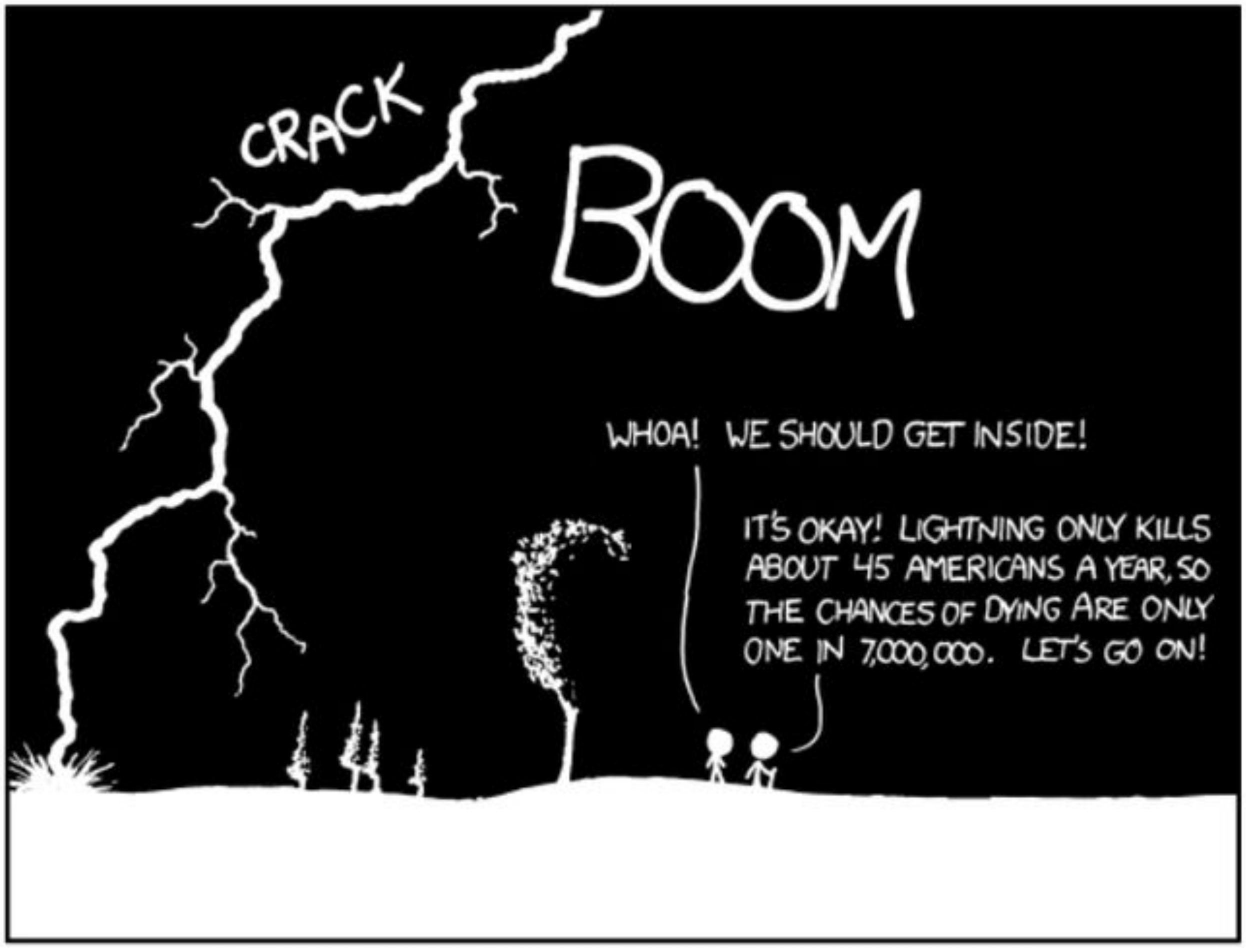
This slide deck posted under Files > Lecture Slides & Handouts > Annotated slides (Diaz)

L25:
Random Variables and Expectations



THE ANNUAL DEATH RATE AMONG PEOPLE WHO KNOW THAT STATISTIC IS ONE IN SIX.

Looking for extra, free study support for EECS 203, 280, or 281?

CSE is hosting collaborative study halls on **April 4-6 and 11-13** for these courses!

IA staff & instructional faculty have been invited to attend to help students solve problems. More info can be found by scanning this QR code

Next 203 “Study Hall” is Monday (4/11), 4:30-6:30pm in EECS 4440

Learning Objectives

After today's lecture (and associated readings, discussion & homework), you should be able to:

- *Events and random variables.*
- *Expectation of a random variable.*
- *Linearity of Expectation*
- *Indicator variables*
- *Expectations of geometric random variables*

Outline:

- **Conditional Probability**
- **Independent Events**
- Birthday Problem

- Random Variables
- Expected Value
  - Linearity of Expectation
  - Independence of random variables
  - Indicator random variables
- Geometric Random Variables
  - aka “The Waiting Time” experiment
- Bernoulli Trials & Binomial Distribution (next time)

Discrete Probability Terminology

- **Experiment:** some procedure that yields an ***outcome***.
  - E.g., flip a penny, a nickel, and a quarter.
- **Sample space:** countable set $\boldsymbol{S}$ of all possible outcomes.
  - E.g., {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
- **Event:** a subset of the sample space $\boldsymbol{S}$.
  - "Penny is heads" : {HHH, HHT, HTH, HTT}
  - "Exactly two tails" : {HTT, THT, TTH}
- **Probability distribution**: $p : \boldsymbol{S} \rightarrow [0,1]$
  - Must have $\sum_{s \in \boldsymbol{S}} p(s) = 1$
  - Probability of event E: $p(E) = \sum_{s \in E} p(s)$
  - When all outcomes in S are equally likely: $p(E) = |E|/|S|$

Conditional Probability

Let E and F be events with $p(F) > 0$. The **conditional probability** of E given F, represented as $p(E \mid F)$, is

