1 Probability, Random Variables, and Probability Distributions

1.1 Basic Probability and Simulations

What is Probability?

- Outcomes are governed by chance, but in many repetitions a pattern emerges.

- Sample Space: all of the possible outcomes

- Event: a specific, desired outcome or set of outcomes

- Notation: probability of Event A $\rightarrow$ P(A)

- Range of Probabilities: the probability of an event is between 0 and 1

- Sum of Probabilities: the probability of the whole sample is 1

Theoretical Probability - What should happen given the sample space.

This is the desired outcomes divided by the total outcomes.

Empirical Probability - When you are performing a simulation or experiment, it is what does happen given the trials.

This is the number of successes divided by trials.

The law of large numbers states that simulated (empirical) probabilities tend to get closer to the true (theoretical probability) as the number of trials increases.
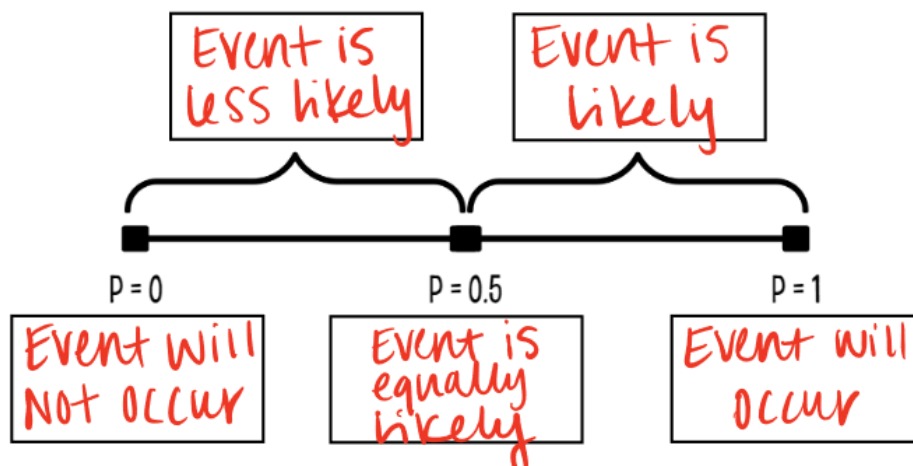
Probability can be written in the forms of fractions, decimals, or percentages.

For our class, we will write probabilities as decimals, rounded to four decimal places.

Example

A bag contains 10 marbles - 2 black, 3 blue, 1 red, and 4 white.

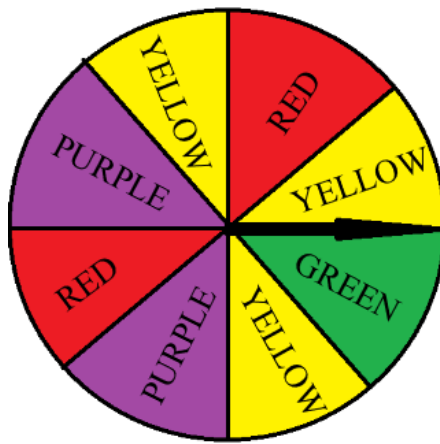(a) What is the probability you will pull out a white marble?

P(White) = 4/10 = 0.40

(b) What is the probability you will pull out a blue marble?

P(Blue) = 3/10 = 0.30

Example

In the spinner below, each wedge is equal in area.

(a) What is the probability that the spinner will land on a red wedge?

P(red) = 2/8 = 0.25

(b) What is the probability that the spinner will land on a yellow wedge?

P(yellow) = 3/8 = 0.375

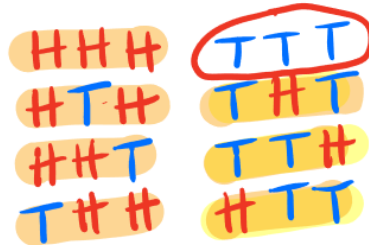The notation for the probability that an event does not occur is $P(A^c)$. We also refer to this as "Not A" in words.

The probability that an event does not occur is $P(A^c) = 1\text{-}P(A)$.

Sometimes, it is easier to compute the complement of an event, than the event itself.

Example

A coin is flipped three times and the results of each flip is noted.

(a) Draw a sample space for this event.

(b) What is the probability a series of three flips will produce exactly one head?

P(1 head) = 3/8 = 0.375

(c) What is the probability that a series of three flips will produce at least one head?

P(at least 1 Head) = 1 - P(no Heads) = 7/8 = 0.875

Myth of "Law of Averages"

- The idea of probability is that randomness is predictable in the long run.
- Probability does not allow us to make short run predictions.
- Probability tells us random behavior evens out in the long run.
- Future outcomes are not affected by past behavior.

Example

A group of 50 high school students were interviewed and asked what their favorite leisure activity is. Use the results in the table below to find and interpret the following probabilities.

	Dance	Sports	Read	Total
Male	2	10	8	20
Female	16	6	8	30
Total	18	16	16	50

(a) What is the probability of picking a male whose favorite leisure activity is dance?

P(male and dance) = 2/50 = 0.04

(b) What is the probability of picking a person whose favorite leisure activity is reading?

P(reading) = 0.32

(c) What is the probability of picking someone who does not consider dance as their favorite leisure activity?

P(not dance) = 1 - P(dance) = 0.64

(d) Given that you've selected a female, what is the probability she likes reading?

P(read — female) = 8/30 = 0.2667

The imitation of chance behavior, based on a model that accurately reflects the situation, is called a simulation. Simulations are usually done with a table of random digits, random number generator, dice, deck of cards,

spinner, etc.

Four Principles of Simulation:

State - Identify the probability calculation.

Must include:

- Identify variable.

- Statement of probability in symbols or words.

Plan - Describe how to use your chance process.

Must include:

- What tool?

- What values are you assigning?

- How many values are you picking each time?

- How many times do you run the simulation?

- What about repeat digits or ignored digits?

- What are you recording?

Do - Perform the simulation (trials typically indicated, perform at least 10 if not).

Must include:

- Simulation data, if number of trials is 10 or less.

- Summary of data for larger trials.

Conclude - Use the results of your simulation to answer the question.

Must include:

- Statement of probability.

- Answer to question.

- Usually about being surprised/reasonable/expected, etc.

Example

What is the probability that a student gets 6 out of 6 questions correct on a true/false quiz written in Greek? (Assume the exam taker does not know any Greek.) Should the instructor be concerned about cheating? Perform a simulation to answer this question.

