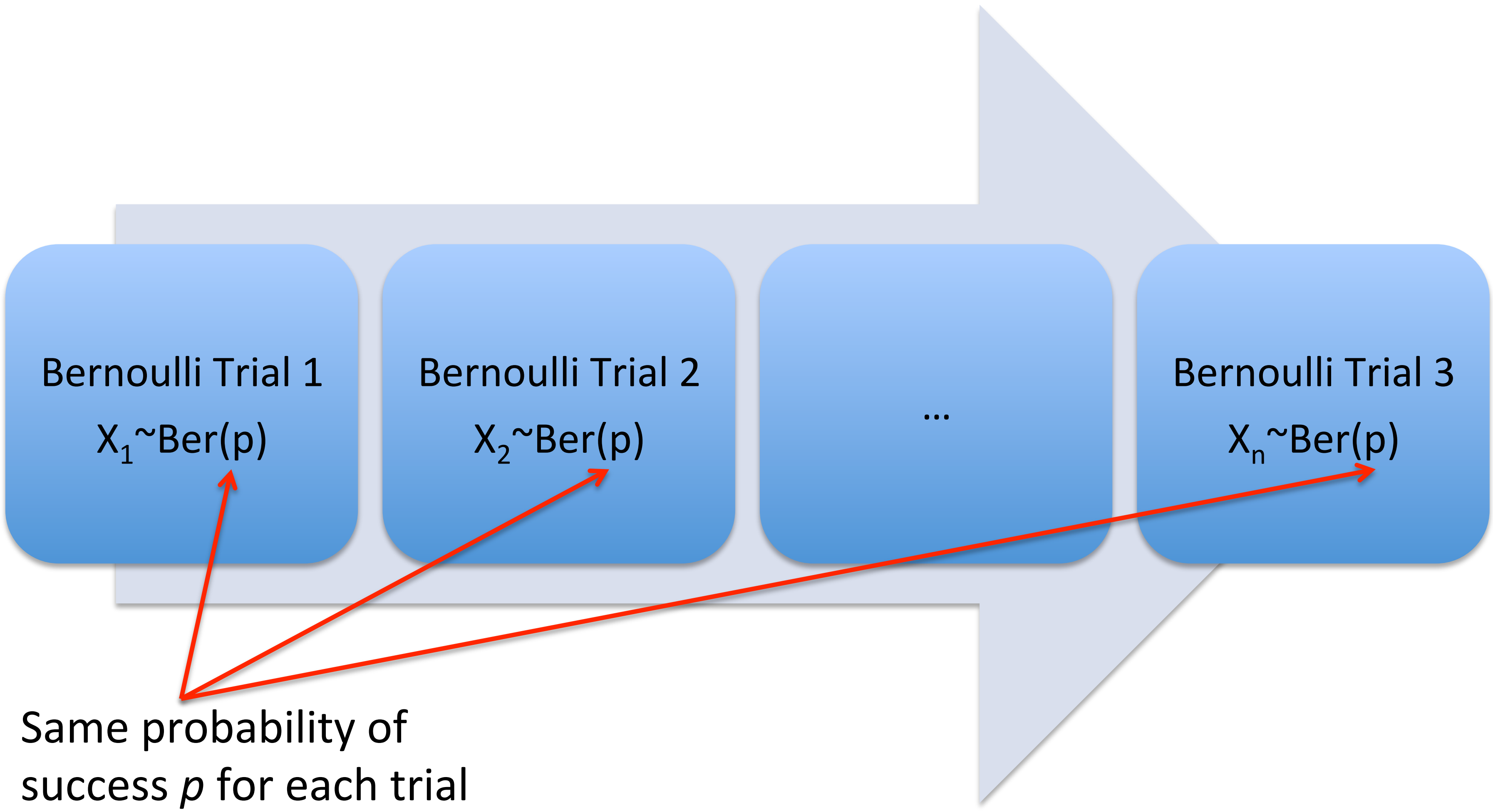



X ~ Bernoulli(0.5)




X ~ Bernoulli(0.2)

# Multiple Independent Bernoulli Trials



# Example – Multiple Coin Tosses

Toss a fair coin 10 times. Each toss is a Bernoulli trial and can have outcome Head (H) or Tail (T). The probability of success for each trial is the same (p=0.5)

We are interested in the total number of heads

Perform the experiment several times:


H T T H H H T H T H = 5H
T H H H T H H T T H = 6H
T H T T T H H H T T = 4H
H H H T T H T H H T = 6H
H H H T H H H T T H = 7H
…

# Example – Multiple Coin Tosses

- Let each $X_1, X_2, \ldots, X_{10}$ each be independent Bernoulli trials with $p=0.5$ ($X_i=1$ if H, $X_i=0$ if T)

- Let $Y=X_1+X_2+\ldots+X_{10}$ (total number of heads)

- How do we find the probability that $Y=4$?