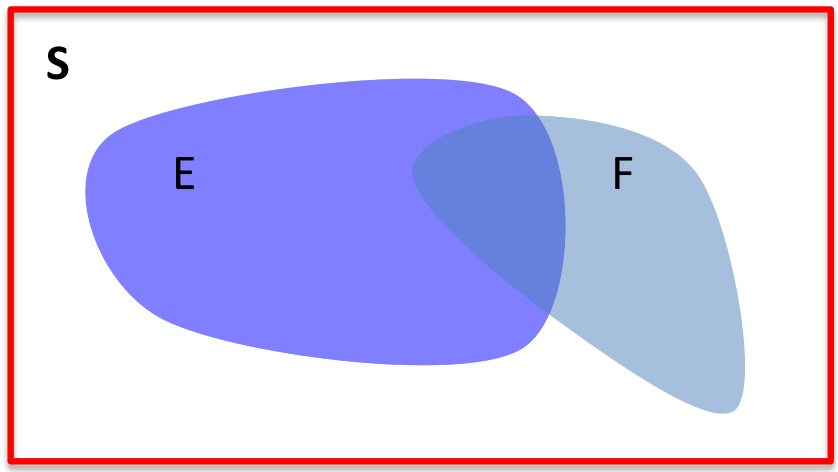
$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{|E \cap F|}{|F|}$$

Read as "probability of E given F"

For a sample space S with *equally likely* outcomes

Meaning of conditional probability:

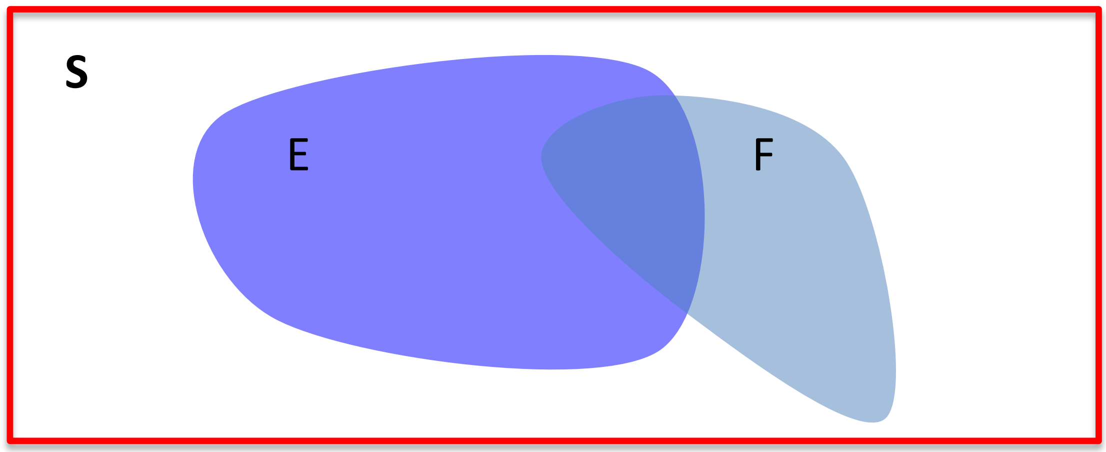
- Given that F happened, what is the probability that E happened?
- Essentially: ***F* becomes our new sample space** (because we are given that F happened)
- The only part of event E that is in our new sample space F is $E \cap F$
- So **our new event is $E \cap F$**



Conditional Probability

The **conditional probability** of event E given event F is

$$p(\_\_\_\_\_\_\_) =$$



Example: Flip a coin 3 times.

- E = total of two Heads
- F = first flip was Tails

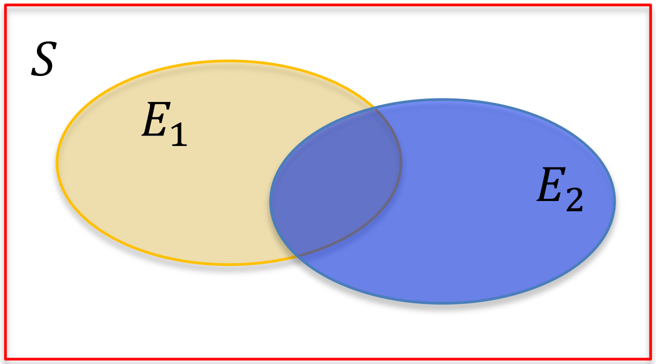
Find p(E), p(F), p(E ∩ F), p(E|F), p(F|E)

Independent Events

- **Independence.** Events E and F are ***independent*** iff

$$p(E|F) = p(E)$$

In words: knowing that F happened doesn't change the probability that E happened.



- **Alternative Definition:** E and F are are ***independent*** iff

$$= p(E \cap F) = p(E) \cdot p(F)$$

Independence of Events

Events E and F are **independent** if and only if

$$p(E|F) = \_\_\_\_\_\_\_\_\_\_$$

(alternatively)

E and F are **independent** if and only if

$$p(E \cap F) = \_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

Are these independent?

- Roll two dice
  - E: the sum of the two dice is 5
  - F: the first die is a 1
  -

- Roll two dice
  - E: the sum of the two dice is 7
  - F: the first die is a 1

Are these independent?

- Roll two dice
  - E: the sum of the two dice is 5
  - F: the first die is a 1

E = {(1,4), (2,3), (3,2), (4,1)},
F = {(1,1),(1,2),(1,3),(1,4),(1,5),(1,6)},

p(E) = 4/36 = 1/9, p(F) = 1/6
*p(E|F) = 1/6 $\neq$ p(E) **Not Independent**
*Or via: p($E \cap F$) = 1/36 $\neq$ 1/9 * 1/6 = p(E)p(F)

- Roll two dice
  - E: the sum of the two dice is 7
  - F: the first die is a 1

p(E) = 6/36 = 1/6
p(F) = 1/6
* p(E|F) = 1/6 = p(E) **Independent**
*Or via: p($E \cap F$) = 1/36 = 1/6 * 1/6 = p(E)p(F)

Outline:

- Conditional Probability
- Independent Events
- **Birthday Problem**
- Random Variables
- Expected Value
  - Linearity of Expectation
  - Independence of random variables
  - Indicator random variables
- Geometric Random Variables
  - aka “The Waiting Time” experiment
- Bernoulli Trials & Binomial Distribution (next time)

Birthday Problem

How many people have to be in a room to assure that the probability that at least two of them have the same birthday is greater than 1/2?

Assume 366 possible birthdays, each equally likely

a) 23
b) 183
c) 365
d) 730

Birthday Problem

How many people have to be in a room to assure that the probability that at least two of them have the same birthday is greater than 1/2?