- What is the probability of guessing 6/6 correct on a T/F quiz? 50% correct or 50% incorrect

- Label digits 0-4 as correct and 5-9 as incorrect. Using a random digit table, select 1 number at a time (repeats allowed) and record if correct or incorrect. Perform this simulation 6 times to represent one quiz. Count how many questions are "correct". Repeat for 20 trials.

# Correct	# Trials
0	0
1	2
2	3
3	8
4	5
5	2
6	0

After conducting the simulation, we see that we had no trials where 6/6 is correct by guessing. We should be suspicious of the result the student got.

1.2   The Addition Rule

Mutually Exclusive Events

- When two events have no outcomes in common, we refer to them as mutually exclusive events.
- P(A and B) = 0
- P(A or B) = P(A) + P(B)

Example

Let's revisit our "rolling the dice" game, where the italicized numbers represent the sum of the dice rolled.

Dice	1	2	3	4	5	6
1	*2*	*3*	*4*	*5*	*6*	*7*
2	*3*	*4*	*5*	*6*	*7*	*8*
3	*4*	*5*	*6*	*7*	*8*	*9*
4	*5*	*6*	*7*	*8*	*9*	*10*
5	*6*	*7*	*8*	*9*	*10*	*11*
6	*7*	*8*	*9*	*10*	*11*	*12*

Let A be rolling a sum of 5 and B be rolling two even numbers.

These events are mutually exclusive because they have no outcomes in common.

P(A) is 4/36, and P(B) is 9/36, so P(A or B) is 13/36.

Example

Identify if the following events are mutually exclusive. If they are not, given an example of where they overlap.

(a) Rolling a sum of 4 and rolling doubles.

Not mutually exclusive. Ex. 2, 2

(b) Rolling an odd sum and rolling a 3.

Not mutually exclusive. Ex. 3, 4

(c) Rolling a 4 and rolling a sum of 12

Mutually exclusive

(d) Rolling an odd number and rolling a sum of 10

Not mutually exclusive. Ex. 5, 5

Example

Find the probability of rolling doubles or a sum of 8.

Dice	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Let A be rolling doubles, and B be rolling a sum of 8.

They are not mutually exclusive events.

P(A) = 6/36 and P(B) = 5/36. We have to subtract P(A and B) = 1/36 from this, and we get P(A or B) = 10/36.

If A and B are any two events resulting from some chance process, then

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Notice how if A and B are mutually exclusive, P(A and B) = 0 and P(A or B) = P(A) + P(B) - 0

How do we know the value of P(A and B)? For now, we will use common sense and drawing out the sample space, but the general multiplication rule will allow us to find this mathematically later.

Example

A card is selected at random from a deck of 52 cards. Determine if the following events are mutually exclusive and then find the probability of the events.
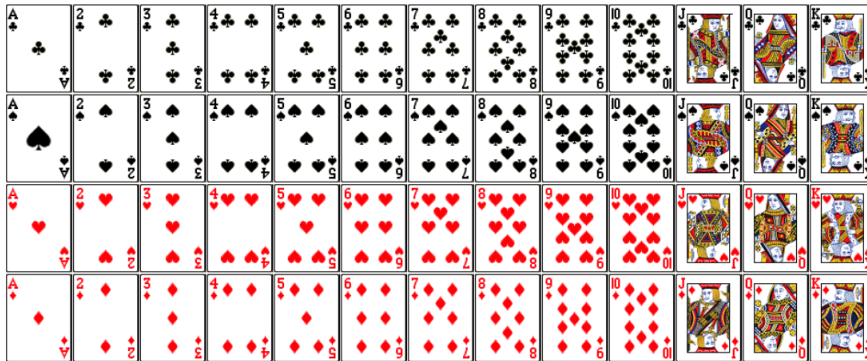
(a) P(red or Queen)

Not mutually exclusive. 26/52 + 4/52 - 2/52 = 28/52.

(b) P(Ace or King)

Mutually exclusive. 4/52 + 4/52 = 8/52

(c) P(Queen or even)

Mutually exclusive. 4/52 + 20/52 = 24/52

(d) P(black or odd)

Not mutually exclusive. 26/52 + 16/52 - 8/52 = 34/52

Example

You were interested in how many students at your school regularly eat breakfast. You conducted a survey, asking, "Do you eat breakfast on a regular basis?" Your random sample consisted of 600 students at the school and the results are shown in the table below.

	Male	Female	Total
Eats Breakfast Regularly	190	110	300
Doesn't Eat Breakfast Regularly	132	168	300
Total	322	278	600

If we select a student from this sample at random, what is the probability that the student is

(a) A female?

P(female) = 278/600

(b) Someone who eats breakfast regularly?

P(Breakfast) = 300/600

(c) A female and eats breakfast regularly?

P(female and breakfast) = 110/600

(d) A female or east breakfast reguarly?

P(female or breakfast)

278/600 + 300/600 - 110/600 = 468/600

1.3   Venn Diagrams and the Multiplication Rule

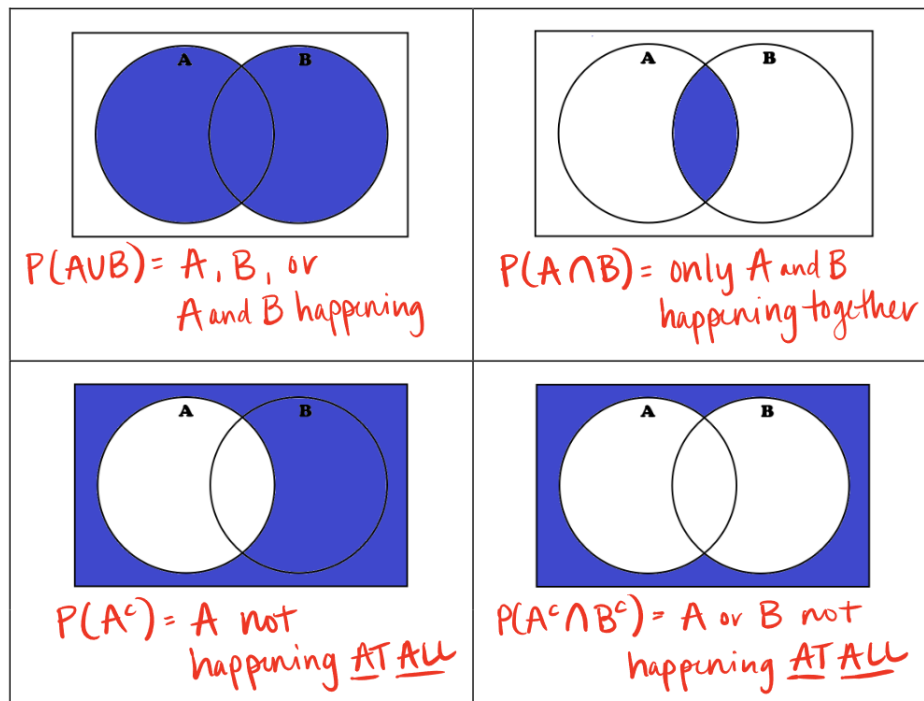Union and Intersection

- We can represent events with a Venn Diagram - a display of potential probabilities.

