Probability and Statistic (for CS Students)

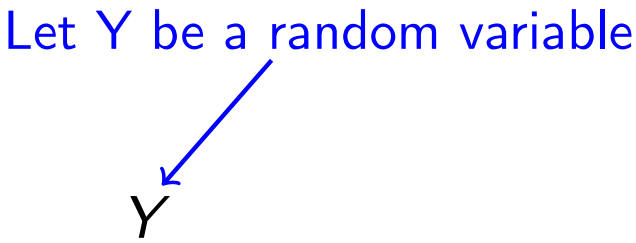
Lecture 9: Bernoulli and Binomial Random Variables

University of Mazandaran

Fall 1405

Review

- A random variable is a variable which takes on values probabilistically
- A discrete random variable is fully described by a probability mass function
- We can also calculate summary statistics such as expectation (and today, variance!)



For example Y is the number of heads in 5 coin flips

$$Y = 2$$

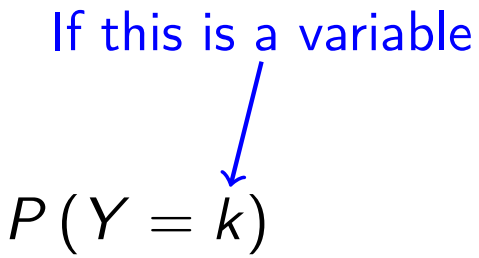
It is an event when Y takes on a value

For example Y is the number of heads in 5 coin flips

$$Y < 2$$

It is also an event when you ask any comparison question

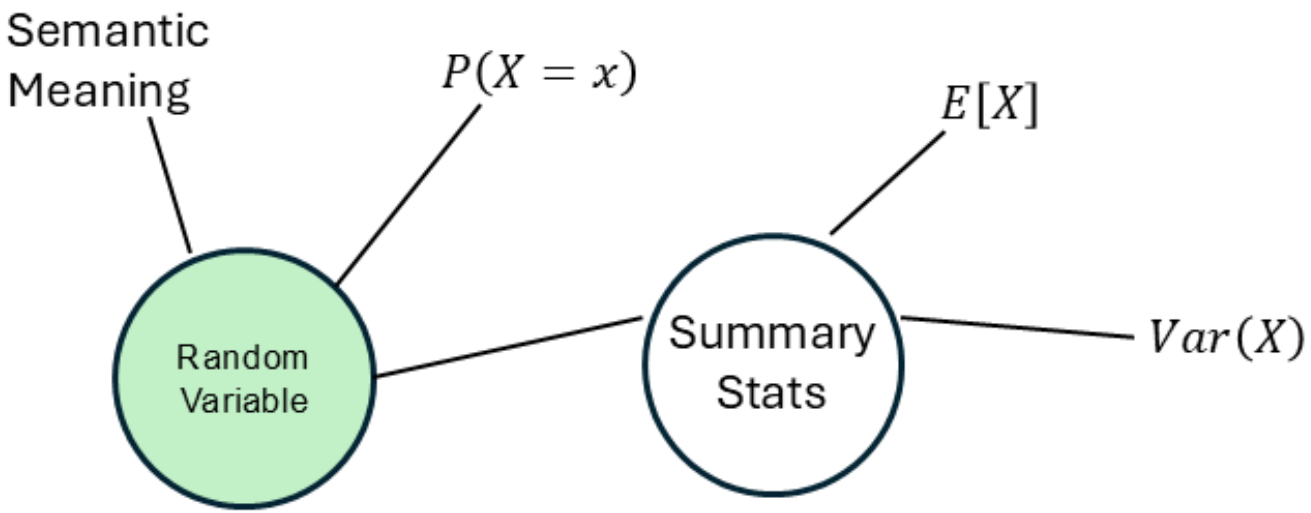
For example Y is the number of heads in 5 coin flips



Then this is a function

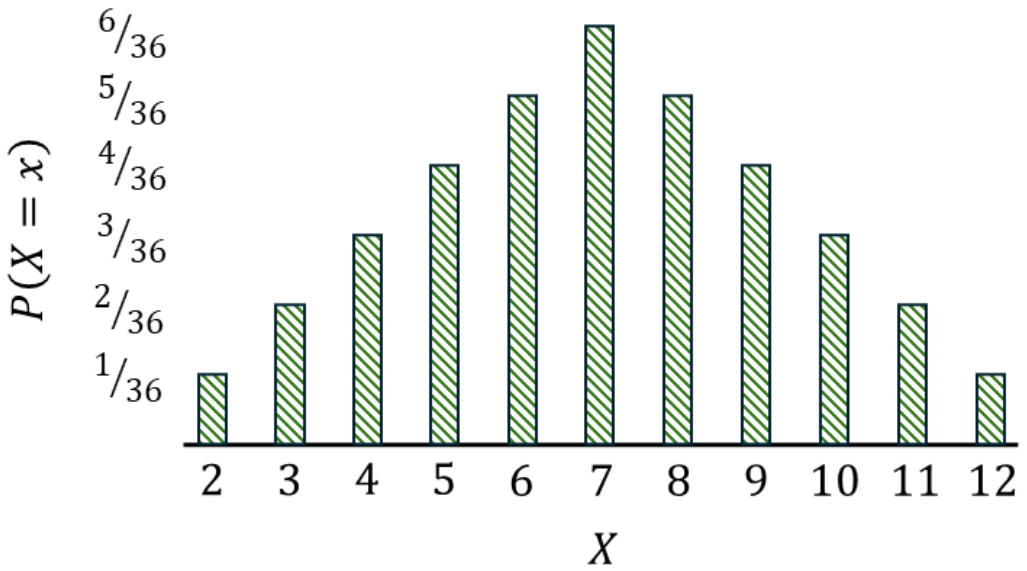
For example Y is the number of heads in 5 coin flips