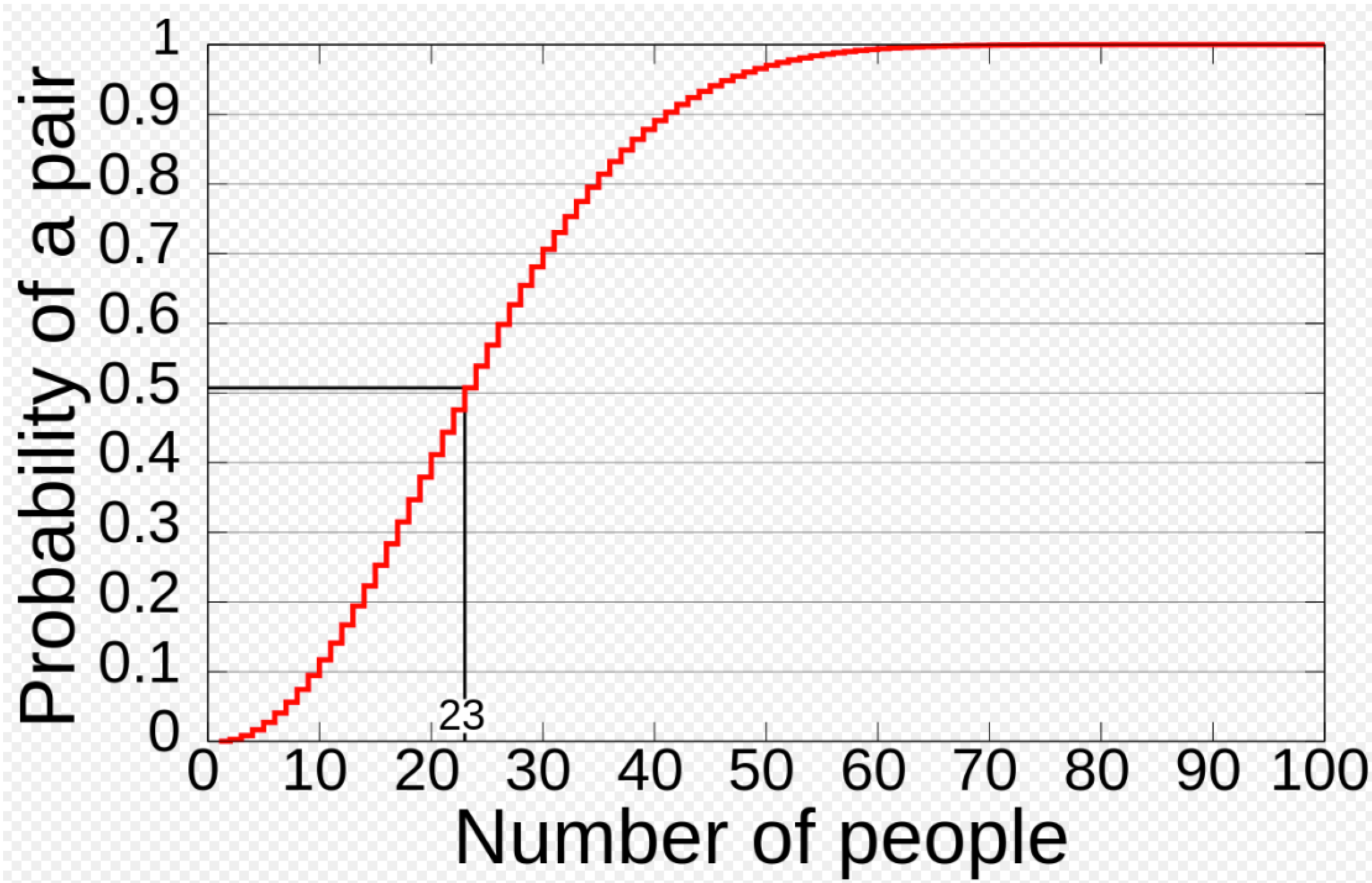
Assume 366 possible birthdays, each equally likely

a) 23
b) 183
c) 365
d) 730

- Let p_n be the probability that no people share a birthday among n people in a room.
  - Then **$1 - p_n$** is the probability that 2 or more share a birthday.
  - We want the smallest n so that $1 - p_n > 1/2$.
- Sample space S: given n people (in some order), #ways of assigning a birthday to each person. $|S| = 366^n$
- E = "no people share a birthday among n people" $|E| = P(366,n)$

$$p_n = p(E) = \frac{P(366,n)}{366^n} \qquad \text{when } \boldsymbol{n} = \mathbf{23},\ 1 - p_n \approx 0.506$$

Birthday Problem



By Rajkiran g - Own work, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=10784025

Outline:

- Conditional Probability
- Independent Events
- Birthday Problem

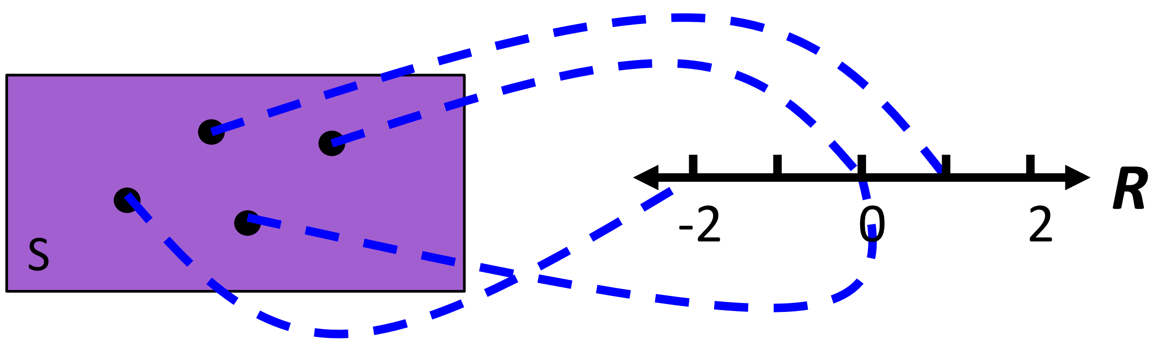
- **Random Variables**
- **Expected Value**
  - Linearity of Expectation
  - Independence of random variables
  - Indicator random variables
- Geometric Random Variables
  - aka “The Waiting Time” experiment
- Bernoulli Trials & Binomial Distribution (next time)

New concept: ***Random Variable***

- Outcomes of an experiment can be anything: heads/tails, numbers rolled on dice, random permutation, random 5-card hand, etc.
  - We often want to interpret the outcome ***numerically*** ("what's it worth?").