- The box around the Venn diagram represents the total sample space where $P(S) = 1$.

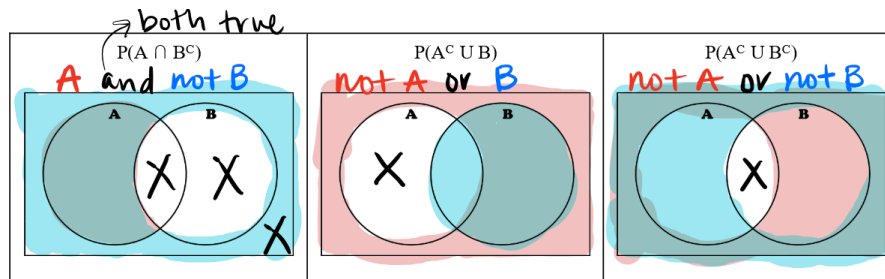- The circles themselves represent the probability of each event.
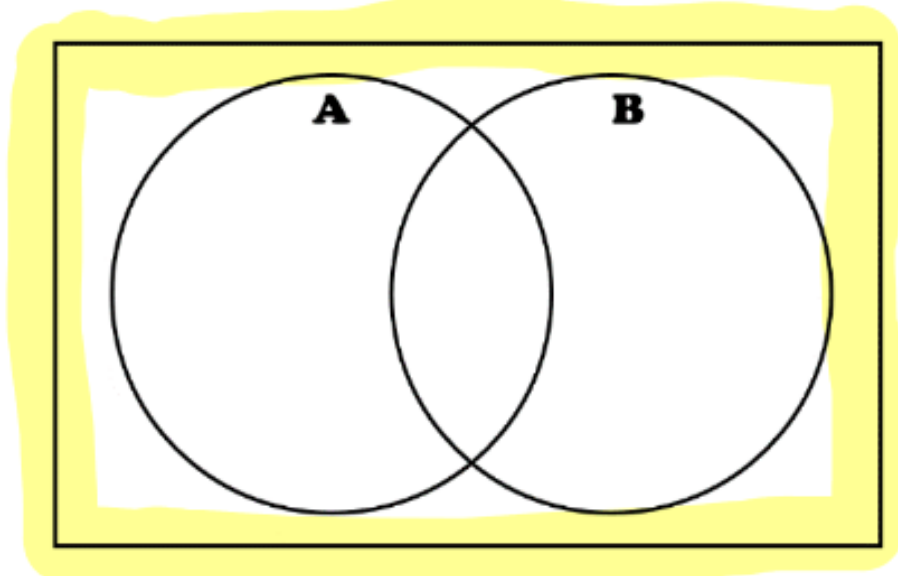
Union: OR = probability that either occurs. The symbol is $\cup$

Intersection: AND = probability both occur. The symbol is $\cap$

Example

Shade the probabilities for the following notations.

$P(A \cap B^c)$ (handwritten: both true, A and not B)

$P(A^c \cup B)$ (handwritten: not A or B)

$P(A^c \cup B^c)$ (handwritten: not A or not B)

Example

In your statistics class, you are interested in exploring the taste preference of your classmates. You have two tastes, sweet and salty. 25 students were asked if they prefer both, only sweet, only salty, or neither. The following two-way table shows the data.

	Yes Sweet	No Sweet	Total
Yes Salty	6	7	13
No Salty	11	1	12
Total	17	8	25

(Venn diagram: A, B with values 11, 6, 7, and 1 outside)
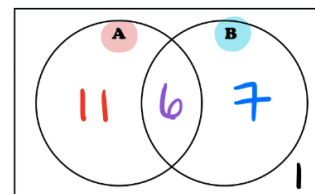
Let A = sweet and B = salty. Find and interpret the following probabilities.

(a) P(A∩B)

$6/25 = 0.24$. The probability a student likes both salty and sweet foods is 24%.

(b) P(A∪B)

$\frac{11+6+7}{25} = 0.96$. The probability a student likes either salty or sweet foods is 96%.

(c) $P(A \cap B^c)$

$\frac{11}{25} = 0.44$. The probability a student likes sweet but not salty foods is 44%.

(d) $P(A^c \cup B)$

$\frac{7+1+6}{25} = 0.56$. The probability a student likes no sweet foods or likes salty foods is 56%.

(e) $P(A \cup B^c)$

$\frac{11+6+1}{25} = 0.72$. The probability a student likes sweet foods or does not like salty foods is 72%.

(f) $P(A^c \cap B^c)$

$\frac{1}{25} = 0.04$. The probability a student does not like sweet foods and does not like salty foods is 4%.

Multiplication rule for independent events

- Two events are independent if they do not influence one another.
- The occurrence of one has no effect on the occurrence of the other.
  - Toss a coin twice.
  - The gender of each child born to the same mother.
  - Trials with replacement.

- If A and B are independent events, then the probability that both A and B occur is found using the multiplication rule.

$$P(A \text{ and } B) = P(A \cap B) = P(A) \cdot P(B)$$

Example

Find the probabilities.

(a) I flip a coin and then roll a dice. What is the probability I flip a head and then roll a 5?

$\left(\frac{1}{2}\right)\left(\frac{1}{6}\right) = 0.0833$

(b) What is the probability that a mother will have 4 girls in a row?

$\left(\frac{1}{2}\right)^4 = 0.0625.$

(c) I have a standard deck of 52 cards. What is the probability I pull a red card, replace it and shuffle, and then pull a black card?