Review

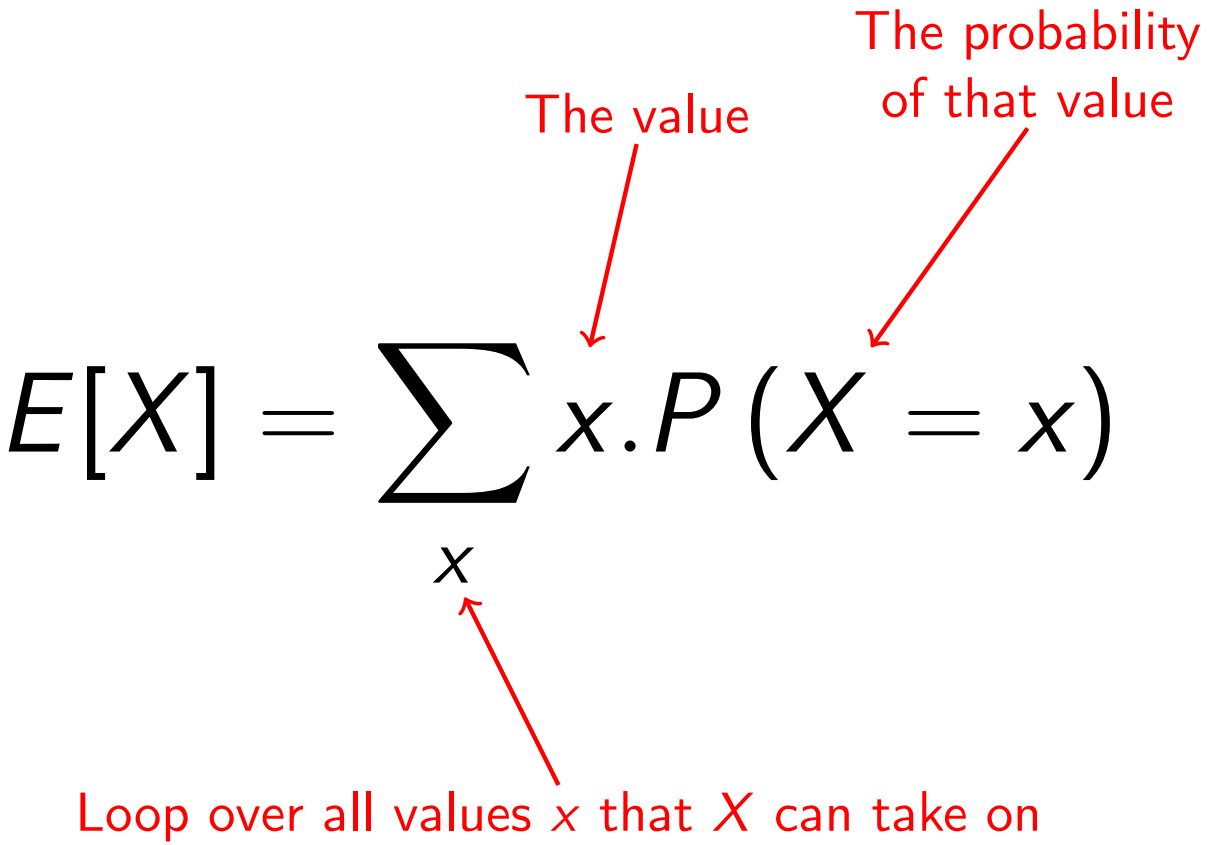

Semantic
Meaning
$P(X = x)$
$E[X]$
Random
Variable
Summary
Stats
$Var(X)$

PMF of sum of two dice

$$P(X = x) = \begin{cases} \frac{x-1}{36} & \text{if } x \in \mathbb{Z}, 1 \leq x \leq 6 \\ \frac{13-x}{36} & \text{if } x \in \mathbb{Z}, 7 \leq x \leq 12 \\ 0 & \text{else} \end{cases}$$



Expected Values



Is Expectation Enough?

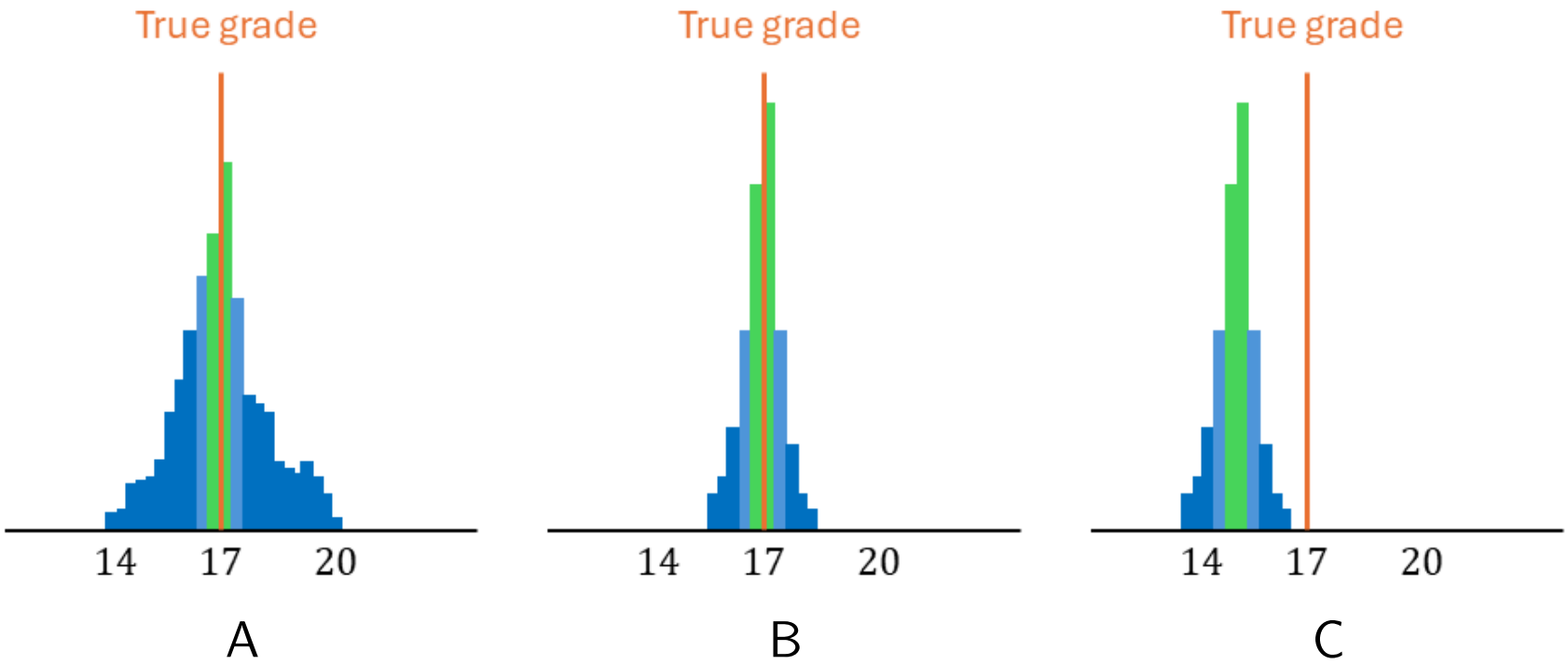
Is Expectation Enough?

No! PMF is complete!

Peer Grading

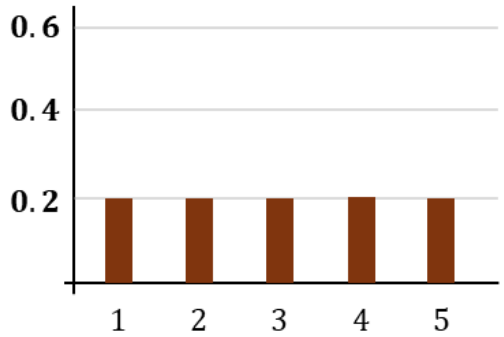
Peer grading on Coursera HCI
31067 peer grades for 3607 students.

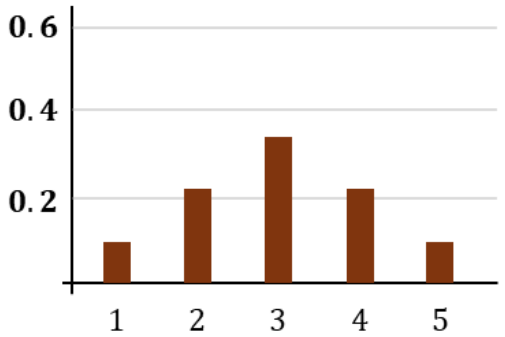
Peer Grading

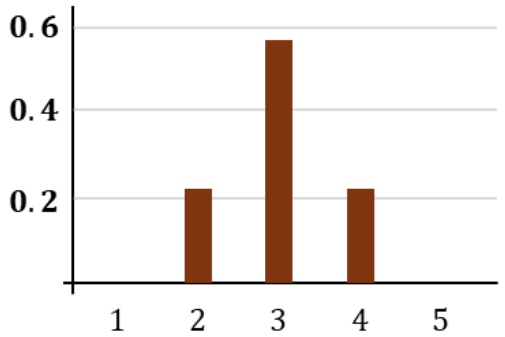


Measure of Spread

Consider the following 3 distributions (PMFs)