> A **random variable** is (technically) a function from a sample space to the real numbers

$X : S \rightarrow \boldsymbol{R}$



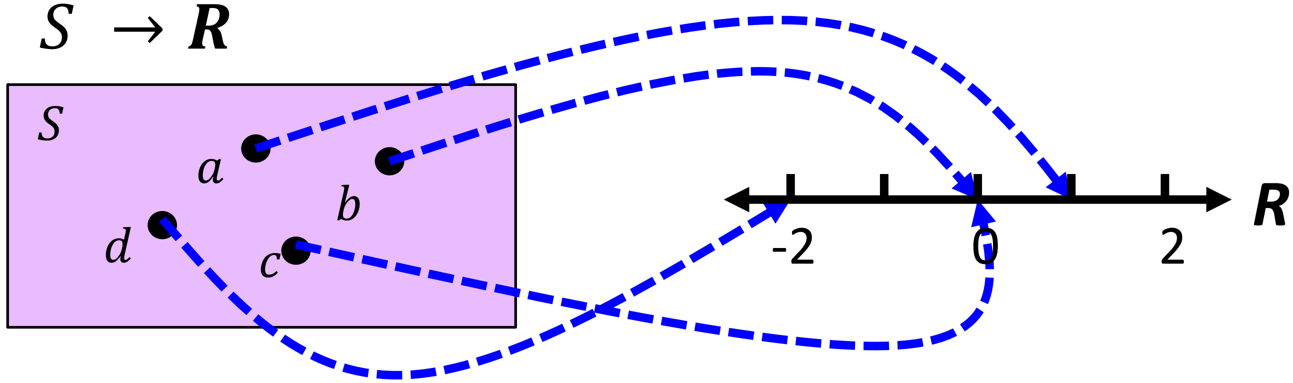
Intuitively:

A random variable is a **numerical measurement** about the outcome of an "experiment."

Random Variables

A **random variable** is a _______________ from a _________________________ to the __________________________

$X : S \to \boldsymbol{R}$



$\boldsymbol{X = r}$ is the set of all outcomes that map to r

Example: Roll 2 dice. Let X = sum of the dice. S is set of pairs.

- X((2,3)) =
- "X = 7" =
- "X > 10" =

$p(X = 7) =$ $p(X > 10) =$

Random Variable: Sums of 2 dice

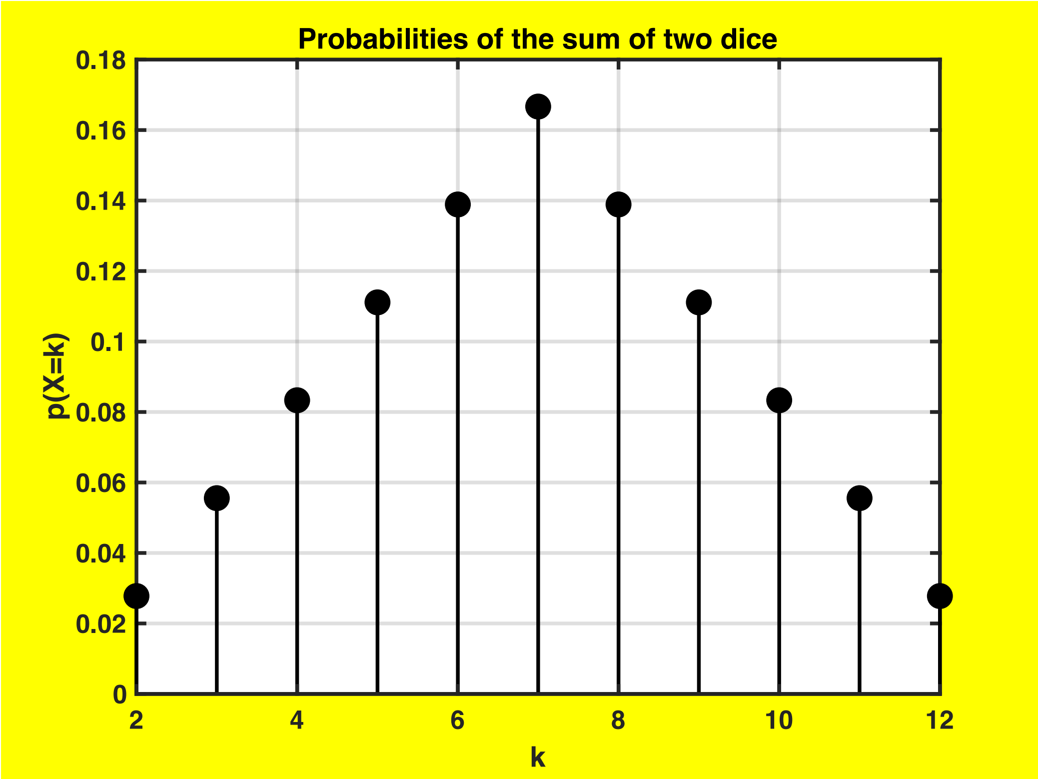
- Roll 2 dice. X(s) = sum of numbers on outcome s.
  (That is, $X = r$ is the event set of all outcomes that map to r.)
  - X((1,1))= 2,
  - X((1,2))= X((2,1))= 3,
  - X((1,3))= X((2,2))= X((3,1))= 4,
  - X((1,4))= X((2,3))= X((3,2))= X((4,1))= 5,
  - X((1,5))= X((2,4))= X((3,3))= X((4,2))= X((5,1))= 6,
  - X((1,6))= X((2,5))= X((3,4))= X((4,3))= X((5,2))= X((6,1))= 7,
  - X((2,6))= X((3,5))= X((4,4))= X((5,3))= X((6,2))= 8,
  - X((3,6))= X((4,5))= X((5,4))= X((6,3))= 9,
  - X((4,6))= X((5,5))= X((6,4))= 10,
  - X((5,6))= X((6,5))= 11,
  - X((6,6))= 12

- Events
  - "X=7" = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}
  - "X>10" = {(5,6), (6,5), (6,6)}
  - "X=7 or X=11" = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1), (5,6), (6,5)}