$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 0.25$

(d) You spin the following spinner 5 times. What is the probability you land on yellow all 5 times?

$\left(\frac{3}{8}\right)^5 = 0.0074.$

Multiplication Rule for Dependent Events

- Two events are dependent if they influence one another.
- The occurrence of one affects the occurrence of the other.
    - Trials without replacement.
- If A and B are dependent events, then the probability that both A and B occur is found using:

$$P(A \text{ and } B) = P(A \cap B) = P(A) \cdot P(B|A)$$

- Where P(B|A) is the probability of B "given" that A has occurred.
- The conditional probability of an event is the probability that one event will happen, if it is known that another event has happened.

> **Example**
>
> Find the probabilities.
>
> (a) I have a standard deck of 52 cards. What is the probability of pulling a spade and then a 5?
>
> $\left(\frac{13}{52}\right)\left(\frac{4}{51}\right) = 0.0196$
>
> (b) I have a standard deck of 52 cards. What is the probability of pulling a jack and then another jack?
>
> $\left(\frac{4}{52}\right)\left(\frac{3}{51}\right) = 0.0045$
>
> (c) In your statistics class, you are interested in exploring the taste preference of your classmates. You have two tastes, sweet and salty. 25 students were asked if they prefer both, only sweet, only salty, or neither. The following two-way table shows the data.
>
> | | Yes Sweet | No Sweet | Total |
> |---|---|---|---|
> | **Yes Salty** | 6 | 7 | 13 |
> | **No Salty** | 11 | 1 | 12 |
> | **Total** | 17 | 8 | 25 |
>
> What is the probability you select someone who likes salty food then select another person who likes salty food?
>
> $\left(\frac{13}{25}\right)\left(\frac{12}{24}\right) = 0.26$
>
> What is the probability you select someone who does not like sweet food then select a person who likes sweet food?
>
> $\left(\frac{8}{25}\right)\left(\frac{17}{24}\right) = 0.2267$

1.4   Conditional Probability and Tree Diagrams

Rule for conditional probability:

- If A and B are dependent events, then the probability that both A and B occur is found using P(A∪B) = P(A)·P(B|A) where P(B|A) is the probability of B "given" that A has already occurred

- The conditional probability of an event is the probability that one event will happen, if it is known that another event has happened.

- We can rearrange the multiplication rule a little and get a formula for conditional probability:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Why does this formula make sense?

- When you know something has happened, it changes the sample space - specifically, it shrinks to the given event.

- If we know A has already occurred, that becomes your "sample space".

- Then, for B to occur, it has to occur with A, which is the intersection of A and B.
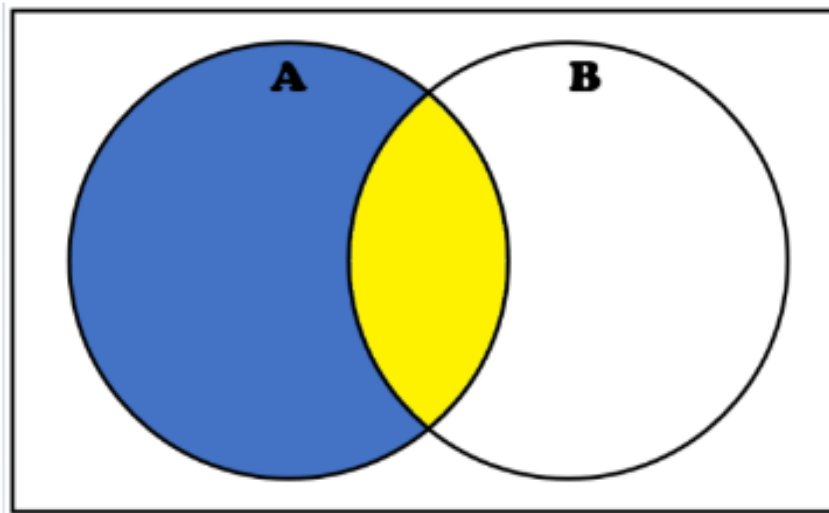
Example

Use the two-way table below to calculate the indicated probabilities.

		Eye Color				
		Brown	Blue	Hazel	Green	Total
	Black	36	9	5	2	52
Hair	Brown	66	34	29	14	143
Color	Red	16	7	7	7	37
	Blonde	4	64	5	8	81
	Total	122	114	46	31	313

(a) If we select a women and find she has brown hair, what is the probability that she will also have hazel eyes?

$$P(\text{Hazel Eyes}|\text{Brown Hair}) = \frac{P(HE \cap BH)}{P(BH)} = \frac{29/313}{143/313} = 0.2028$$

(b) If we select a woman and find she has brown eyes, what is the probability that she will also have black hair?

$$P(\text{Black Hair}|\text{Brown Eyes}) = \frac{P(BH \cap BE)}{P(BE)} = \frac{36}{122} = 0.2951$$

Showing Independence:

- If we draw two cards with replacement, we know that those events are independent.

- If we draw two cards without replacement, we know that those events are dependent.

However, in the real world, it can be difficult to determine if two events are independent or dependent. We do have a way to prove independence, however, using either of the following formulas.

$$P(A \cap B) = P(A) \cdot P(B) \text{ or } P(A) = P(A|B)$$

Example

Below is a two-way table that shows produce and how many packages are purchased based on whether it is fresh or frozen.

	Fresh	Frozen	Total
Blueberries	20	36	56
Pineapple	21	39	60
Avocado	9	15	24
Total	50	90	140

(a) Are "blueberries" and "fresh" independent?

$P(B) = P(B|\text{Fresh})$, 0.4=0.4, so independent.

(b) Are "avocados" and "frozen" independent?

$P(\text{Frozen} \cap \text{Avocados}) = P(\text{Frozen}) \cdot P(A)$, and $\frac{15}{140} \neq \frac{90}{140} \cdot \frac{24}{140}$. They are not independent.

