Week 2: Probability Events

Jessie Yeung

STA 220

Winter 2024

Overview

- This week will cover Module 2
  - Module 2: Probability Events
- Topics include:
  - Probability and Events
  - Conditional Probability
  - Independence of Events
  - Bayes' Rule

Randomness and Probabilities

Randomness

- Why do we need to talk about probability? It all comes down to *randomness*.
- Randomness can also be introduced from random sampling or random assignments of treatments
  - Random sampling and random assignment help us make better conclusions from studies
- Randomness can come in from the error of our measurements
  - When we are collection information or measuring quantities, there will always be error attached to what we are measuring
- So randomness becomes introduced into statistics

Examples of Randomness

- The result of rolling a 6-sided die
- The time that a bus arrives at the bus stop
- The height of the next person who walks into the room
- The sex of the next baby born in a hospital
- The maternal mortality rate in Canada next year

Randomness and Probability

- Randomness is the reason why we need to talk about probability in statistics
- We never know definitively what will happen in an experiment
    - if we did, why would we do an experiment?
- But we can generally have a good idea of what outcomes of our experiment are more likely to happen
    - If we are flipping a fair coin, I cannot know for certain if I will get a head or a tail, but I do have some idea that both have the <u>*same chance*</u> of appearing.
    - If I flip a coin that has a better chance of landing on heads, then I still don't know if I will get a head or tail, but I know I am <u>*more likely*</u> to get the head
- **Randomness**: Individual outcomes are uncertain, but there is structure to how often outcomes occur in very large numbers of repetitions

What is Probability?

- Probability therefore is what we use to talk about the *chances* of some outcome of an experiment happening:
  - e.g. when flipping a fair coin, we have a 50% chance of landing on heads.
- When we then talk about the collection of chances for each possible outcome of our experiment, we get a **probability distribution**.
- **Important**: this is not the same thing as saying "*If I flip my coin twice, I have to get one head and one tail*".
  - **Probability distribution** is a model describing what I should expect to see if I flip my coin <u>forever</u>, (in a long run)

Probability Notation

- We can write $P(A)$ = probability of the event A
- For a probability to make sense, it must be between 0 and 1
  - If $P(A) = 1$, this means A will always happen
  - If $P(A) = 0$, this means A will never happen
- Just for convenience, we often use capital letters to refer to an event we are interested in
  - e.g. A = getting at least one head in two coin flips

Sample Space and Events

- A **sample space** is a set of all possible outcomes of a random experiment and is usually denoted S.
- Subsets of the sample space are considered **events**.
- Eg. for rolling a 6-sided die, $S = \{1,2,3,4,5,6\}$
  - Probability Distribution: $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$
  - Example of a Simple Event: $A = \{1\}$
  - Example of a Compound Event: $B = \{1,2,3\}$
- $0 \leq P(A) \leq 1$ for all events A
- The probability that the event will be in the sample space is 1. In other words, $P(S) = 1$

Sample Space

- We roll a 6-sided die, and then flip an unfair coin. What is our sample space?

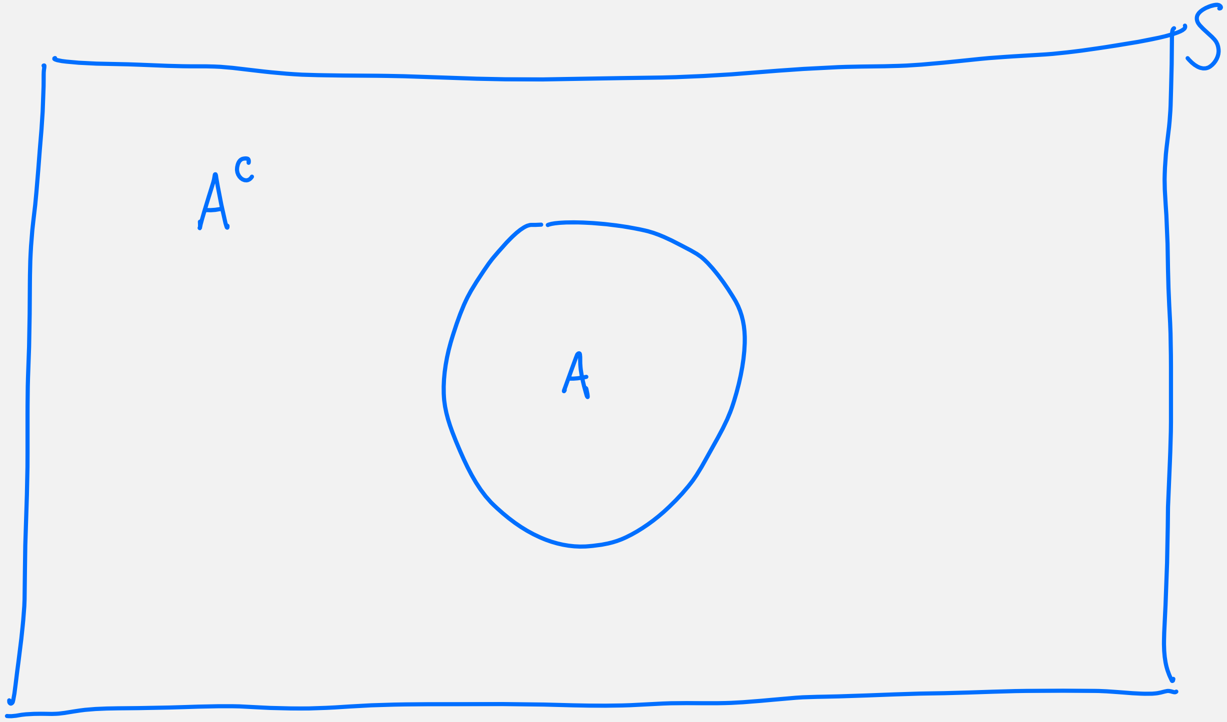
$S = \{1H, 2H, 3H, 4H, 5H, 6H, 1T, 2T, 3T, 4T, 5T, 6T\}$