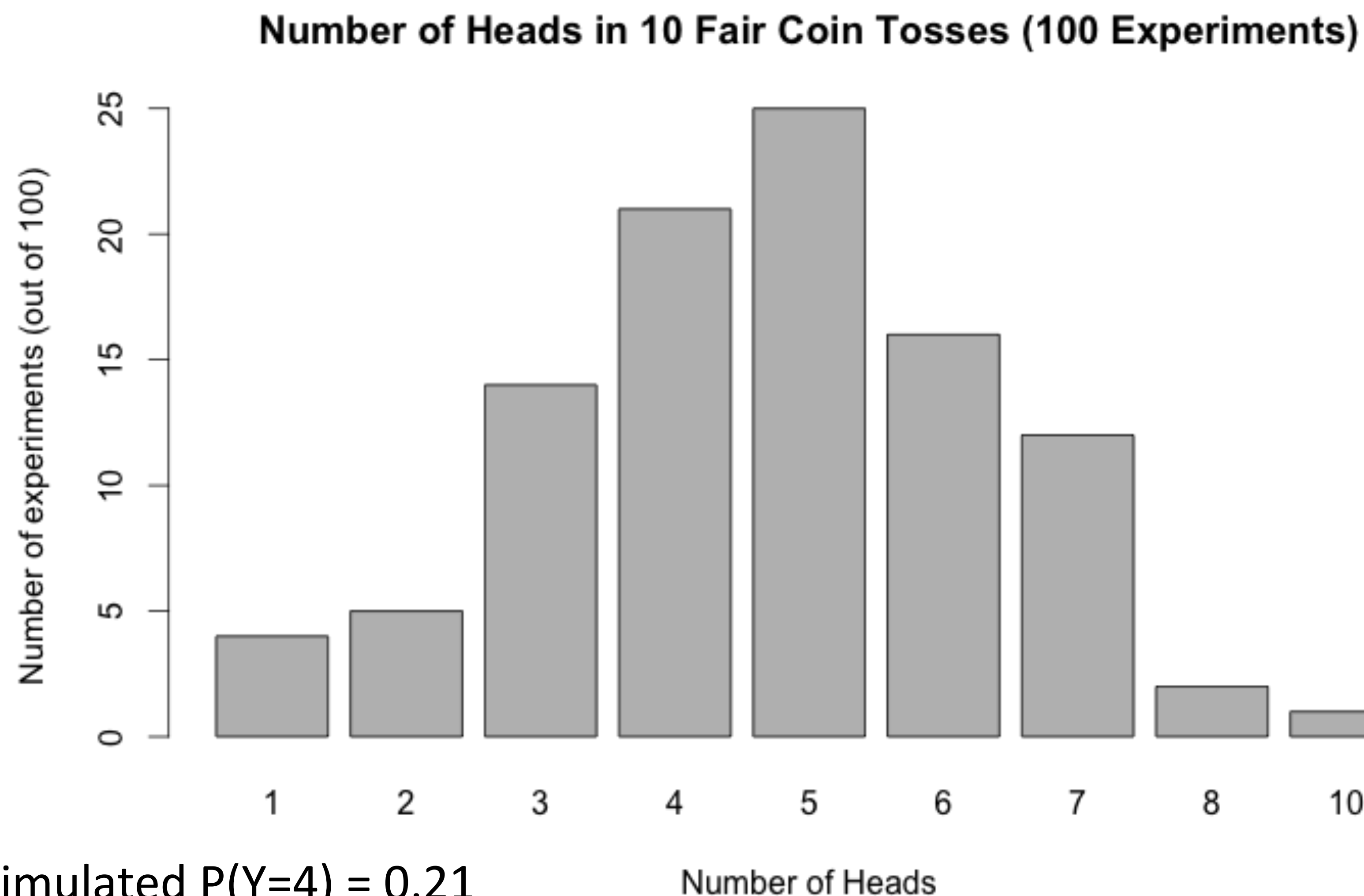
- What is the expected value (mean) of Y?