Tree Diagrams

Example

A company manufacturing electronic components for home entertainment systems buys electrical connectors from three suppliers. The company prefers to use supplier A because only 2% of those connectors prove to be defective, but supplier A can deliver only 55% of the connectors needed. The company must also purchase connectors from two other suppliers, 40% from supplier B and the rest from supplier C. The rates of defective connectors from B and C are 4% and 5%, respectively. You buy one of these components.
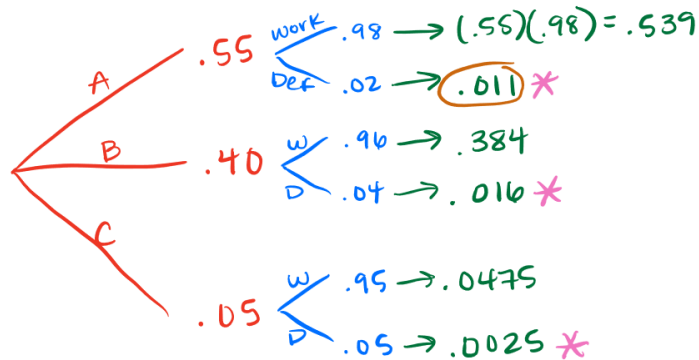
(a) What is the probability that you use supplier B to and the connector is defective?

P(B∩Defective) $= (0.40)(0.04) = 0.016$

(b) What's the probability that you find that the connector is defective?

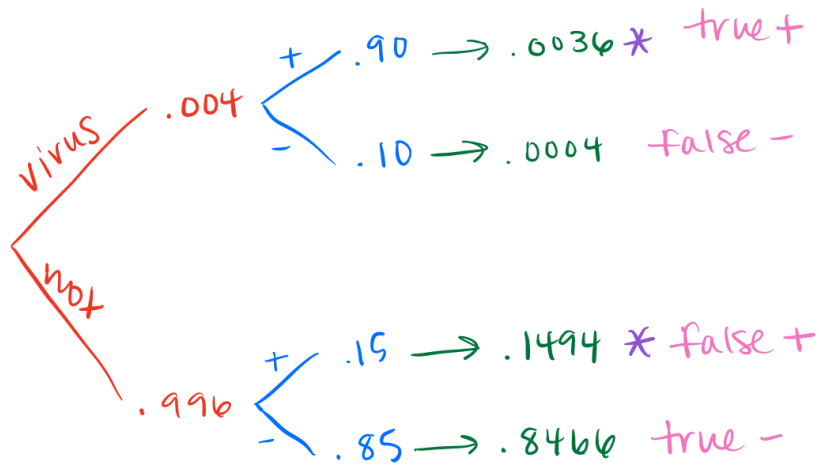P(Defective) $= 0.011 + 0.016 + 0.0025 = 0.0295$

(c) Given that the component is defective, what is the probability it came from supplier A?

P(A|Defective) $= \frac{P(A \cap \text{defective})}{P(\text{defective})} = \frac{0.011}{0.0295} = 0.3729$

> **Example**
>
> The probability that a person has a particular virus is 1 out of 250. If they have the virus, the probability they test positive for the virus is 0.90. If they do not have the virus, the probability they test positive for the virus is 0.15.
>
> virus .004
> + .90 → .0036 * true +
> − .10 → .0004 false −
>
> not .996
> + .15 → .1494 * false +
> − .85 → .8466 true −
>
> (a) Given that a person tests positive for the virus, what is the probability they have the virus?
>
> $P(\text{virus}|+) = \frac{P(\text{virus} \cap +)}{P(+)} = \frac{0.0036}{0.0036 + 0.1494} = 0.0235$
>
> (b) What is the probability of a person testing and getting a false positive?
>
> $P(\text{false}+) = P(\text{no virus} \cap +) = 0.1494$

1.5   Discrete and Continuous Random Variables

A random variable takes numerical values that describe the outcomes of some chance process. Random variables involve the same probability rules we have already learned, but we will extend those rules to be able to model more events.

Example

Suppose we toss a coin 3 times.

This is the sample space.

The total number of outcomes is 8, the probability of any one outcome is $1/8$.

Let X be the number of heads out of 3 coin tosses.

Symbols	Meaning	Probability
$X = 0$	no heads out of 3 tosses	1/8
$X = 1$	1 heads out of 3 tosses	3/8
$X = 2$	2 heads out of 3 tosses	3/8
$X = 3$	3 heads out of 3 tosses	1/8

With these 4 X values, and their corresponding probabilities, we can create a frequency table and a histogram to describe this event.
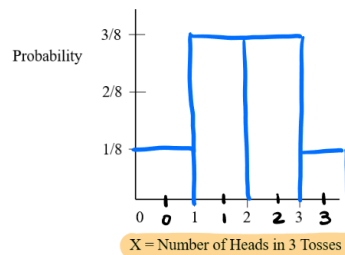
- The frequency table and histogram above each represent a probability distribution.

- A probability distributions tells us the value that our random variable can take and the probability associated with each value.

Every probability distribution must satisfy each of the following requirements.

1. There is a numerical (not categorical) random variable X, and each of the values of X has an associated probability with it.

2. The sum of all the probabilities in the distribution $\sum P(X)$ must equal 1.

3. $0 \leq P(X) \leq 1$ for all probabilities in the distribution.

Notation:

Notation	What it means	Example
$P(X = 2)$	What is the probability that the random variable equals 2?	$P(X = 2)$ $\boxed{3/8}$
$P(X \geq 2)$	What is the probability that the random variable is greater than or equal to 2?	$P(X \geq 2)$ P(x=2)+P(x=3) $\frac{3}{8} + \frac{1}{8} = \frac{4}{8} = \boxed{\frac{1}{2}}$
$P(X < 1)$	What is the probability that the random variable is less than 1?	$P(X < 1)$ P(x=0) = $\boxed{\frac{1}{8}}$

X	0	1	2	3
P(X)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

A discrete random variable takes on a fixed set of possible values with whole number outcomes.

- Random variables are usually capital letters (X, Y, Z, etc.)

- Random variables can be discrete or continuous (continuous is next lesson)

- Random variables must be numeric in value.

    - In our example, even though a result of Head or Tails is categorical, we described it numerically.

The mean of a discrete random variable X, is the mean outcome for infinitely many trials.

We think of this as the "expected value" because it is the average value we would expect to get if the trials could continue indefinitely. (Not what we expect to get in a single trial)

To find the mean, μ_x, or expected value, $\sum(x)$, of a discrete random variable X, multiply each possible value by its probability, then add all the products.