The fact that the coin is unfair does not impact the sample space. It would only impact the probability distribution.

Complement of an Event

- Suppose we have an event A, but we actually want to know the probability that A does **not** happen. This is called the complement of A
- We can find this by: $P(A^c) = 1 - P(A)$
- Eg. Let A denote the event of raining tomorrow
  - $P(A)$ is the probability that it will rain tomorrow
  - $P(A^c)$ is the probability that it will not rain tomorrow
  - Further suppose that $P(A) = 0.3$. Then, $P(A^c) = 1 - 0.3 = 0.7$

Complement of an Event Visualized



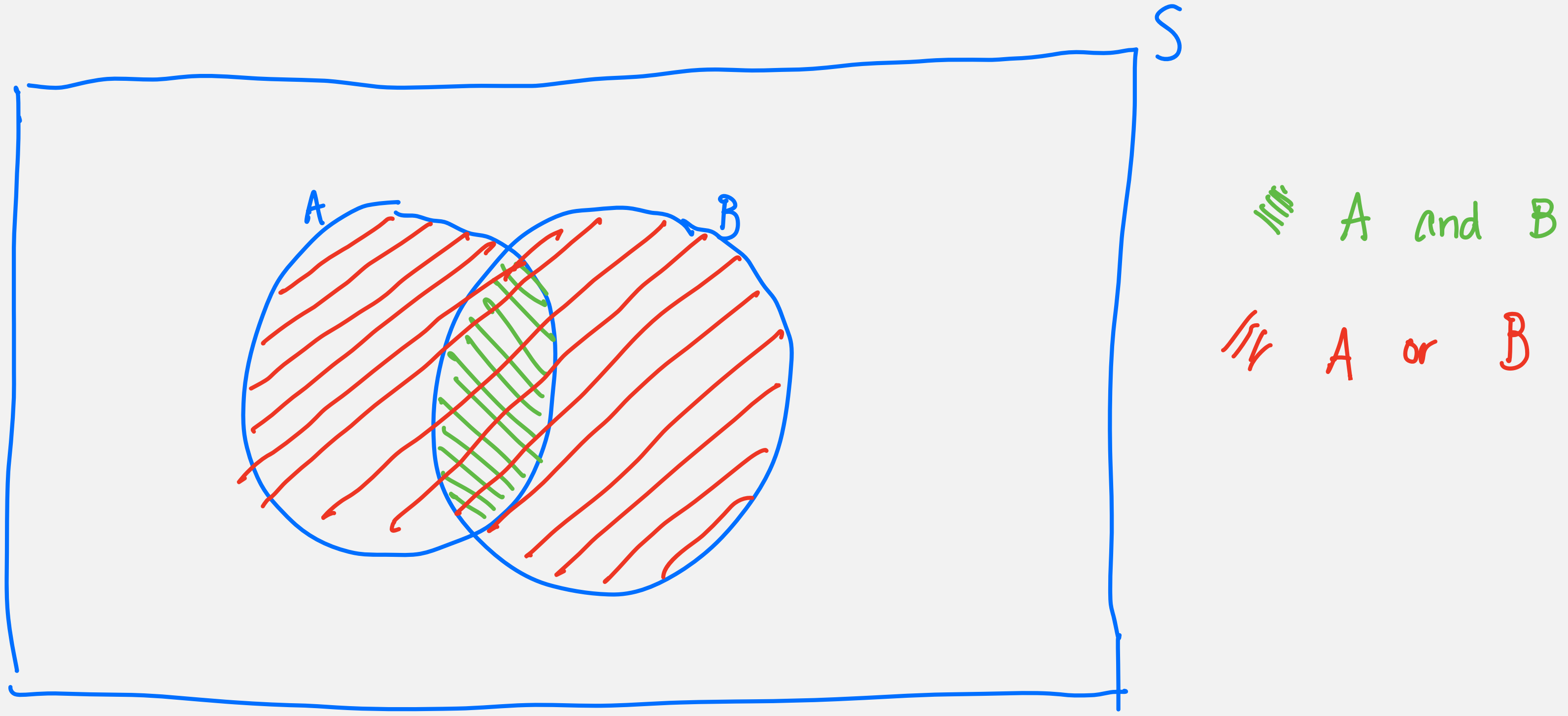
Addition Rule

- Sometimes we want to know the probability of combinations of events
- To do this, we use what we call the Addition Rule:

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

- When we use "***or***", we mean that either A can happen or B can happen or both A and B can happen
- When we use "***and***", we mean only the case where both A and B happen
- So to find the probability of A or B, we need to have 3 probabilities to work with

Addition Rule Visualized



Check your understanding!

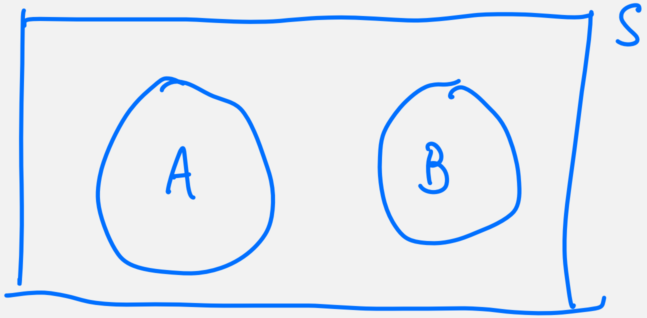
What happens when I try to use the Addition Rule and events A and B have no outcomes in common?

A. I can't compute P(A or B)

B. I can't compute P(A and B)

C. I get that P(A or B) = 0

D. I get that P(A and B) = 0



$$P(A \text{ or } B) = P(A) + P(B) - \overset{0}{\cancel{P(A \text{ and } B)}}$$

$$= P(A) + P(B)$$

Example: Rolling 2 dice

Find the following probabilities for rolling 2 dice:

a) The sum of the dice is not 6

b) The sum of the dice is at least 4

		Die #2					
		1	2	3	4	5	6
Die #1	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

$S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Sum	2	3	4	5	6	7	8	9	10	11	12
Prob	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

a) Let $A = \{2, 3, 4, 5, 7, 8, 9, 10, 11, 12\}$ $\qquad A^c = \{6\}$

Addition Rule : $P(A) = P(2 \text{ or } 3 \ldots \text{ or } 12)$

$$= \underbrace{P(2) + P(3) + \ldots + P(12)}_{\text{All disjoint}}$$

$$= 31/36$$

Complement: $P(A) = 1 - P(A^c) = 1 - P(6) = 1 - \frac{5}{36} = \frac{31}{36}$

b) Let $B = \{4, 5, 6, 7, 8, 9, 10, 11, 12\}$

$B^c = \{2, 3\}$

$P(B) = \frac{33}{36}$

Example: Rolling 2 dice

- Suppose now you are asked: What is the probability that the sum is (A) not a 6 or (B) at least a 4?
- We will use the Addition rule: $P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$
- We already have $P(A)$ and $P(B)$ from our previous example.
- Therefore I need to know what values of the sum are common to both A and B
- A=sum is not 6
- B=sum is at least 4
- So P(A and B) =

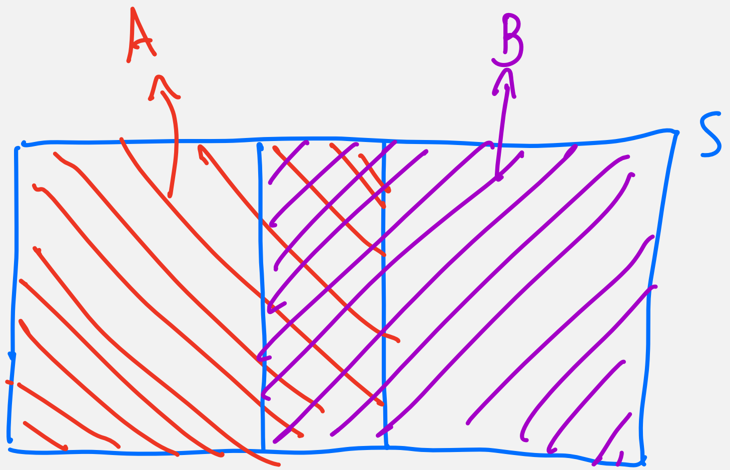
A and B = {4, 5, 7, 8, 9, 10, 11, 12}

P(A and B} = $\frac{28}{36}$

Example: Rolling 2 dice

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

$$= \frac{31}{36} + \frac{33}{36} - \frac{28}{36} = \frac{36}{36} = 1$$



Conditional Probability

Conditional Probability

- Conditional probability is very related to our discussion on contingency tables from last week (relationships between two qualitative variable)
- When we looked at the conditional distribution, we were ignoring/covering up part of our table because we were only interested in one of the groups:
  - e.g. out of only the non-coffee drinkers, what proportion drink tea?