- Probability of events
  - p(X=7) = 1/6
  - p(X >10) = 3/36 = 1/12
  - p(X=7 or X = 11) = 8/36

Random variable: Sum of two dice

- X((1,1))= 2,
- X((1,2))= X((2,1))= 3,
- X((1,3))= X((2,2))= X((3,1))= 4,
- X((1,4))= X((2,3))= X((3,2))= X((4,1))= 5,
- X((1,5))= X((2,4))= X((3,3))= X((4,2))= X((5,1))= 6,
- X((1,6))= X((2,5))= X((3,4))= X((4,3))= X((5,2))= X((6,1))= 7,
- X((2,6))= X((3,5))= X((4,4))= X((5,3))= X((6,2))= 8,
- X((3,6))= X((4,5))= X((5,4))= X((6,3))= 9,
- X((4,6))= X((5,5))= X((6,4))= 10,
- X((5,6))= X((6,5))= 11,
- X((6,6))= 12.



Random Variable Examples

Experiment: roll a (fair) die.

I'll pay you \$5 for a 1, \$10 for a 2, but you pay me \$3 for any other number rolled.

Define the random variables X, Y, Z as follows:

X = the number you rolled,   Y = amount you win/lose

Z = 1 if you win money, 0 if not

s	$p(s)$	$X(s)$	$Y(s)$	$Z(s)$

Random Variable Examples

Experiment: roll a (fair) die.

I'll pay you \$5 for a 1, \$10 for a 2, but you pay me \$3 for any other number rolled.

Define the random variables X, Y, Z as follows:

X = the number you rolled,   Y = amount you win/lose

Z = 1 if you win money, 0 if not

s	$p(s)$	$X(s)$	$Y(s)$	$Z(s)$
1	1/6	1	5	1
2	1/6	2	10	1
3	1/6	3	-3	0
4	1/6	4	-3	0
5	1/6	5	-3	0
6	1/6	6	-3	0

Should you play this game?
How much will you win (or lose) on average?

Random Variable Examples

Experiment: roll an unfair die. p(6) = ½ and uniform probability over other rolls

I'll pay you $5 for a 1, $10 for a 2, but you pay me $3 for any other number rolled.

Define the random variables X, Y, Z as follows:

X = the number you rolled, Y = amount you win/lose

Z = 1 if you win money, 0 if not

s	$p(s)$	$X(s)$	$Y(s)$	$Z(s)$
1		1	5	1
2		2	10	1
3		3	-3	0
4		4	-3	0
5		5	-3	0
6		6	-3	0

Should you play this game?
How much will you win (or lose) on average?

Expected Value

- Let p be a probability distribution over S.
- The ***expected value*** of $X: S \to \boldsymbol{R}$ is the average of X over S, according to ("weighted by") the distribution p:

$$E(X) = \sum_{s \in S} p(s) \cdot X(s) = \sum_{r \in Range(X)} p(X = r) \cdot r$$

This sum is over the ***outcomes*** in sample space S

This sum is over the possible ***values*** that X can take.

Expected Value

The ***expected value*** of $X: S \to \boldsymbol{R}$ is the ____________ ____________ of X

Two ways to find E(X):

- $E(X) =$ (weighted sum over outcomes)

- $E(X) =$ (weighted sum over range of X)

Binomial RV (next time)

- Let B = # of successes in n Bernoulli trials
- E(B) = ____________

Geometric RV

- Let X = # of Bernoulli trials until first success
- E(X) = ____________

Exercises: Expected Value

- Expected value of a roll of a (fair) die:

- Expected amount you win with the unfair die from previous example

Exercises: Expected Value

- Expected value of the sum of 2 (fair) dice.
  - |S| = 36
  - X = {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}, so |X| = 11

$$E(X) = \sum_{s \in S} p(s) \cdot X(s)$$ would add up 36 values (every outcome in S)

$$E(X) = \sum_{r} p(X = r) \cdot r$$ would add up 11 values (every value of X)

A bit more algebra than is convenient, but either way
E(X) = 7

Outline:

- Conditional Probability
- Independent Events
- Birthday Problem

- Random Variables
- Expected Value
  - **Linearity of Expectation**
  - **Independence of random variables**
  - **Indicator random variables**
- Geometric Random Variables
  - aka “The Waiting Time” experiment
- Bernoulli Trials & Binomial Distribution (next time)

Linearity of Expectation

makes your life easier

- ***Theorem.*** If X and Y are random variables,
  - $(X + Y)(s) = X(s) + Y(s)$ is also a random variable, and

$$E(X + Y) = E(X) + E(Y)$$

Corollary: $aX + b$ is also a random variable, and
$$E(aX + b) = aE(X) + b$$

- **Proof.**

$$\begin{aligned} E(X + Y) &= \textstyle\sum_s p(s)(X(s) + Y(s)) \\ &= \textstyle\sum_s (p(s)X(s) + p(s)Y(s)) \\ &= \textstyle\sum_s p(s)X(s) + \sum_s p(s)Y(s) \\ &= E(X) + E(Y) \end{aligned}$$

Linearity of Expectation

- The expected value of the sum of random variables is the

  ____

$$E(X + Y) =$$

$E(aX + b) =$ for any constants a, b

  - Does linearity require that X and Y be *independent*?
  - Does $E(XY) = E(X)E(Y)$?

Example : X = sum of two dice, i.e., X((a,b)) = a+b

Linearity of Expectations… Examples

- The expected value of the sum of random variables is the sum of their expectations.

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

$$E(a \cdot X + b) = a \cdot E(X) + b$$ for any constants a,b

- Example 1: X = sum of two dice, i.e., X((a,b)) = a+b
  - X_1 = outcome of 1^{st} die, X_2 = outcome of 2^{nd} die
  - $X = X_1+X_2$.
  - $E(X) = E(X_1+X_2)$
  - $= E(X_1) + E(X_2)$
  - $= 7/2 + 7/2$
  - $= 7$

Indicator Variables

- You can convert a ***random variable*** X to an ***event***, e.g., "$X = 5$" or "$X \in \{1,3,5\}$."
- You can also convert an ***event*** into a ***random variable***.
- If F is an event, the ***indicator variable*** I_F is:
  - $I_F = \begin{cases} 1 & \text{if } F \text{ holds} \\ 0 & \text{if } F \text{ does not hold} \end{cases}$
- What is $E(I_F) =$???
  - $E(I_F) = 0 \cdot p + 1 \cdot \mathrm{p}(I_F = 1) = p(F)$.