X	X_1	X_2	. . .	X_n
P(X)	P_1	P_2	. . .	P_n

$$mu_x = x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots x_n \cdot p_n = E(x) = \sum x_i \cdot p_i$$

> **Example**
>
> Northwestern University posts the grade distributions for its courses online. Students in Statistics 101 in a recent semester received 26% A's, 42% B's, 20% C's, 10% D's, and 2% F's.
>
> (a) Create a probability distribution for the random variable X, which represents the student's grade on a four point scale (with A = 4)
>
> | X | 4 | 3 | 2 | 1 | 0 |
> |------|-----|-----|-----|-----|-----|
> | P(X) | .26 | .42 | .20 | .10 | .02 |
>
> (b) Find and interpret $P(X \geq 3)$ if X represents a randomly selected Statistics 101 student.
>
> P(X$\geq$ 3) = P(X=3)+P(X=4) = 0.68. The probability that a randomly selected Stats 101 student gets a grade of A or B is 68%.
>
> (c) Write the event "the student got a grade worse than C" in terms of values of the random variable X. What is the probability of this event?
>
> $P(x < 2)$ = P(x=1) + P(x=2) = 0.10 + 0.02 = 0.12
>
> (d) Compute the mean of the random variable X. Interpret this value in context.
>
> $\mu_x = (4)(0.26) + (3)(0.42) + (2)(0.20) + (1)(0.10) + (0)(0.02) = 2.8$. If many students took this course, we would expect an average grade of 2.8, which is between a B and C.

The standard deviation of a random variable X is a measure of how much the values of the variable typically vary from the expected value.

X	x_1	x_2	...	x_n
P(X)	p_1	p_2	...	p_n

$$\sigma_x^2 = (x_1 - \mu_x)^2 \cdot p_1 + (x_2 - \mu_x)^2 \cdot p_2 + \cdots + (x_n - \mu_x)^2 \cdot p_n$$

$$\sigma_x^2 = Var(X) = \sum (x_1 - \mu_x)^2 \cdot p_i$$

$$\sigma_x = \sqrt{\sum (x_i - \mu_x)^2 \cdot p_i}$$

> **Example**
>
> For the Northwestern example, find and interpret the standard deviation of the grade probability distribution.
>
> $$\sigma_x = \sqrt{(4 - 2.8)^2(.26) + (3 - 2.8)^2(.42) + (2 - 2.8)^2(.2) + (1 - 2.8)^2(.1) + (0 - 2.8)^2(.02)} = 1$$
>
> If many students took this course, we would expect the typical deviation from the mean to be about 1 letter grade.

Continuous Random Variable

- Continuous Random Variables take on all possible values in an interval of numbers. The probability distribution of X is described by a density curve.

- The probability of any event is the area under the density curve and above, below, or between the values x that define the event.

- Area under a density curve is always equal to 1.

- The probability is on the y-axis in a distribution, so finding the area equates to finding the probability of getting an interval of numbers.

- The probability at an event is 0 for a continuous random variable because the area under a point is 0.

- The mean and standard deviation of a continuous random variable are found using integration, so we will not discuss their formulas here.

Continuous random variables are usually represented as functions in practice. The continuous random variable we will explore is the normal distribution, and we will revisit it at the start of Unit 5.

Why do we have random variables?

This is an important question to keep in mind as we continue to look at different random variables. We need random variables because they represent more things than a single measurement can. One-variable statistics used data, where random variables can use theory and apply it to real situations.

Calculator: Mean and Standard Deviation from a Probability Distribution Table

- Enter outcomes into L1 and probabilities into L2

- Run 2nd - VARS - 1VarStats with List:L1 and FreqList:L2

- Mean will be represented by $\overline{x}$

- Standard deviation will be represented by σ_x

1.6   Combining Random Variables

> **Example**
>
> Let X and Y be two independent variables with the following probability distribution.
>
> | X | 1 | 2 | 3 |
> |------|-----|-----|-----|
> | P(X) | 0.3 | 0.2 | 0.5 |
>
> | Y | 5 | 10 |
> |------|-----|------|
> | P(Y) | 0.4 | 0.6 |
>
> If we let S = X + Y, create the probability distribution below.
>
> | S | 5 | 6 | 7 | 8 | 11 | 12 | 13 |
> |------|-----|-----|----|-----|-----|-----|----|
> | P(S) | .12 | .08 | .2 | .18 | .12 | .3 | |
>
> Now we can find the following:
>
> - E(S) = 10.2
> - E(X) = 2.2
> - E(Y) = 8
> - $\sigma_S = 2.6$
> - $\sigma_X = 0.87$
> - $\sigma_Y = 2.45$

There is a shortcut for finding means and standard deviations of combined random variables without having to create a new distribution!

If X and Y are two independent random variables:

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

$$E(X - Y) = E(X) - E(Y)$$

$$\sigma_{X+Y} = \sigma_{X-Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}$$

With the problem above, if you insert the values, you can see that they are the same.

Example

The head folks at Google decided to try a random salary assignment for their workers this year! They have the following probability distribution, where X = salary.

X	50,000	60,000	70,000	80,000
P(X)	0.2	0.3	0.4	0.1

(a) What is the expected salary for a Google employee? What is the standard deviation?

$\mu_x = \$64,000$, $\sigma_x = \$9165.15$

(b) Google decides to give everyone a $500 bonus. Create a new probability distribution and find the expected value and standard deviation. Notice anything?

X+500	50500	60500	70500	80500
P(X+500)	.2	.3	.4	.1

The expected value increased by $500, but the standard deviation remained the same.

(c) Google decides that it will pay cash at the end of the year for this salary. They will give out $100 bills. Let X = number of $100 bills they will give out. Create a new probability distribution (use the bonus distribution) and find the expected value and standard deviation. Notice anything?

$\frac{X+500}{100}$	505	605	705	805
$P\left(\frac{X+500}{100}\right)$	.2	.3	.4	.1

In this case, both the mean and standard deviation were divided by 100.