	Coffee - No	Coffee - Yes	Total - Tea
Tea - No	44	29	73
Tea - Yes	40	45	85
Total - Coffee	84	74	158

Conditional Probability

- When we find a conditional probability, we are actually dealing with two events:
  - A = event that we are interested in (e.g. tea drinkers)
  - B = event that we already know has happened and want to condition on (e.g. non-coffee drinkers)
- Since we already know that event B happened, we can ignore anything that is not B
- We do this by conditioning on B
  - This means we remove anything in our collection of outcomes of our experiment that is not included in B.

Example: Coin Flipping

- Let's flip a coin 3 times. Then we have 8 possible outcomes:

$$\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

- Suppose we are interested in the event A = two or more heads
- What is $P(A)$?

$$P(A) = \frac{4}{8} = \frac{1}{2}$$

Example: Coin Flipping

- Let's flip a coin 3 times. Then we have 8 possible outcomes:

  {HHH, HHT, HTH, HTT, ~~THH~~, ~~THT~~, ~~TTH~~, ~~TTT~~}

- Suppose we are interested in the event A = two or more heads
- Now let's say we have event B = heads on first toss.
- Our probability of at least two heads (P(A)) will now change if we include the fact that we already know we got a head on the first toss:
  - If we condition on B, then we remove all outcomes that don't have a head on the first toss.
- What is $P(A|B)$?

$P(A|B) = \frac{3}{4}$

"given"

"conditional on"

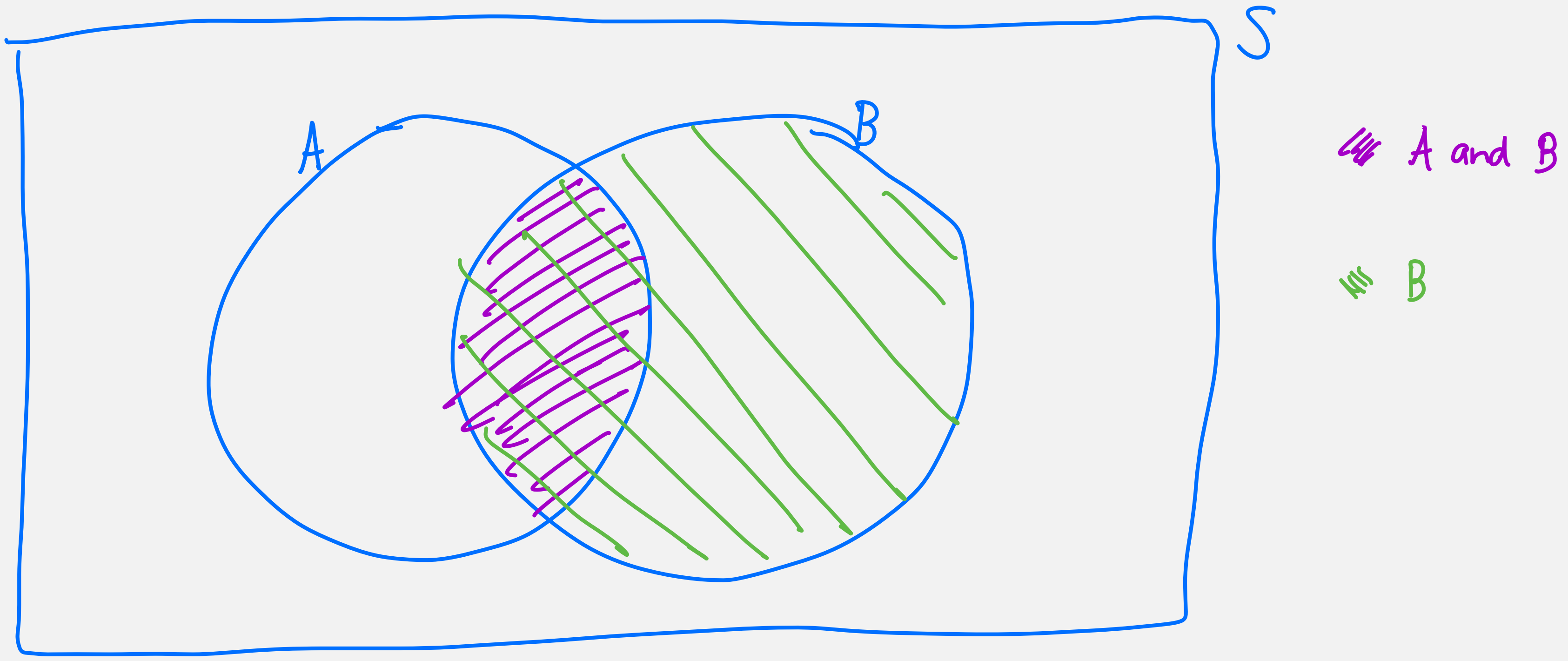
Notation – Conditional Probability

- The general definition of conditional probability has the following form:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \qquad \text{where} \quad P(B) > 0$$