Indicator Variables

- If F is an event, the ***indicator variable*** I_F is:

  - $I_F =$

I **indicates** whether event F happened or not

- $E(I_F) =$
- Together with Linearity of Expectation, Indicator variables can help us determine the expected number of _______________ in repeated trials of _______________ experiment.

- **Example:** Y = number of 7s in 12 rolls of two dice. Find E(Y).
  - Consider each roll individually
    - $Y_i = \{$

$E(Y_i) =$

$E(Y) =$

More fun with Indicator Variables

- See “Balls and Bins” slides at the end of the main Lec 25 slides.

Does "linearity" work with multiplication too?

- Linearity of Expectation:

$$E(X+Y) = E(X) + E(Y)$$

- ***This holds <u>regardless</u> of whether X and Y are independent!***
- Multiplication is just "repeated addition", so by LoE we have (?)

$$E(XY) = E(X) \cdot E(Y)$$

Not necessarily!

- ***This holds <u>only if</u> X and Y are independent!***
- <u>Example:</u> Suppose you have a six-sided die where the number opposite $x \in \{1, \dots, 6\}$ is $7 - x$. X = number on top; Y = number of bottom. ***Note: X and Y are not independent!***
  - $E(XY) = \frac{1}{6}(1 \cdot 6 + 2 \cdot 5 + 3 \cdot 4 + 4 \cdot 3 + 5 \cdot 2 + 6 \cdot 1) = \frac{56}{6}$.
  - $E(X)E(Y) = \left(\frac{7}{2}\right)^2 = \frac{49}{4}$. Note: $E(X+Y) = 7 = E(X) + E(Y)$.

Independent Random Variables

- ***Events*** E, F are ***independent*** if $\mathrm{p}(E \cap F) = \mathrm{p}(E) \cdot \mathrm{p}(F)$.
- ***Random variables*** X, Y are ***independent*** if any events derived from X, Y are independent.
  - E.g., $E =$ event that $X \leq 5$; $F =$ event that Y is an odd integer.
  - If X, Y are independent then $\mathrm{Pr}(E \cap F) = \mathrm{Pr}(E) \cdot \mathrm{Pr}(F)$

Multiplying Random Variables

- Random variables X and Y on a sample space S are ***independent*** if the events "X(s)=a" and "Y(s)=b" are independent, for all numbers a,b.

Linearity of expectation: works even if $\mathbf{X_1}$, $\mathbf{X_2}$ are dependent!

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

$$E(a \cdot X + b) = a \cdot E(X) + b$$

$$\mathbf{E(X_1\, X_2) = E(X_1)\, E(X_2)}$$

only if $\mathbf{X_1}$, $\mathbf{X_2}$ are independent.

Outline:

- Conditional Probability
- Independent Events
- Birthday Problem

- Random Variables
- Expected Value
  - Linearity of Expectation
  - Independence of random variables
  - Indicator random variables
- **Geometric Random Variables**
  - **aka “The Waiting Time” experiment**
- Bernoulli Trials & Binomial Distribution (next time)

Geometric Random Variables

- Suppose you have a biased coin comes up Heads with probability $p \in (0,1)$ and Tails with probability $1 - p$.
- You flip repeatedly until you get the first Head. Let X be the total number of flips (including the last Head).
- What is the probability distribution, p(X = k)?
  - What is the sample space S?
  - What is p(s) and X(s) for each outcome?

Geometric Random Variables

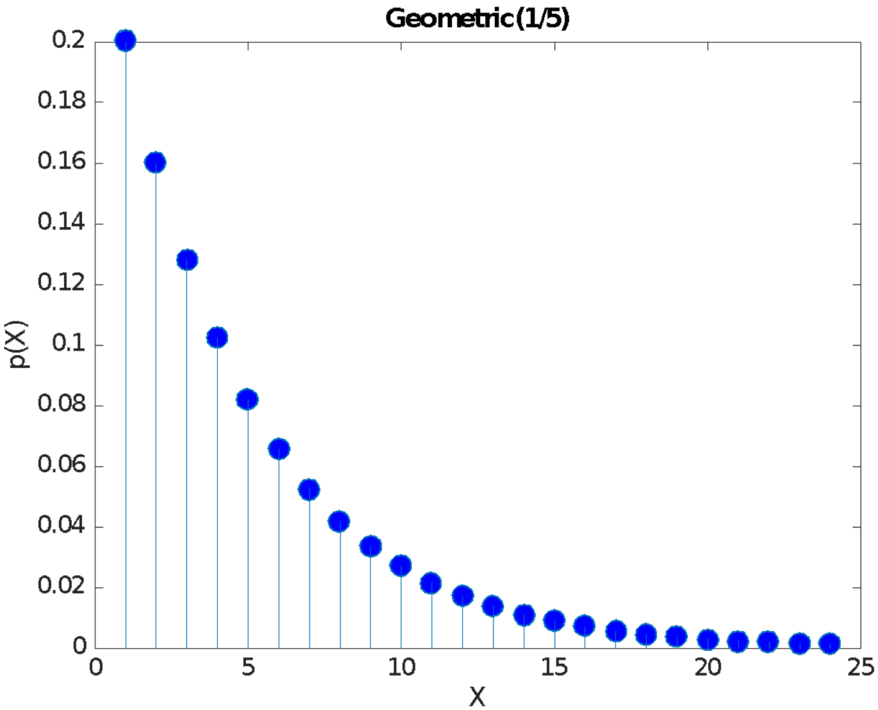
- Suppose you have a biased coin comes up Heads with probability $p \in (0,1)$ and Tails with probability $1 - p$.
- You flip repeatedly until you get the first Head. Let X be the total number of flips (including the last Head).
- What is the probability distribution, p(X = k)?
  - What is the sample space S?
  - What is p(s) and X(s) for each outcome?