If X is a random variable and a and b are both constants:

$$E(a + bX) = a + bE(X)$$

$$\sigma_{a+bX} = |b|E(x)$$

> ### Example
>
> For an upcoming concert, each customer may purchase up to 3 child tickets and 3 adult tickets. Let C be the number of child tickets purchased by a single customer. The probability distribution of the number of child tickets purchased by a single customer is given in the table below.
>
> | C | 0 | 1 | 2 | 3 |
> |---|---|---|---|---|
> | P(C) | 0.4 | 0.3 | 0.2 | 0.1 |
>
> (a) Compute the mean and the standard deviation of C.
>
> $\mu_C = 1$ child ticket, $\sigma_C = 1$ child ticket.
>
> (b) Suppose the mean and the standard deviation for the number of adult tickets purchased by a single customer are 2 and 1.2, respectively. Assume that the number of child tickets and adult tickets purchased are independent random variables. Compute the mean and the standard deviation of the total number of adult and child tickets purchased by a single customer.
>
> $\mu_A = 2$, $\sigma_A = 1.2$.
>
> $T = A + C$, $\mu_T = \mu_A + \mu_C = 3$ tickets.
>
> $\sigma_T = \sqrt{\sigma_A^2 + \sigma_C^2} = \sqrt{(1.2)^2 + (1)^2} = 1.562$ tickets.
>
> (c) Suppose each child ticket costs $15 and each adult ticket costs $25. Compute the mean and the standard deviation of the total amount per purchase.
>
> S = 25A = 15C,
>
> $\mu_S = 25\mu_A + 15\mu_C = \$65$
>
> $\sigma_S = \sqrt{(25 \cdot \sigma_A)^2 + (15 \cdot \sigma_C)^2} = \$33.54$

> ### Example
>
> Each full carton of Grade A eggs consists of 1 randomly selected empty cardboard container and 12 randomly selected eggs. The weights of such full cartons are approximately normally distributed with a mean of 840 grams and a standard deviation of 7.9 grams.
>
> The weights of the empty cardboard containers have a mean of 20 grams and a standard deviation of 1.7 grams. It is reasonable to assume independence between the weights of the empty cardboard containers and the weights of the eggs. It is also reasonable to assume independence among the weights of the 12 eggs that are randomly selected for a full carton.
>
> Let the random variable X be the weight of a single randomly selected Grade A egg.
>
> (a) What is the mean of X?
>
> 1 container + 12 eggs, full carton has $\mu = 840$, $\sigma = 7.9$, empty container has $\mu = 20$, $\sigma = 1.7$
>
> $840 = 20 + x_1 + \cdots + x_{12}$ so, $840 = 20 + 12X$, or $X = 68.33$ grams.
>
> (b) What is the standard deviation of X?
>
> $7.9 = \sqrt{(1.7)^2 + \sigma_{x1}^2 + \cdots + \sigma_{x12}^2} \implies 62.41 = (1.7)^2 + 12\sigma_x^2$.
>
> Solving for σ_x gives 2.227 grams.

1.7 The Binomial Distribution

Bernoulli Trials:

Requirements for a Bernoulli Trial:

1. Two Possible Outcomes

2. Probability of Success is the Same for Each Trial

3. Trials are Independent

If we take a Bernoulli Trial and then we are interested in the number of successes in a specific number of trials, we create a Binomial Probability Distribution.

There are four requirements for a setting to follow the binomial probability distribution. As long as these four things are true, we can use the binomial probability model for calculating the probabilities.

- Binary:
  - Each trial falls into one of two categories - we call them "success" or "failure".
  - Success does not necessarily mean a positive outcome, but instead, the outcome we are looking for.
- Independent Trials
  - Each trial is independent of the next.
  - Reasonable Assumption for coins, cards with replacement, spinners, rolling a die, etc.
  - What about sampling without replacement, when it can't be avoided? This is where the "10% condition" comes in.
    * The 10% condition says that if you are sampling from a large enough population, you can proceed with sampling without explacement as though the trials were independent.
    * Ex. Sampling from a deck of cards without replacement would violate this but sampling people from a large town would be okay.
- Number of Trials is Fixed
  - We are counting the number of successes from a set number of trials (called n).
- Same Probability
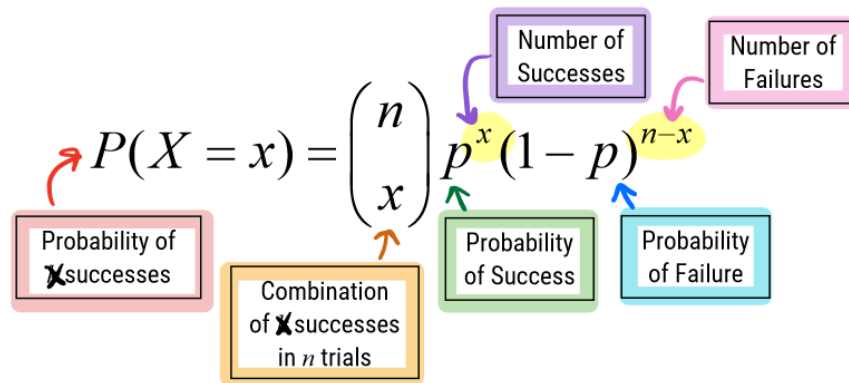 - The probability of success (called p) is the same for each trial.

If all four points are satisfied, you can compute the probability you get a certain number of successes in a specified number of trials:

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

where $\binom{n}{x} = nC_x = \frac{n!}{x!(n-x)!}$. (MATH - PROB - 3:nCr) and n is the number of independent trials, p is the probability of success, and x is the number of successes.

The mean is $\mu_x = np$, the standard deviation is $\sigma_x = \sqrt{np(1-p)}$

Why this formula?

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

Number of Successes

Number of Failures

Probability of X successes

Combination of X successes in n trials

Probability of Success

Probability of Failure

Example

Explain how each of the following situations folows the binomial probability distribution model.

(a) You toss a coin 10 times and count the number of tails you get.

X is the number of tails.

- Binary: Tails/Heads
- Independence: Coin flips are independent
- Number of Trials: 10
- Probability: 0.50

Find and interpret P(X=4).

$P(X = 4) = \binom{10}{4}(0.50)^4(0.50)^{10-4} = 0.2051$. The probability you get 4 tails in 10 coin flips is 20.51%.

(b) Type A blood is found in 43% of the population. You have a group of 40 students.

X is the number of students with Type A blood.

- Binary: Type A/Not Type A
- Independence: Assume many students in population
- Number of Trials: 40
- Probability: 0.43

Find and interpret P(X=15)

$P(X = 15) = \binom{40}{15}(0.43)^{15}(0.57)^{25} = 0.1008$. The probability you get 15 students out of 40 who have Type A blood is 10.08%.

Distributions in your calculator often have "pdf" and "cdf" after them and there is an important difference.