# Simulation

- Use a random number generator to draw from {0, 1} with 0.5 probability of each; repeat 10 times
- Sum the result
- Repeat this process many times
- I did this in R for 100 runs:

```
y <- NULL
for (i in 1:100){
    y <- c(y, sum(sample(c(0,1), size=10, prob=c(0.5,0.5),
            replace=TRUE)))
}
```

# Results of the Coin Toss Simulation




Simulated P(Y=4) = 0.21
Simulated mean = 4.71

# Binomial Distribution

- We can find exact probabilities for Y using the Binomial distribution
- The number of successes Y in a series of *n* **independent** Bernoulli trials (each with the same success probability *p*) is a Binomial random variable
- We say that Y ~ Bin(*n*, *p*)

# X ~ Binomial(n,p)

- Probability mass function

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

for $x$ = 0, 1, 2, ..., n

- Mean and Variance

$$\mu_X = np$$

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

Where did that come from??

# Example – Multiple Coin Tosses

- Let each $X_1, X_2, \ldots, X_{10}$ each be independent Bernoulli trials with p=0.5 ($X_i=1$ if H, $X_i=0$ if T)

- Let $Y=X_1+X_2+\ldots+X_{10}$ (total number of heads)

- How do we find the probability that Y=4?

  P(Y=4) = 0.2051

- What is the expected value (mean) of Y?

  E(Y) = 5

# Independence in Binomial Experiments

- In the coin toss example, we clearly have independent Bernoulli trials
- Recall the light bulb example where we draw one component with p= 0.2. Will the probability of success change as the experiment goes on?
  - Depends on the size of the population
  - Rule of thumb, treat as independent if sample size is less than or equal to 5% of the population size

# Example

- We have a lot of 1,000 components that has a failure rate of 17%.
- Can we model the number of failed components in a random sample of 10 with a Binomial distribution?

- How about a sample of 100?

# Binomial Distribution - Illustration

- http://www.math.uah.edu/stat/applets/SpecialCalculator.html (choose 'Binomial' from the dropdown menu)
- Shape of the distribution (PMF) is determined by n and p
  - Symmetric if p = 0.5
  - Skewed left if p is large
  - Skewed right if p is small
  - Lower spread if n is large

# Recall Exercise 2.4.12

Suppose we have a large collection of components to test (success = S and failure = F), each with failure probability of 0.2. Let X represent the number of successes among a sample of three components.

X here is a Binomial random variable with n = 3 and p = 0.8 , so
X ~ Binomial(3, 0.8)

What is the mean of X?

E(X) = np = 3(0.8) = 2.4

What is the variance of X?

Var(X) = np(1-p) = 3(0.8)(0.2) = 0.48

What is P(X=3)?

$$P(X=3) = \binom{3}{3} 0.8^3 (1-0.8)^{(3-3)} = 0.8^3 = 0.512$$

# Example – 2 System Designs

One design for a system requires the installation of two identical components. The system will work if at least one of the components works.

An alternative design requires four of these components, and the system will work if at least two of the four components work.

If the probability that a component works is 0.9, and if the components function independently, which design has the greater probability of functioning?

Design 1: 0.81
Design 2: 0.9963

# When *p* is Unknown

- So far, we've dealt with examples where we have knowledge of the value of the success probability *p*
- In practice we often do not know *p* even though we know the process will be a series of independent Bernoulli trials
- How can we estimate it?
  - Sample *n* individuals from the population and count the number of successes X
  - Our best guess for *p*:

$$\hat{p} = \text{sample proportion} = \frac{\text{number of successes}}{\text{number of trials}} = \frac{X}{n}$$

# Example - Polling

- We want to estimate the percent of the population that support a certain candidate
- We take a random sample of 200 voters and find that 112 support the candidate.
- Estimate the probability *p* that a randomly drawn voter from the population will support the candidate

$$\hat{p} = \text{sample proportion} = \frac{112}{200} = 0.56$$

# How 'good' is our estimate?

- In other words, how much bias and uncertainty is there in our estimate?

- Bias = $\mu_{\hat{p}} - p = E(X/n) - p = \frac{np}{n} - p = 0$

- Uncertainty =

In practice, substitute $\hat{p}$ for p

$$\sigma_{\hat{p}} = \sigma_{X/n} = \frac{\sigma_X}{n} = \frac{\sqrt{np(1-p)}}{n} = \sqrt{\frac{p(1-p)}{n}}$$

Bigger sample -> smaller uncertainty!

# Polling Example Revisited

- 112 out of 200 in the sample supported the candidate
- Find the uncertainty in our estimate of p

Recall that $\hat{p} = 0.56$

$$\sigma_{\hat{p}} = \sqrt{\frac{0.56(1-0.56)}{200}} = 0.0351$$

# Next

- Two more discrete distributions:
  - Poisson
  - Geometric

- Continuous distributions