- Since we are conditioning on B, this means that we are only considering how many times A happens out of the number of times B happens
  - i.e. out of only the outcomes in B, how many of them are also outcomes in A?
- So we first look at all the outcomes that are overlapping A and B (in both), and find their probability (out of everything)
- Then we look at just the ones in B, and find that probability (out of everything)

Conditional Probability Visualized



Example: Back to Coin Flipping

- So if we go back to our coin flipping example, the total possible outcomes was

$$\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

- We had event A = two or more heads, and event B = head on first toss
- If we use our formula, we need to find two probabilities:
  - P(A and B) = 3/8
  - P(B) = 1/2
- So then P(A|B) = $\frac{P(A \text{ and } B)}{P(B)} = \frac{3/8}{1/2} = 3/4$

- Note this is the same as when we first removed any outcomes that did not happen in B, and then finding how many were in A, out of what was left.

Example: Back to Coin Flipping

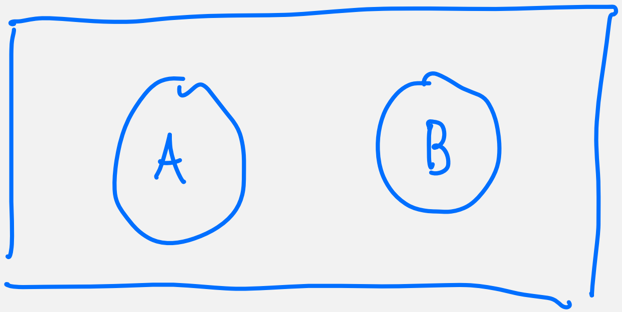
- We could also have looked at this problem with a contingency table:
- We will let the columns of the table be the proportion of possible outcomes that are either in A or not in A
- Then we let the rows of the table be the proportion of possible outcomes that are either in B or not in B.

{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

	A = 2 or more heads	not A = 1 or fewer heads	Row Total
B = head on first toss	3/8	1/8	4/8
not B = tail on first toss	1/8	3/8	4/8
Column Total	4/8	4/8	8/8 = 1

Check your understanding!

TRUE or FALSE: When A and B have no outcomes in common, $P(A|B) = P(A)$.



$P(A \text{ and } B) = 0$

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{0}{P(B)} = 0$$

False

* Would be true if $P(A) = 0$, but, the statement is false if it's describing A and B in general

Independent vs. Mutually Exclusive Events

Independence of Events

- The general idea of independence between 2 events is when knowing one event already happened, it does not influence whether or not the other event happens.
- As an example, suppose my experiment is to roll a die and flip a coin.
  - Event A = roll a 3
  - Event B = land on heads
- Does knowing that I flipped a head with my coin give me any information (that I didn't already have) about the chance that I will roll a 3? i.e. are A and B independent?
  - the key to determining independence is this idea of providing information that didn't already exist.

Independence: Definition

Mathematically we can show independence in 2 ways:

1. Conditional probability: We can say that events A and B are independent when

$$P(A|B) = P(A)$$

- This means that knowing B doesn’t give me any information about the chances of A
- Similarly, if knowing A happens doesn’t give me information about the chances of B happening, then

$$P(B|A) = P(B)$$

- This is the more intuitive concept for how we check independence
- You only need to check one of the formulas since one implies the other

Example: Roll dice/toss coin

- Again, A = {roll a 3} and B = {flip a head}
- Let's take this back to contingency tables:

{(H,1), (H,2), (H,3), (H,4), (H,5), (H,6), (T,1), (T,2), (T,3), (T,4), (T,5), (T,6)}

	A = roll a 3	not A = did not roll a 3	
B = flip a head	$1/12$	$5/12$	$6/12$
not B = flip a tail	$1/12$	$5/12$	$6/12$
	$2/12$	$10/12$	12/12

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{1/12}{6/12} = 1/6$$

$$P(A) = \frac{2}{12} = \frac{1}{6}$$

Since $P(A|B) = P(A)$, events A and B are independent.

Independence: Definition

The second way to check if events A and B are independent involves the definition of conditional probability:

2. Multiplication Rule: We can say that events A and B are independent when

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

- Recall that a conditional probability can be written as

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

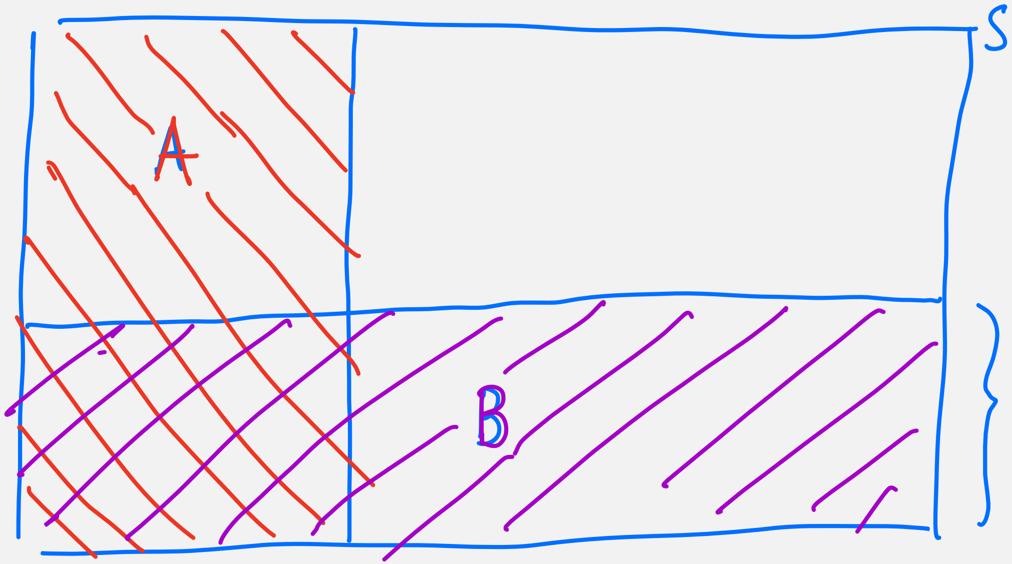
- If A and B are independent, then by the multiplication rule,

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{P(A) \times P(B)}{P(B)} = P(A)$$

- So by using the multiplication rule, we get back out first definition of independence

Independent Events Visualized

$P(B) = P(B|A) = 0.5$

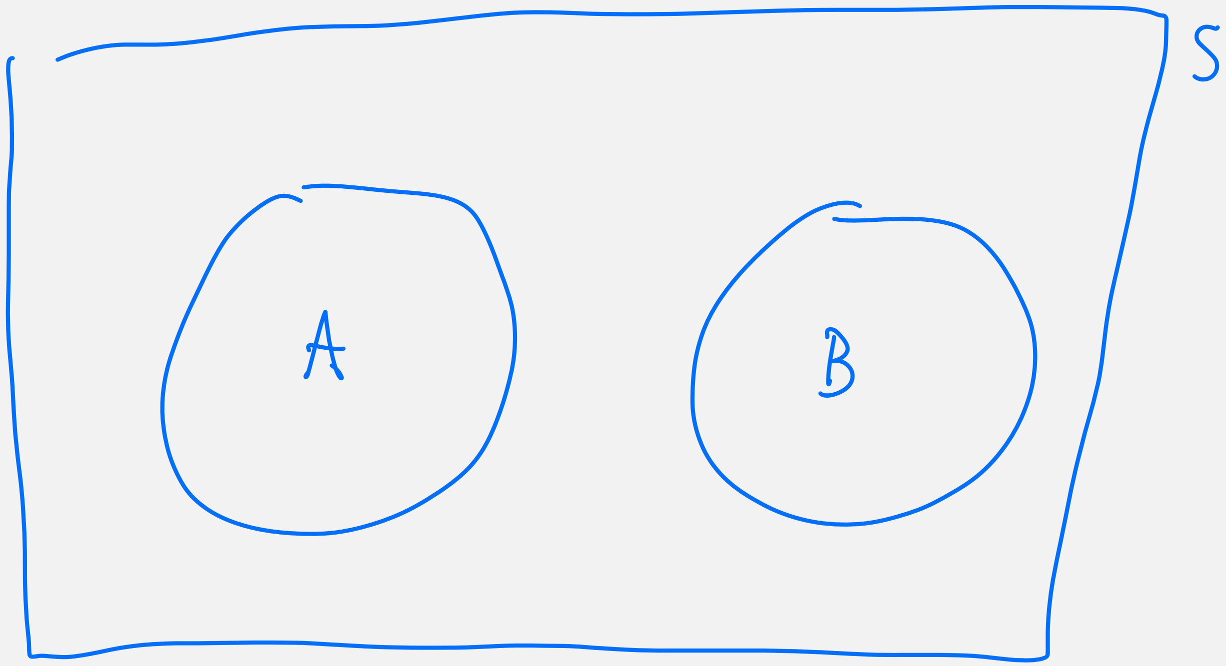


$P(B) = 0.5$

Mutually Exclusive Events

- Another concept that sometimes comes up is the idea of **mutually exclusive events** which is where two events have no outcomes in common
- This is also sometimes referred to as **disjoint events**
- The key difference between independent events and disjoint/mutually exclusive events is what is going on in the overlap between the events, i.e. P(A and B):
  - **Independent events**: there are common outcomes in the overlap, but they can be split nicely into P(A)P(B)
  - **Mutually exclusive events**: there are no common outcomes in the overlap, i.e. the overlap is empty.
    - If we have no outcome in both A and B, then both A and B cannot happen at the same time, and so P(A and B) = 0 if A and B are mutually exclusive

Mutually Exclusive/Disjoint Events Visualized



Example: Blood Types

- Canadian blood services says about 46% of Canadians have Type O blood, 42% have Type A, 9% have Type B, and the rest have Type AB.
  - If we examine one person, are the events that the person is Type A (A) and that the person is Type B (B) mutually exclusive, independent or neither?
  - i.e. Can someone be both Type A and Type B?
  - Since it is not possible to have 2 blood types, then P(A and B) = 0 so they are mutually exclusive events.