- pdf = probability distribution function and is used when you are trying to calculate $P(X = k)$
- cdf = cumulative distribution function and is used when you are trying to calculate $P(X \leq k)$

Specifically for binomial distributions, we want

- 2nd - VARS - A:binompdf(trials, probability of success) for $P(X = k)$
- 2nd - VARS - B:binomcdf(trials, probability of success) for $P(X \leq k)$

A binomial random variable is discrete so how you set up your inequality matters.

Example

Let $X =$ number of students that pass the AP Statistics Exam out of a class of 20. The probability that a random student will pass is 0.62. Find the probability that

(a) All students pass hte exam.

P(x=20) = binompdf(trials: 20, p: 0.62, x: 20) = 0.0007

(b) Six or fewer students pass hte exam

P(x≤ 6) = binomcdf(trials:20, p: 0.62, x:6) = 0.0037

(c) At least 12 students pass the exam

P(x≥ 12) = 1-P(x≤ 11) = 1 - binomcdf(trials:20, p:0.62, x:11) = 0.6659

Example

Suppose the probability that any random freshman girl will agree to try out for a high school dance team is 10% and one of the team captains asks 25 random freshman girls to try out.

(a) What is the probability that exactly 1 will say yes?

$P(x = 1) =$ binompdf(trials: 25, p: 0.10, x: 1) = 0.1994

(b) What is the probability that at least 1 will say yes?

$P(x \geq 1) = 1$ - $P(x = 0) = 1$ - binompdf(trials: 25, p: 0.10, x: 0) = 0.9282

(c) On average, how many freshman girls will agree to try out for the team? Interpret in context.

mean $= np = 2.5$ girls. On average, if we select many groups of 25 freshman girls, we would expect 2.5 girls to try out.

(d) What is the standard deviation of this binomial distribution? Interpret in context.

$\sigma_x = \sqrt{np(1 - p)} = 1.5$ girls. The number of girls who try out typically differs from the mean by 1.5 girls.

1.8   The Geometric Distribution

When we perform many independent trials of the same chance process and are interested in the occurrence of the first success, a geometric setting arises.

Before attempting to calculate a geometric setting, check to make sure all of the following conditions have been met.

- Binary - the possible outcomes can be classified as a success or failure

- Independent - trials must be independent (the result of one trial tells us nothing about the result of another trial)

- First - you are interested in the number of trials until the first success.

- Success - the probability of success remains consistent trial to trial

---

Example

You roll a six-sided fair die. Let $X = $ number of trials until you roll a six. Complete the probability distribution table below.

X	1	2	3	4+
$P(X)$	$\frac{1}{6}$	$\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)$	$\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)$	$1-\left(\begin{array}{c}.1667+\\.1389+\\.1157\end{array}\right)$
	.1667	.1389	.1157	.5787

If X has a geometric distribution with a probability of success p on each trial and x represents the trial that you get your first success. The probability that X equals x is given by

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

Where $P(X = x)$ is the probability of first success on x trial, $(1 - p)$ is the probability of failure, p is the probability of success, and $x - 1$ is the number of failures.

Example

You roll a six-sided fair die. Let $X = $ number of trials until you roll a six.

(a) What is the probability that you get your first six on the 5th roll?

$P(X = 5) = \left(\frac{5}{6}\right)^4 \left(\frac{1}{6}\right) = 0.0804$

(b) What is the probability that you get your first six within 3 rolls?

$P(X \le 3) = P(X = 1) + P(X = 2) + P(X = 3) = 0.1667 + 0.1389 + 0.1157 = 0.4213.$

(c) What is the probability that it takes at least 4 rolls to get your first six?

$P(X \ge 4) = 1 - P(X \le 3) = 1 - 0.4213 = 0.5787.$

Specifically for geometric distributions, we want

- 2nd - VARS - E:geometpdf(probability of success, k) for $P(X = x)$

- 2nd - VARS - F:geometcdf(probability of success, k) for $P(X \le x)$

Note: Your calculator can only do less than or equal to so all inequalities must be written in this manner.

> **Example**
>
> Let X = number of basket attempts needed for a college basketball player to make his first free throw. The player has an 82% chance of making a random free throw. Find the probability that
>
> (a) He makes his first basket on his fifth atempt.
>
> $P(X = 5)$ = geometpdf(p = 0.82, x = 5) = 0.0009
>
> (b) It takes less than 4 attempts to make his first basket.
>
> $P(X < 4)$ = geometcdf(p = 0.82, x: 3) = 0.9942
>
> (c) It takes at least 3 attempts to make his first basket.
>
> $P(X \geq 3) = 1 - P(X \leq 2)$ = 1 - geometcdf(p: 0.82, x: 2) = 0.0324
>
> (d) He will make 6 baskets in a row before he misses.
>
> $P(X = 7) = (0.82)^6(0.18)$ = geometpdf(p: 0.18, x: 7) = 0.0547

Like all probability distributions, the geometric distribution has a mean and a standard deviation. If X $G(p)$ then:

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

$$\sigma_x = \frac{\sqrt{1-p}}{p}$$

> **Example**
>
> Drew decides to place a $10 bet on Number 3 in consecutive spins of a roulette wheel until he wins. On any spin, Drew has a 1/38 chance that the ball will land in the Number 3 slot. Let X = number of spins until first win.
>
> (a) How many spins do you expect it to take until Drew wins? What does this typically vary by?
>
> $\mu_x = \frac{1}{1/38}$ = 38 spins.
>
> $\sigma_x = \frac{\sqrt{1-\frac{1}{38}}}{\frac{1}{38}}$ = 37.4967 spins
>
> Recall: if an event has a probability less than 5%, we consider it to be rare/unusual/surprising if it were to actually happen.
>
> (b) Would you be surprised if it took 10 or fewer spins for Drew to win? Calculate a probability to support your answer.
>
> $P(X \leq 10)$ = geometcdf(p: $\frac{1}{38}$, x: 10) = 0.2341, non unusual.
>
> The probability of this occurring is approximately 23.41%, so it would not be surprising.
>
> (c) Would you be surprised if it took more than 13 spins of roulette wheel until he won? Calculate a probability to support your answer.
>
> $P(X > 13) = 1 - P(X \leq 13) = 1-$geometcdf(p: $\frac{1}{38}$, x: 13) = 0.7070.
>
> The probability of this occurring is approximately 70.70%, so it would not be surprising.