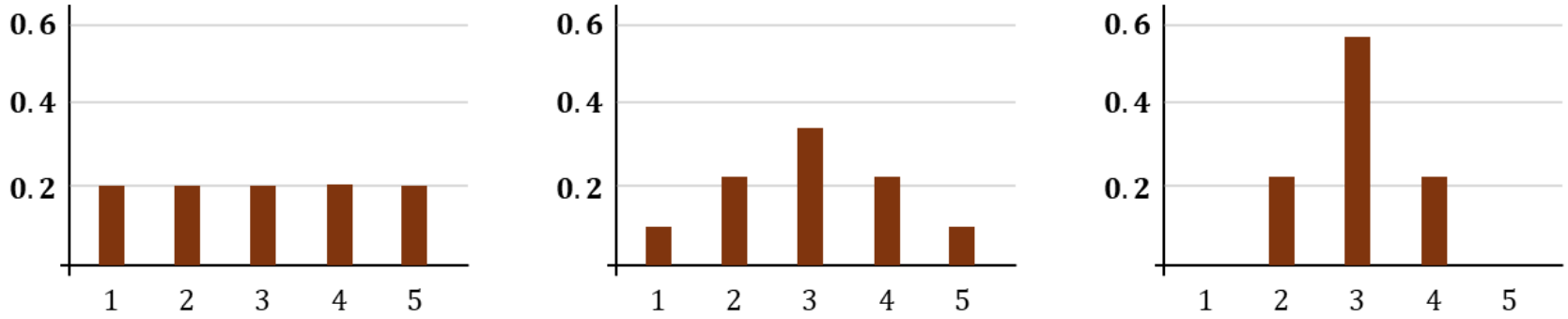






Measure of Spread

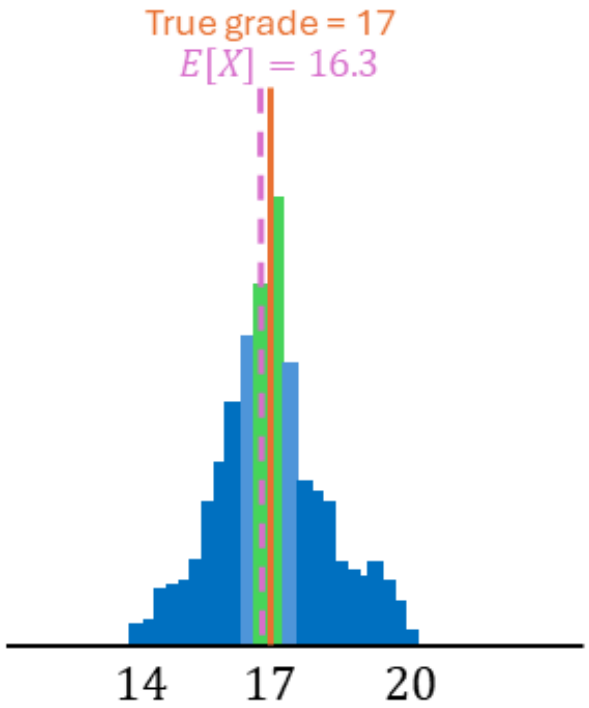
Consider the following 3 distributions (PMFs)



All have the same expected value, $E[X] = 3$
But "spread" in distributions is different
Invent a formal quantification of "spread"

Peer Grading in Coursera HCI

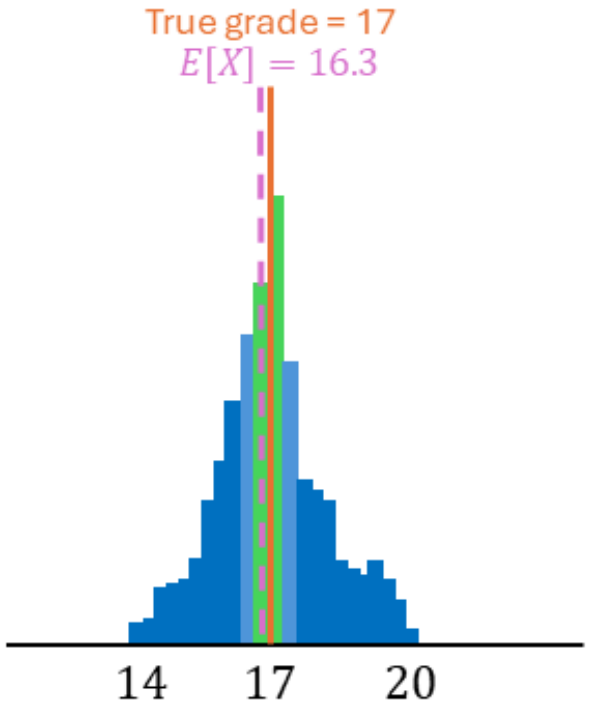
Let X be a random variable that represents a peer grade



Peer Grading in Coursera HCI

Let X be a random variable that represents a peer grade

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



Variance

if X is a random variable with mean μ, then the variance of X, denoted by $\mathrm{Var}(X)$, is:

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

Variance is a formal definition of the spread of a random variable

Variance

if X is a random variable with mean μ, then the variance of X, denoted by $\mathrm{Var}(X)$, is:

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

Variance is a formal definition of the spread of a random variable

Also known as the 2nd moment, or square of the standard deviation.

Computing Variance

$$
\begin{aligned}
\mathrm{Var}(X) &= E[(X-\mu)^2] \\
&= \sum_x (x-\mu)^2 p(x) \\
&= \sum_x \left(x^2 - 2\mu x + \mu^2\right) p(x) \\
&= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\
&= E[X^2] - 2\mu E[x] + \mu^2 \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - \mu^2 \\
&= E[X^2] - (E[X])^2
\end{aligned}
$$

Example

UMZ wants to equip its classrooms with interactive whiteboards. Two contractors are among the options and their monthly supply costs over previous 6 months are:

- **Contractor A**: \$960, \$1030, \$980, \$1000, \$1040, \$990
- **Contractor B**: \$500, \$1300, \$900, \$600, \$1500, \$1200

Compute the mean and variance of the monthly cost for each contractor and suggest one to the dean?

Example

UMZ wants to equip its classrooms with interactive whiteboards. Two contractors are among the options and their monthly supply costs over previous 6 months are:

- **Contractor A**: \$960, \$1030, \$980, \$1000, \$1040, \$990
- **Contractor B**: \$500, \$1300, \$900, \$600, \$1500, \$1200

Compute the mean and variance of the monthly cost for each contractor and suggest one to the dean?

$$E[A] = \sum_{i=1}^{6} x_i p(x_i) = (960 + 1030 + 980 + 1000 + 1040 + 990) \times \frac{1}{6} = 1000$$

$$E[B] = \sum_{i=1}^{6} x_i p(x_i) = (500 + 1300 + 900 + 600 + 1500 + 1200) \times \frac{1}{6} = 1000$$

$$Var[A] = E\left[(A - E[A])^2\right] = 766.66$$

$$Var[B] = E\left[(B - E[B])^2\right] = 133333.3$$

Bernoulli Random Variable

Experiment results in “Success" or “Failure"

- X is random indicator variable (1 =success, 0 =failure)

Bernoulli Random Variable

Experiment results in “Success" or “Failure"

- X is random indicator variable (1 =success, 0 =failure)
- $P(X = 1) = p(1) = p$ and $P(X = 0) = p(0) = 1 - p$

Bernoulli Random Variable

Experiment results in “Success" or “Failure"

- X is random indicator variable (1 =success, 0 =failure)
- $P(X = 1) = p(1) = p$ and $P(X = 0) = p(0) = 1 - p$
- X is Bernoulli Random Variable: $X \sim \text{Bern}(p)$

Bernoulli Random Variable

Experiment results in “Success" or “Failure"

- X is random indicator variable (1 =success, 0 =failure)
- $P(X = 1) = p(1) = p$ and $P(X = 0) = p(0) = 1 - p$
- X is Bernoulli Random Variable: $X \sim \text{Bern}(p)$

Is distributed as a

Bernoulli Random Variable

Experiment results in “Success" or “Failure"

- X is random indicator variable (1 =success, 0 =failure)
- $P(X = 1) = p(1) = p$ and $P(X = 0) = p(0) = 1 - p$
- X is Bernoulli Random Variable: $X \sim \text{Bern}(p)$

Is distributed as a

parameter: success probability

Bernoulli Random Variable

Experiment results in “Success" or “Failure"

- X is random indicator variable (1 =success, 0 =failure)
- $P(X = 1) = p(1) = p$ and $P(X = 0) = p(0) = 1 - p$
- X is Bernoulli Random Variable: $X \sim \text{Bern}(p)$
- $E[X] = p$

Bernoulli Random Variable

Experiment results in “Success" or “Failure"

- X is random indicator variable (1 =success, 0 =failure)
- $P(X = 1) = p(1) = p$ and $P(X = 0) = p(0) = 1 - p$
- X is Bernoulli Random Variable: $X \sim \text{Bern}(p)$
- $E[X] = p$
- $Var(X) = p(1 - p)$

Again Coins!