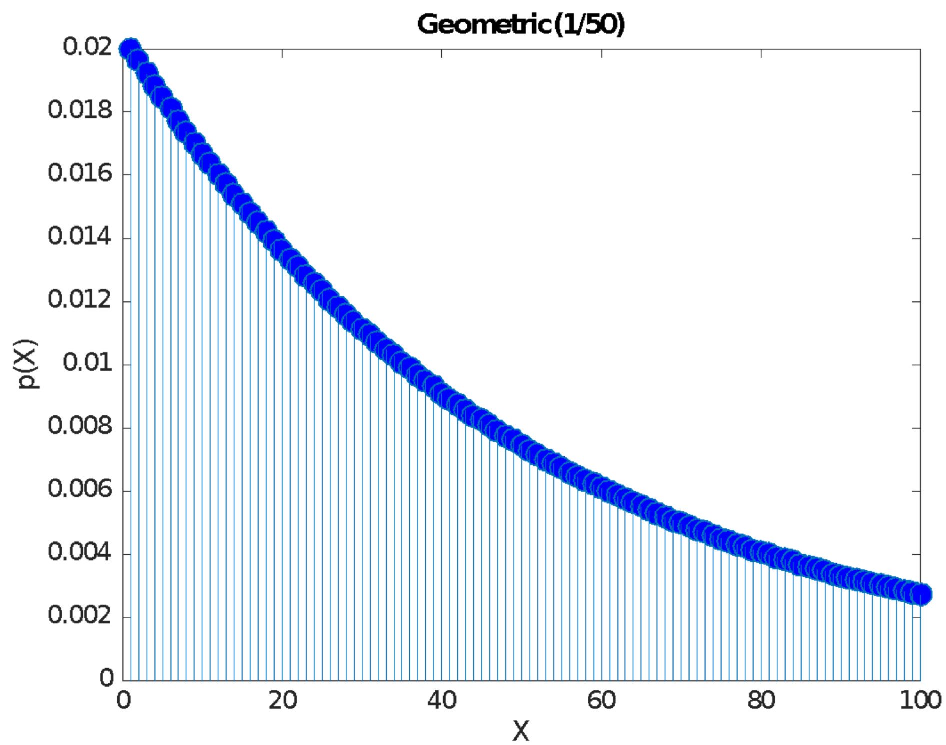
- We have a (countably) infinite sample space:

s	p(s)	X(s) = # of flips until first Head
– H	p	1
– TH	$(1-p)p$	2
– TTH	$(1-p)^2 p$	3
– …		
– $T^{k-1}H$	$(1-p)^{k-1}p$	k
– …		

$$\boxed{p(X = k) \; = (1-p)^{k-1}p}$$

Geometric Random Variables





Geometric Random Variables

- Suppose you have a biased coin comes up Heads with probability $p \in (0,1)$ and Tails with probability $1 - p$.
- You flip repeatedly until you get the first Head. Let X be the total number of flips (including the last Head).
- The probability distribution is $p(X = k) \ = (1-p)^{k-1}p$

- What is E(X)?
- Consider the first flip. It's either heads or tails.
  - If it's tails, then X = 1.
  - If it's heads, then we add 1 to our total and repeat from scratch.

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

Solve for E(X) and you get…

$$E(X) = 1/p$$

Geometric Random Variables

- Suppose you have a biased coin comes up Heads with probability $p \in (0,1)$ and Tails with probability $1 - p$.
- You flip repeatedly until you get the first Head. Let X be the total number of flips (including the last Head).
  - Geometrically Distributed

A r.v. X has a ***geometric distribution*** with parameter p if:

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

If X is geometrically distributed with parameter p then:

$$E(X) = 1/p$$

Watching Seinfeld Reruns

- Every night a Seinfeld episode is drawn ***uniformly at random*** from the 180 shows and broadcast.
- What is the ***expected number of nights*** you need to watch to see all episodes?

- This is a tricky problem if you don’t start right!
  - Use linearity of expectations.
  - You know the expectation of a geometric distribution.

Seinfeld Reruns

- Let $X_{i \to j}$ be the number of days you have to watch to go from having watched i distinct shows to having watched j distinct shows. Then,

$$X_{0\to180} = X_{0\to1} + X_{1\to2} + \cdots + X_{178\to179} + X_{179\to180}$$

By Linearity of Expectations...

$X_{k\to k+1}$ is a geometric r.v. with success probability $\dfrac{180-k}{180}$

$$E[X_{k\to k+1}] = \frac{180}{180-k}$$

$$E[X_{0\to180}] = \frac{180}{180} + \frac{180}{179} + \frac{180}{178} + \cdots + \frac{180}{2} + \frac{180}{1}$$

$$\approx 180 \cdot 5.77 < 1039$$

Outline:

- Probability Recap, so far
- Conditional Probability
- Independent Events
- Birthday Problem

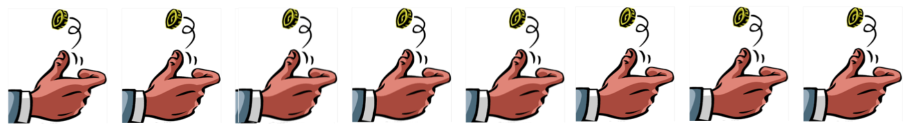
- Random Variables
- Expected Value
  - Linearity of Expectation
  - Independent Random Variables
  - Indicator RVs
- **(Bernoulli Trials & Binomial Distribution)**
- Geometric Distribution
  - Aka "The Waiting Time" experiment

- Independence of random variables? – probably skip this?
- Geometric random variables and their expectations
  - The coupon collector problem
- Indicator random variables
  - Balls'n'bins problems.

Coin-toss Example

A coin whose probability of getting heads is 2/3 is tossed 8 times.

What is the probability of exactly 3 heads in the 8 tosses?



One sequence with 3 heads: T H H T T H T T

Probability of this sequence: $\frac{1}{3}\ \frac{2}{3}\ \frac{2}{3}\ \frac{1}{3}\ \frac{1}{3}\ \frac{2}{3}\ \frac{1}{3}\ \frac{1}{3} = (2/3)^3 * (1/3)^5$

- So, what is the probability of any particular sequence with 3 heads? $(2/3)^3 * (1/3)^5$
- How many of sequences with 3 heads are there? $\binom{8}{3}$

Probability of getting 3 heads in the 8 tosses:

$$P(3\ heads) = \binom{8}{3} \left({}^{2}/_{3}\right)^{3} \left({}^{1}/_{3}\right)^{5}$$

Coin-toss Example

A coin whose probability of getting heads is 2/3 is tossed 8 times.
What is the probability of exactly 3 heads in the 8 tosses?