Example: Blood Types

- Canadian blood services says about 46% of Canadians have Type O blood, 42% have Type A, 9% have Type B, and the rest have Type AB.
- If we examine two people, are the events that the first person is Type A (A*) and the second person is Type B (B*) mutually exclusive, independent or neither?
  - now we are dealing with two people, so it is entirely possible that P(A* and B*) is now non-zero.
  - If we assume that these two people are not familial relatives, then these two people would be considered independent

Example: Blood Types (Continued)

Canadian blood services says about 46% of Canadians have Type O blood, 42% have Type A, 9% have Type B, and the rest have Type AB. <u>Among 4 potential donors,</u>

a) What is the probability that they are all Type O?

Let O1, O2, O3, O4, denote the event that the 1st, 2nd, 3rd, 4th person has type O.

$P(O1 \text{ and } O2 \text{ and } O3 \text{ and } O4) = P(O1) \times P(O2) \times P(O3) \times P(O4)$

$= 0.46^4$

$= 0.0448$

b) What is the probability no one is Type AB?

$P(\text{no one is Type AB}) = 0.97^4$

$= 0.8853$

Check your understanding!

$P(A) = 0.3$

$P(C^c) = 0.78 \Rightarrow P(C) = 0.22$

$P(A \text{ and } C) = 0.2$

Info 1: The probability of event A occurring is 0.3

Info 2: The probability of event C not occurring is 0.78

Info 3: The probability of event A and C occurring is 0.2

a) What is the probability that event A occurs if we know C occurred?

b) What is the probability that event C occurs if we know A occurred?

c) Are events A and C independent?

d) Ignore Info 2. If they were independent, what would be the value of P(C)?

a) $P(A|C) = \dfrac{P(A \text{ and } C)}{P(C)} = \dfrac{0.2}{0.22} = 0.9091$

b) $P(C \mid A) = \frac{P(C \text{ and } A)}{P(A)} = \frac{0.2}{0.3} = \frac{2}{3}$

c) Check if $P(A \text{ and } C) = P(A) \times P(C)$

$P(A \text{ and } C) = 0.2$

$P(A) \times P(C) = 0.3 \times 0.22$

$= 0.066$

Not independent !

d) $P(A \text{ and } C) = P(A) \times P(C)$

$\Rightarrow 0.2 = 0.3 \times P(C)$

$\Rightarrow P(C) = \frac{2}{3}$

Bayes' Rule

$P(A|B) \longrightarrow P(B|A)$

Bayes' Rule

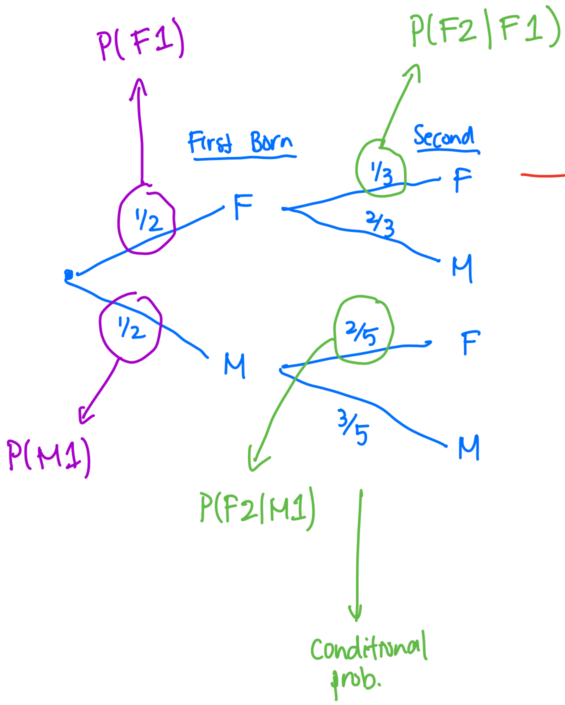
- This is a handy rule that will let us use conditional probabilities that we already know and "reverse" them to find a probability of interest
- It is actually a direct extension of the definition of conditional probability

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

- The difference is that in most conditional probability questions up to now, you have been given a value for P(B)
  - now you need to find it from other conditional probabilities

Bayes' Rule and Tree Diagrams

- To see why Bayes' rule looks like it does, it makes sense to start with tree diagrams
- Tree diagrams show conditional probabilities:
  - There is a 50% of having a female firstborn child
  - If we have a female firstborn, then the probability the second child is female is 1/3
  - If we have a male firstborn, then the probability the second child is female is 2/5