1. n independent trials of a Bernoulli experiment (eg flipping a coin)
2. Each trial has a probability of p, of being a success (eg a head)
3. What is the probability of exactly k successes? (eg k heads)

Many Random Variables Follow This Pattern

Examples:

- # of heads in n coin flips
- # of 1's in randomly generated bit string of length n
- # of disk drives crashed in 1000 computer cluster
- # of people who vote for a candidate

Exactly k heads in n coin flips

X: R.V. that models the number of heads in n coin flips
Probability of exactly k heads:

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

Exactly k heads in n coin flips

X: R.V. that models the number of heads in n coin flips
Probability of exactly k heads:

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

X is called a binomial random variable

Declare a Random Variable to be Binomial

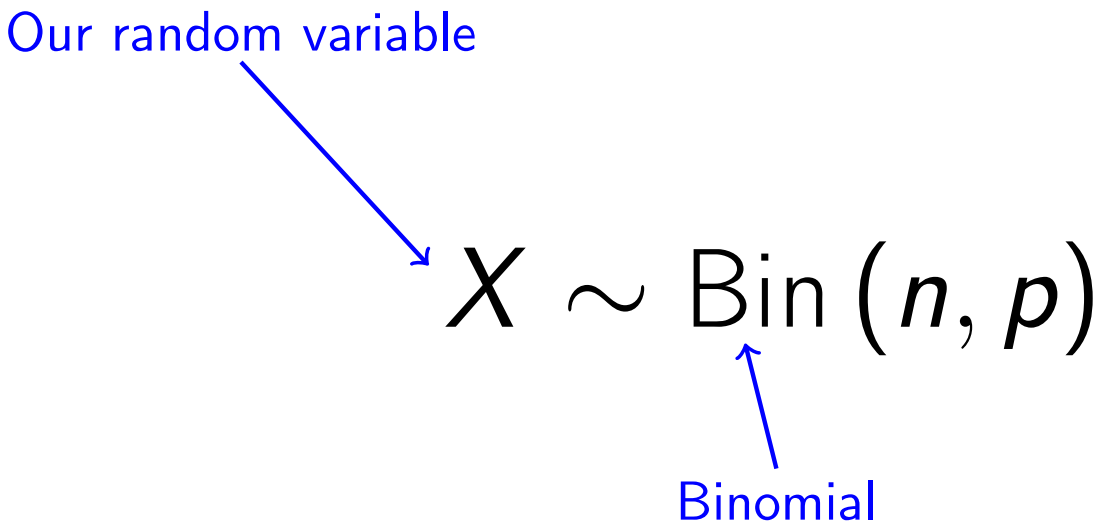
$$X \sim \text{Bin}(n, p)$$

Declare a Random Variable to be Binomial

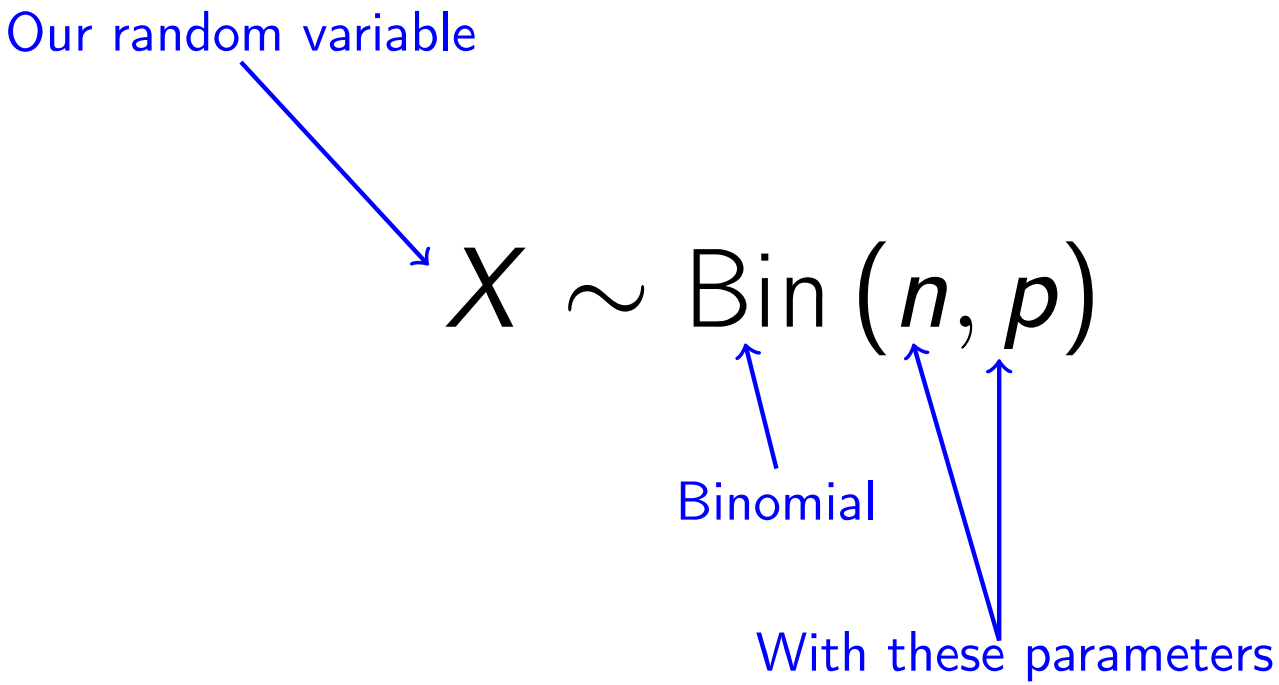
Our random variable

$$X \sim \text{Bin}(n, p)$$

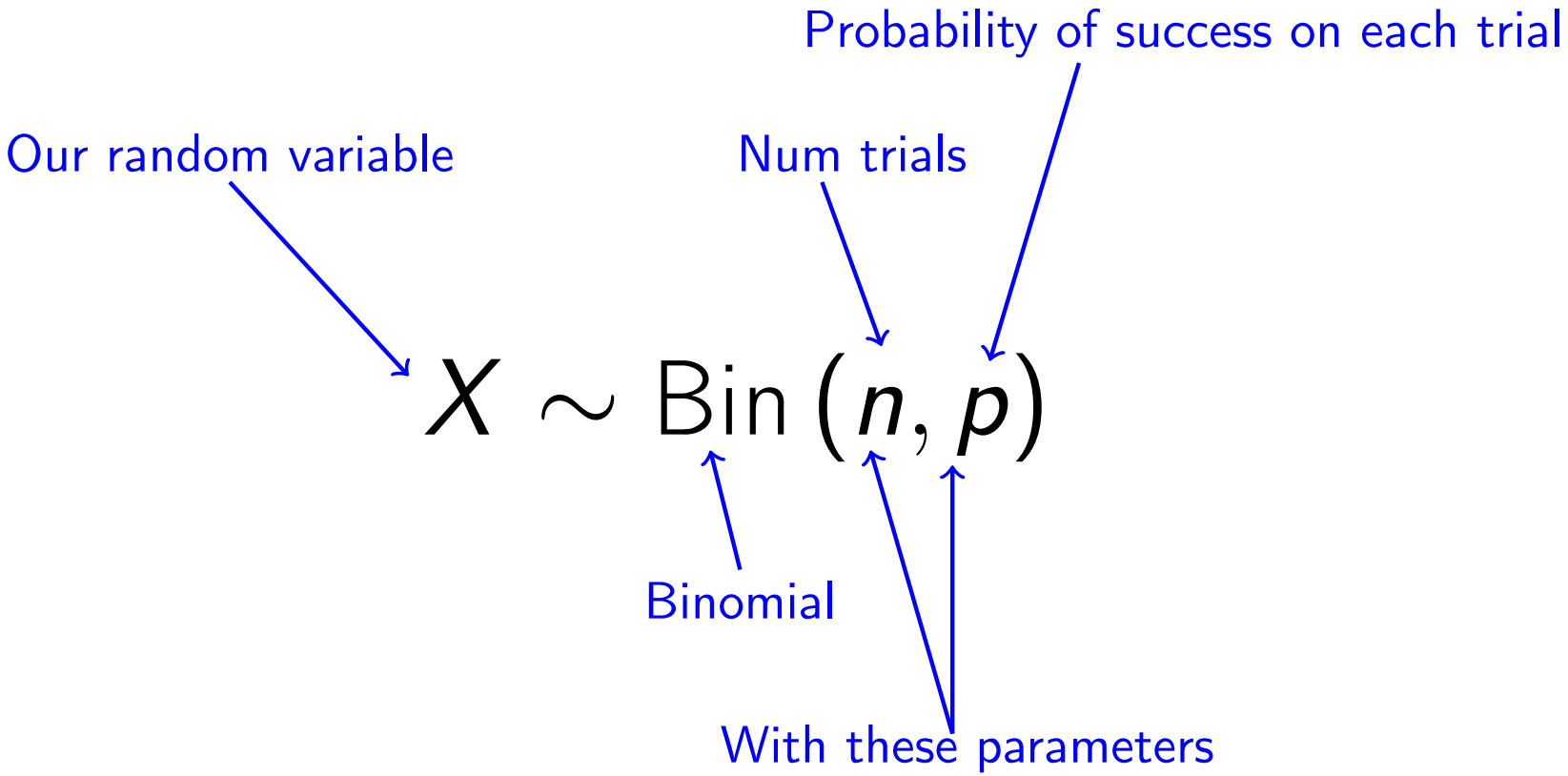
Declare a Random Variable to be Binomial



Declare a Random Variable to be Binomial



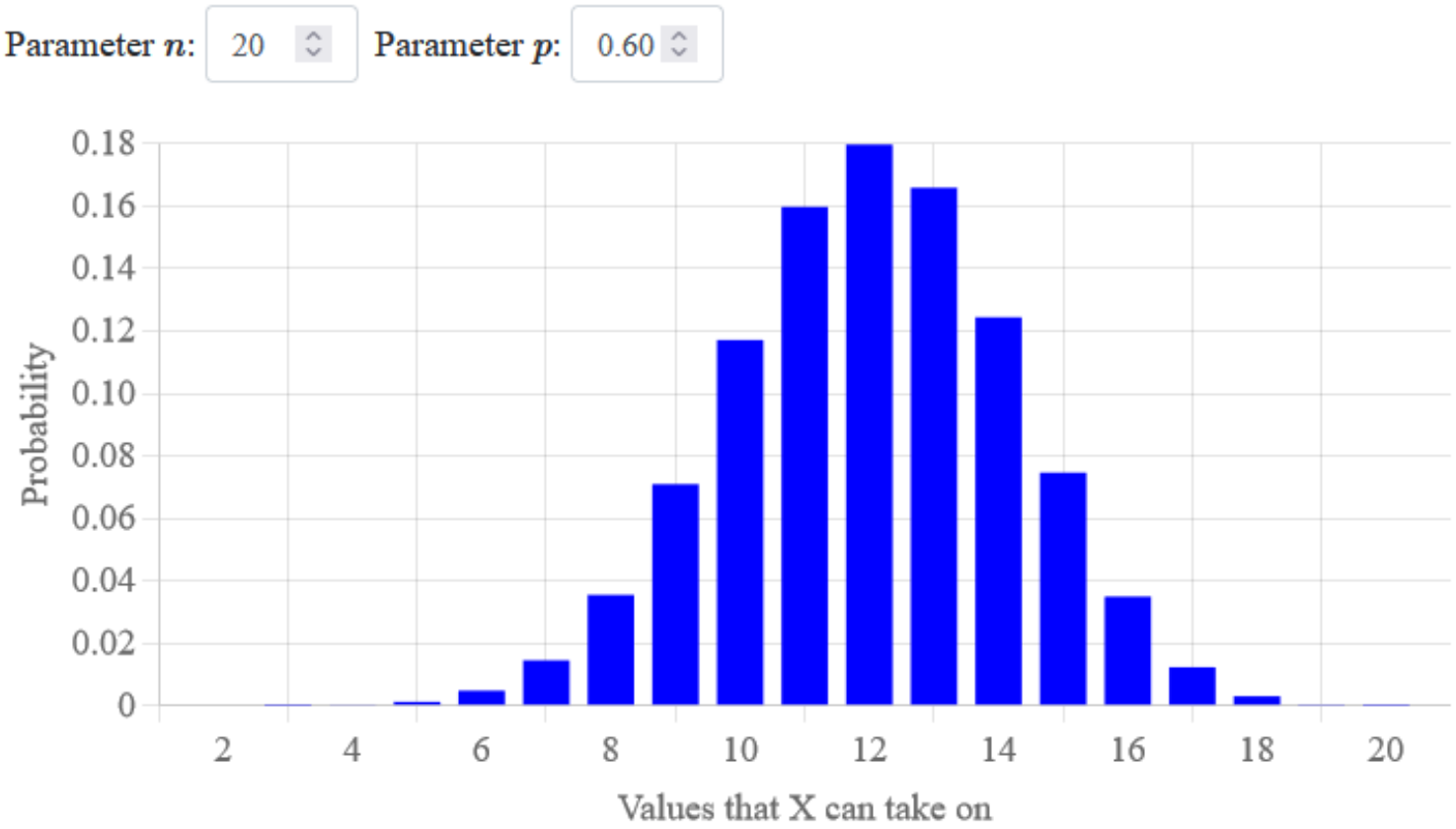
Declare a Random Variable to be Binomial