One sequence with 3 heads: T H H T T H T T
Probability of this sequence: $\frac{1}{3}\frac{2}{3}\frac{2}{3}\frac{1}{3}\frac{1}{3}\frac{2}{3}\frac{1}{3}\frac{1}{3} = (2/3)^3 * (1/3)^5$

- So, what is the probability of any particular sequence with 3 heads? $(2/3)^3 * (1/3)^5$
- How many of sequences wit

Probability of getting 3 head

$$P(3\ heads) = \binom{8}{3}(2$$

Take-aways from the example:

- The *Experiment* consists of repeating a simpler experiment *n* times.
- This *simpler experiment* is "simpler" because it's sample space consists of two outcomes.
- Events of interest may consist of the *number of times* one of the two simpler experimental outcomes occurs.

Bernoulli Trials, Binomial Experiment, and Binomial Distribution

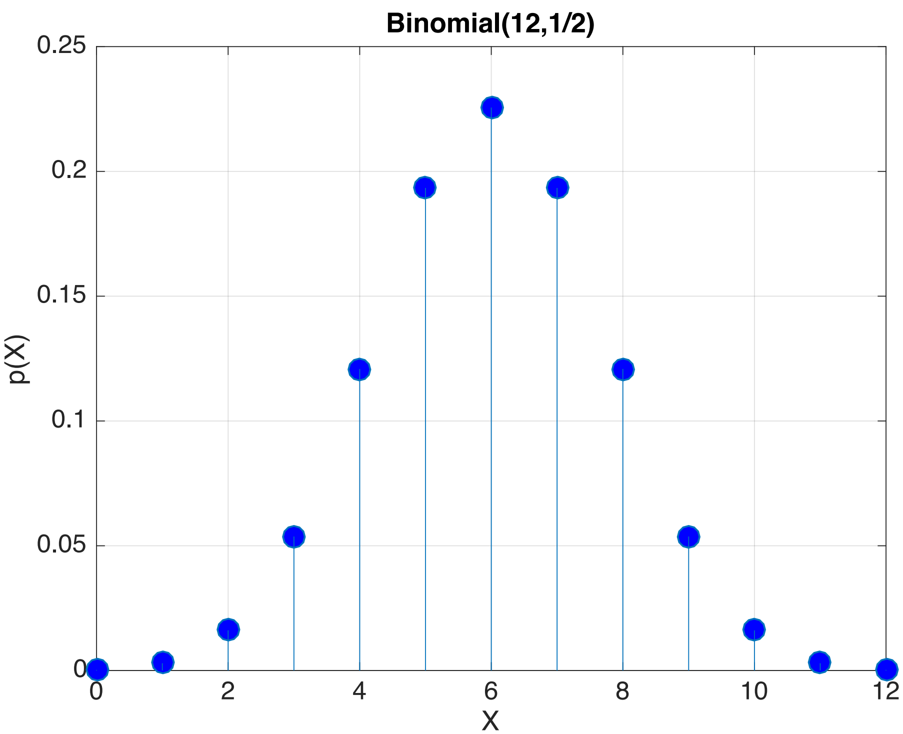
- Bernoulli Trial
  - Experiment that has exactly 2 outcomes, S = {success, failure}
    - p(success) = p
    - p(failure) = q = (1-p)
- Binomial Experiment
  - Repeat the Bernoulli trial *n* times where each trial has the same probability of success and all trials are mutally independent
- Binomial Distribution
  - Let event be that *k* successes occur in *n* Bernoulli trials. Then

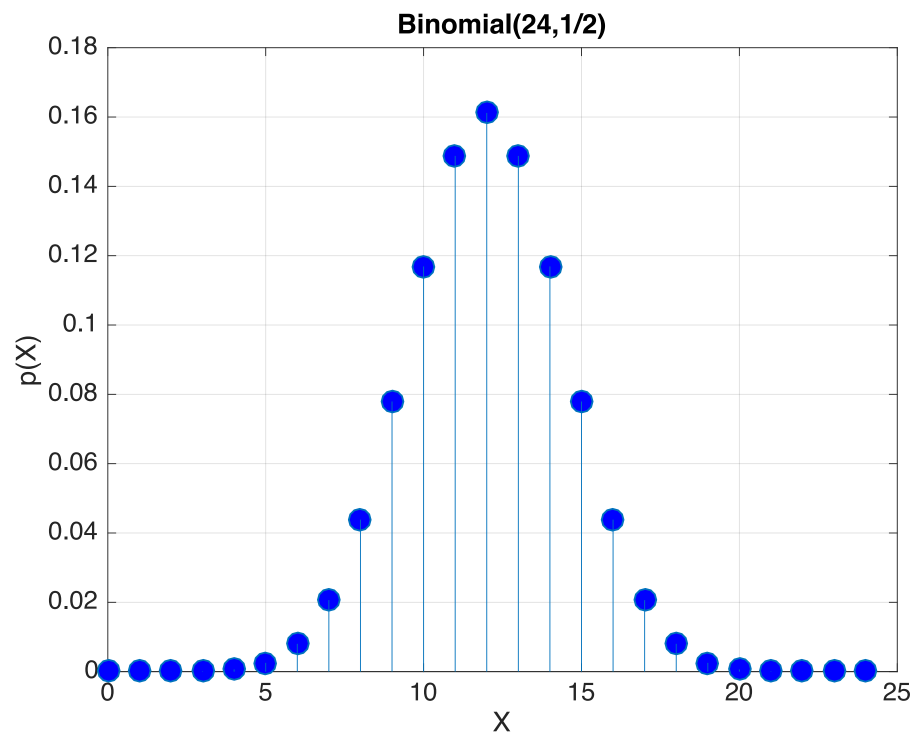
$$\boxed{\text{p(exactly } k \text{ successes in } n \text{ trials)} = \binom{n}{k} p^k q^{n-k}}$$

Random Variable: Binomial Distribution

- X is the number of successes in *n* Bernoulli trials where p(success) = 1/2

$$P(X = k) = \binom{n}{k} \frac{1}{2^n}$$





Random Variable: Binomial Distribution

- X is the number of successes in *n* Bernoulli trials where probability of a success is *p*

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