$P(F1 \text{ and } F2) = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$

Tree Diagrams

- In the child example on the previous slide, suppose we want to know the probability of a female first child given that we had a female second child, $P(F1|F2)$
  - This is the reverse of what we were given, which was $P(F2|F1)$
  - So we want:

$$P(F1|F2) = \frac{P(F1\ and\ F2)}{P(F2)}$$

  - The numerator we get from multiplying down the branch with female first and second
  - The bottom part requires a bit more thought – we have 2 pieces
    - Female second child if we already had female first, $P(F2|F1)$
    - Female second child if we had a male child first, $P(F2|M1)$

Tree Diagrams and Bayes' Rule

- The bottom part we are looking for is P(second child female)
- We saw that 2 of our tree branches end at second child being female
  - so I need to add both of these together to get the total probability of having the second child female.
- Bayes' rule just writes out these conditional probability notations:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$ Bayes'

- Notice that the numerator is equal to $P(A \text{ and } B)$

Numerator: $P(B|A) = \frac{P(B \text{ and } A)}{P(A)} \Rightarrow P(B|A) \times P(A) = P(A \text{ and } B)$

Tree Diagrams and Bayes' Rule

- Bayes' rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

P(A and B)

P(B)

- Notice that the numerator is equal to $P(A \, and \, B)$
- Notice that the denominator is equal to $P(B)$

Denominator : $P(B|A)P(A) + P(B|A^c)P(A^c)$

$= P(A \text{ and } B) + P(A^c \text{ and } B)$

$= P(B)$

Back to the Example:

- Bayes' rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

- Therefore,

$$P(F1|F2) = \frac{P(F2|F1)P(F1)}{P(F2|F1)P(F1) + P(F2|M1)P(M1)}$$

- All we did was substitute A for F1 and B for F2
- Notice that A^c becomes $M1$ since that's the complement of $F1$

$$P(F1|F2) = \frac{\frac{1}{3}\times\frac{1}{2}}{\frac{1}{3}\times\frac{1}{2}+\frac{2}{5}\times\frac{1}{2}} = \frac{5}{11}$$

Example: Thrombosis

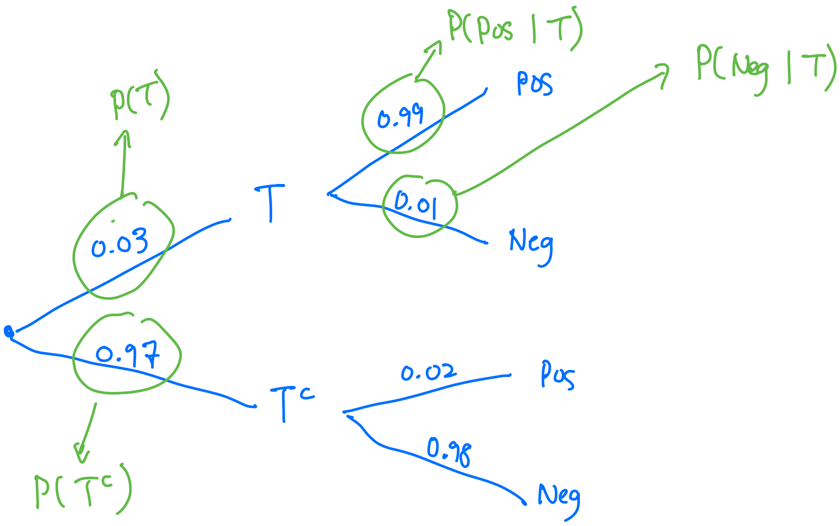
A genetic test is used to determine if people have a predisposition for thrombosis (i.e. blood clotting that blocks blood flow). It is believed that 3% of people actually have this predisposition. The genetic test has probability 99% of giving positive result when person actually has it. It also has probability 98% of giving a negative result when the person does not have the predisposition. What is the probability that someone who tests positive for the predisposition actually has it?

T = has Thrombosis predisposition

Pos = $\oplus$ result

Neg = $\ominus$ result

$P(T) = 0.03$ $\quad$ $P(Pos \mid T) = 0.99$ $\quad$ $P(Neg \mid T^c) = 0.98$



Want $P(T \mid Pos)$

$$P(T \mid Pos) = \frac{P(Pos \mid T)\, P(T)}{P(Pos \mid T)\, P(T) + P(Pos \mid T^c)\, P(T^c)}$$

$$= \frac{0.99 \times 0.03}{0.99 \times 0.03 + 0.02 \times 0.97}$$

$$= 0.605$$

Reminder : In the denominator, we are calculating $P(Pos)$.

In this case, $P(Pos) = 0.0491$

Bayes' Rule

- Once you have created a tree diagram, you should know how to use Bayes rule to calculate any probability related to events A and B.

Videos and Practice Problems

- In this lecture, we covered all of Module 2: Probability Events
- Try practice problems posted for Module 2

Next Week

- Second quiz is due on Sunday at 11:59 PM
- Next week we will continue our discussion on probability with the introduction of random variables