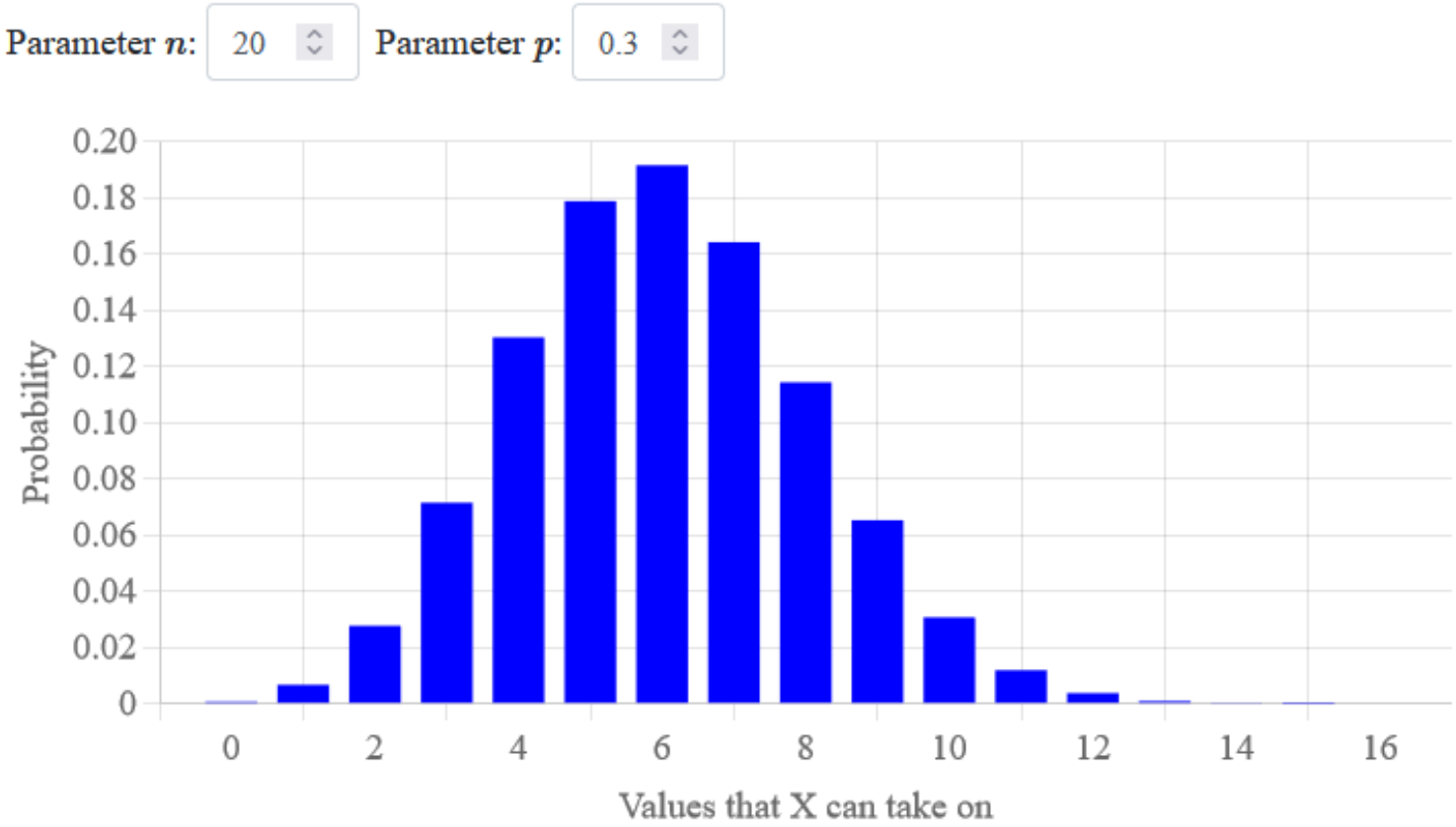
Binomial PMF

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

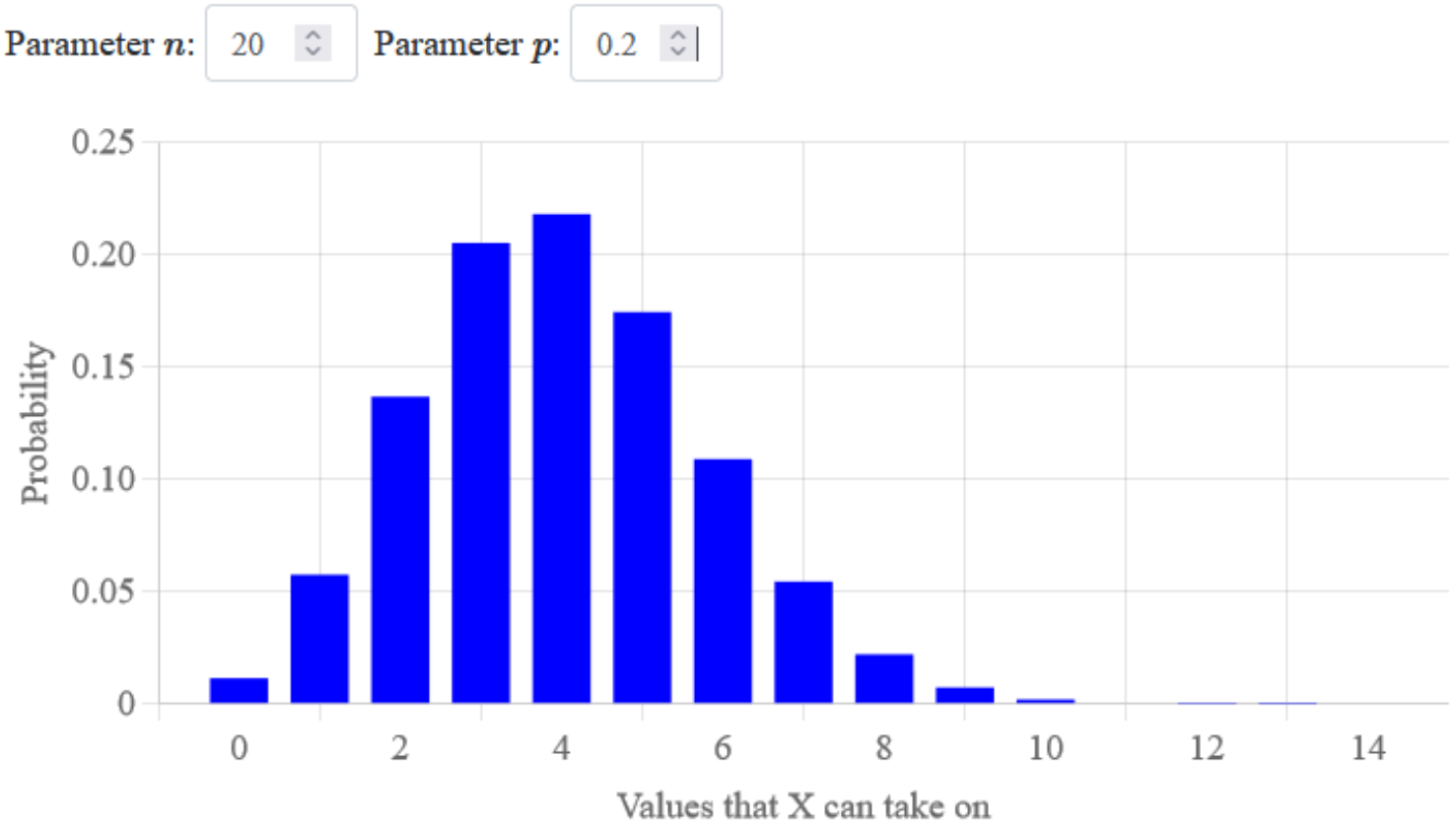
Binomial PMF



Binomial PMF



Binomial PMF



Coins, now with Binomial

Three fair ("heads" with $p = 0.5$) coins are flipped

- X is number of heads
- $X \sim Bin\,(n = 3, p = 0.5)$

$$P\,(X = 0) = \binom{3}{0} p^0\,(1-p)^3 = \frac{1}{8}$$

$$P\,(X = 1) = \binom{3}{1} p^1\,(1-p)^2 = \frac{3}{8}$$

$$P\,(X = 2) = \binom{3}{2} p^2\,(1-p)^1 = \frac{3}{8}$$

$$P\,(X = 3) = \binom{3}{3} p^3\,(1-p)^0 = \frac{1}{8}$$

Adds Click

1000 ads served, each clicked with $p = 0.01$, otherwise ignored.
What is the probability of 10 clicks?

$\boldsymbol{H}$: number of clicks

$$\boldsymbol{H} \sim \text{Bin}\,(n = 1000, p = 0.01)$$

$$P\,(\boldsymbol{H} = k) = \binom{1000}{k} (0.01)^k\,(0.99)^{1000-k}$$

$$P\,(\boldsymbol{H} = 10) = \binom{1000}{10} (0.01)^{10}\,(0.99)^{990}$$

Adds Click

1000 ads served, each clicked with $p = 0.01$, otherwise ignored.
What is the probability of less than 10 clicks?

Adds Click

1000 ads served, each clicked with $p = 0.01$, otherwise ignored.
What is the probability of less than 10 clicks?

$$P\,(\text{less than } 10) = P\,(\boldsymbol{H} = 0 \text{ or } \boldsymbol{H} = 1 \text{ or } \cdots \text{ or } \boldsymbol{H} = 9)$$

Adds Click

1000 ads served, each clicked with $p = 0.01$, otherwise ignored.
What is the probability of less than 10 clicks?

$$P(\text{less than } 10) = P(\boldsymbol{H} = 0 \text{ or } \boldsymbol{H} = 1 \text{ or } \cdots \text{ or } \boldsymbol{H} = 9)$$
$$\equiv P(\boldsymbol{H} = 0) + P(\boldsymbol{H} = 1) + P(\boldsymbol{H} = 2) + P(\boldsymbol{H} = 3) + P(\boldsymbol{H} = 4)$$
$$+ P(\boldsymbol{H} = 5) + P(\boldsymbol{H} = 6) + P(\boldsymbol{H} = 7) + P(\boldsymbol{H} = 8) + P(\boldsymbol{H} = 9)$$

Adds Click

1000 ads served, each clicked with $p = 0.01$, otherwise ignored.
What is the probability of less than 10 clicks?

$$
\begin{aligned}
&P(\text{less than } 10) = P(\boldsymbol{H} = 0 \text{ or } \boldsymbol{H} = 1 \text{ or } \cdots \text{ or } \boldsymbol{H} = 9) \\
&\equiv P(\boldsymbol{H} = 0) + P(\boldsymbol{H} = 1) + P(\boldsymbol{H} = 2) + P(\boldsymbol{H} = 3) + P(\boldsymbol{H} = 4) \\
&+ P(\boldsymbol{H} = 5) + P(\boldsymbol{H} = 6) + P(\boldsymbol{H} = 7) + P(\boldsymbol{H} = 8) + P(\boldsymbol{H} = 9) \\
&= \sum_{k=0}^{9} P(\boldsymbol{H} = k)
\end{aligned}
$$

Adds Click

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

```
from scipy import stats
stats.binom.pmf(k,n,p)
stats.binom.pmf(10,1000,0.01)
```

How Many Servers Crash?

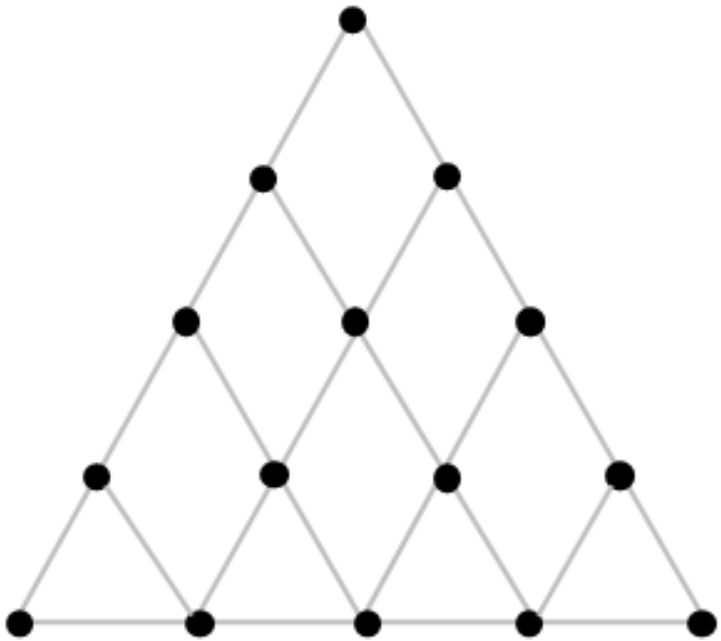
7 runs of program, each run crashes with probability 0.3. What is the chance of exactly 3 crashes?

How Many Servers Crash?

7 runs of program, each run crashes with probability 0.3. What is the chance of exactly 3 crashes?

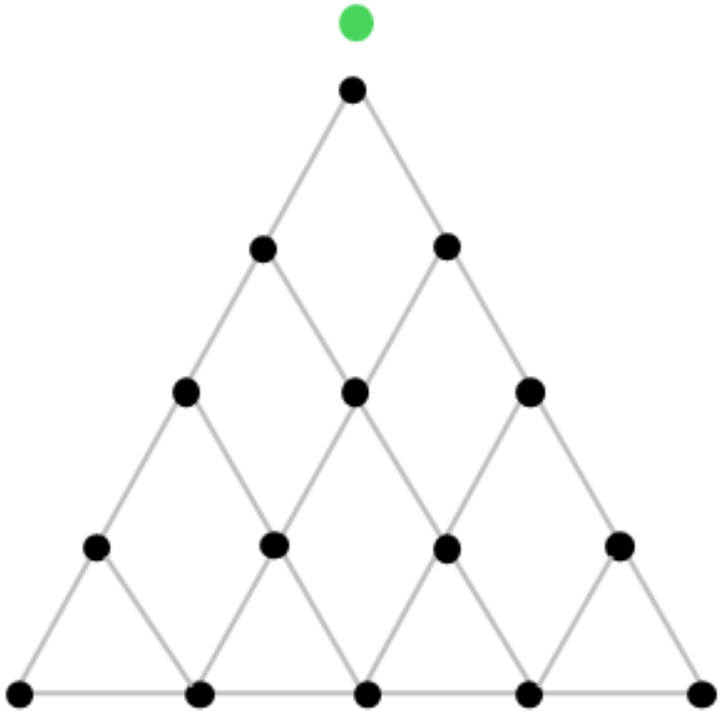
$$X : \# \text{ of crashes}$$
$$X \sim \text{Bin}\,(n = 7, p = 0.3)$$
$$P\,(X = 3) = \binom{7}{3} (0.3)^3 (0.7)^4$$

Galton Board



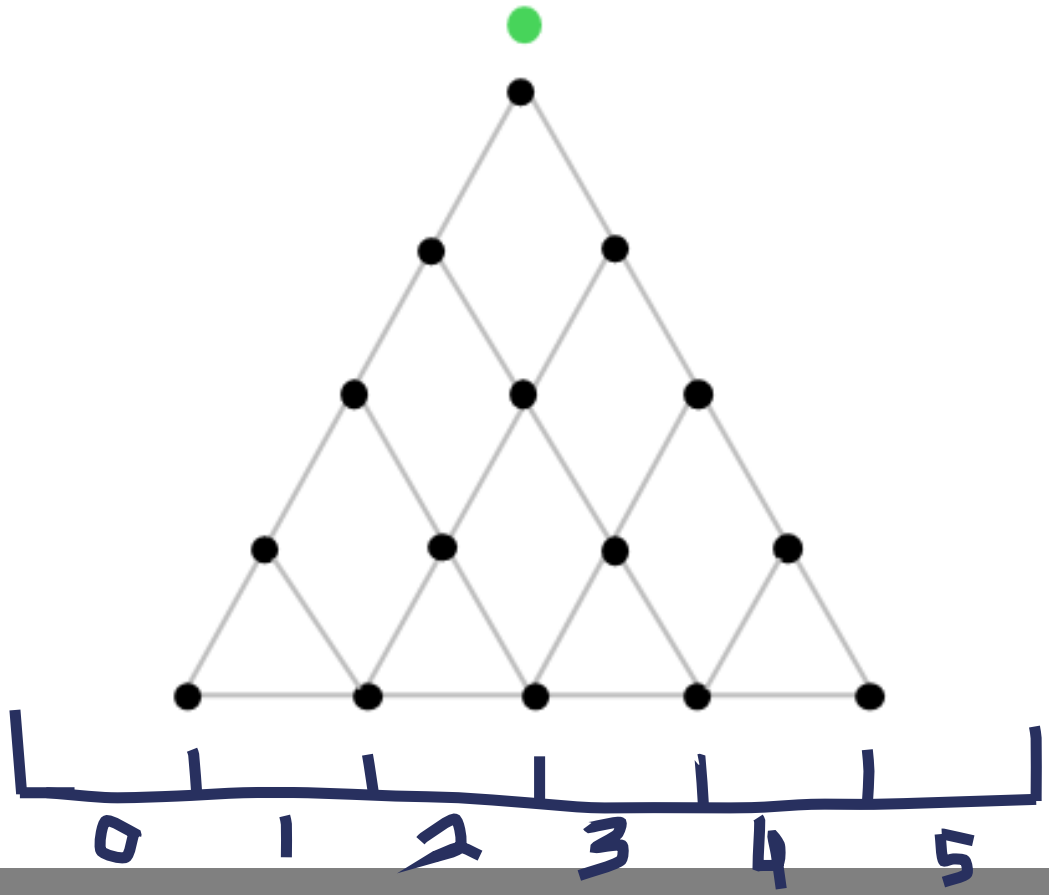
Galton Board

We drop a ball on the pins. There is equal chance of going left or right.



Galton Board

We drop a ball on the pins. There is equal chance of going left or right.
Each pin represents an independent event.



Galton Board

Assume X be the random variable that represents the number of rightward (leftward) moves.
Calculate the probability of a ball landing in a bucket.

Galton Board

Assume X be the random variable that represents the number of rightward (leftward) moves.
Calculate the probability of a ball landing in a bucket.

$$X \sim \text{Bin}\,(n = 5, p = 0.5)$$

Galton Board

Assume X be the random variable that represents the number of rightward (leftward) moves.
Calculate the probability of a ball landing in a bucket.

$$X \sim \text{Bin}\,(n = 5, p = 0.5)$$

X could also represent bucket index (why?)

Bernoulli vs Binomial

Bernoulli is an indicator RV

Binomial is the sum of n Bernoullis

Expectation of Binomial

Let $X \sim \text{Bin}(n, p)$. Let Y_i be 1 if trial i was a success.
$Y_i \sim \text{Bern}(p)$

Expectation of Binomial

Let $X \sim \text{Bin}(n, p)$. Let Y_i be 1 if trial i was a success.
$Y_i \sim \text{Bern}(p)$

$$E[X] = E[\sum_{i=1}^{n} Y_i]$$

Expectation of Binomial

Let $X \sim \text{Bin}(n, p)$. Let Y_i be 1 if trial i was a success.
$Y_i \sim \text{Bern}(p)$

$$E[X] = E[\sum_{i=1}^{n} Y_i]$$

Proof:

$$\text{Since } X = \sum_{i=1}^{n} Y_i \Longrightarrow E[X] = \sum_{i=1}^{n} E[Y_i]$$

Expectation of Binomial

Let $X \sim \text{Bin}(n, p)$. Let Y_i be 1 if trial i was a success.
$Y_i \sim \text{Bern}(p)$

$$E[X] = E[\sum_{i=1}^{n} Y_i]$$

Proof:

$$\text{Since } X = \sum_{i=1}^{n} Y_i \Longrightarrow E[X] = \sum_{i=1}^{n} E[Y_i] = n.p$$

Variance of Binomial

Let $X \sim \text{Bin}(n, p)$. Let Y_i be 1 if trial i was a success.
$Y_i \sim \text{Bern}(p)$

$$Var(X) = Var(\sum_{i=1}^{n} Y_i) = \sum_{i=1}^{n} (Var(Y_i))$$

$$\Longrightarrow Var(X) = n.p.(1-p)$$

Summary of Bernoulli and Binomial

Binomial Random Variable

Notation: $X \sim \text{Bin}(n, p)$

Description: Number of "successes" in n identical, independent experiments each with probability of success p.

Parameters: $n \in \{0, 1, \ldots\}$, the number of experiments.
$p \in [0, 1]$, the probability that a single experiment gives a "success".

Support: $x \in \{0, 1, \ldots, n\}$

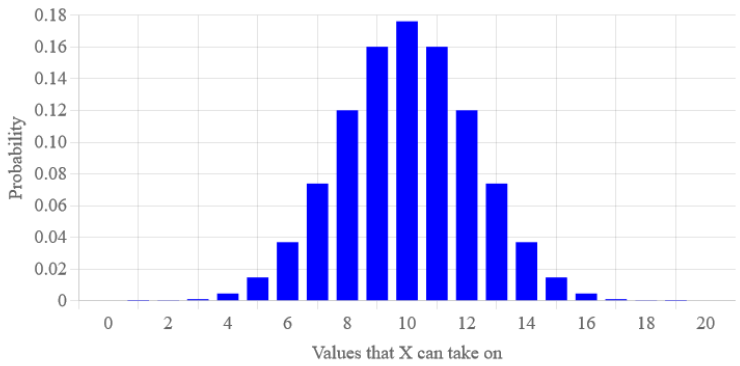
PMF equation: $\mathrm{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$

Expectation: $\mathrm{E}[X] = n \cdot p$

Variance: $\mathrm{Var}(X) = n \cdot p \cdot (1-p)$

PMF graph:

Parameter n: 20 Parameter p: 0.5



Bernoulli Random Variable

Notation: $X \sim \text{Bern}(p)$

Description: A boolean variable that is 1 with probability p

Parameters: p, the probability that $X = 1$.

Support: x is either 0 or 1

PMF equation: $\mathrm{P}(X = x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = 0 \end{cases}$

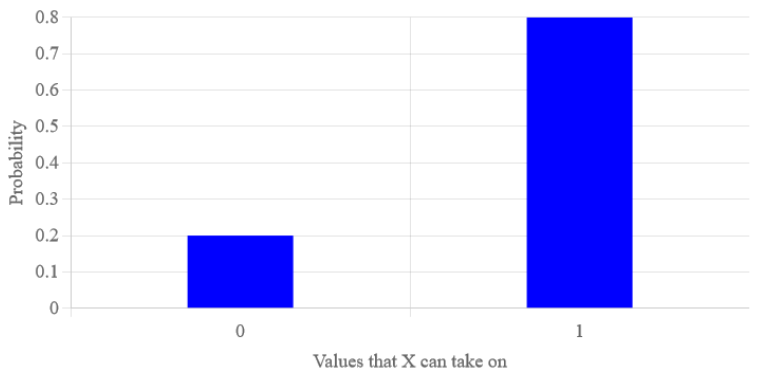
PMF (smooth): $\mathrm{P}(X = x) = p^x (1-p)^{1-x}$

Expectation: $\mathrm{E}[X] = p$

Variance: $\mathrm{Var}(X) = p(1-p)$

PMF graph:

Parameter p: 0.80



Feedback Time!

Help us improve the course quality by giving us feedback here

