Week Four: Basic Probability II

CSC 217

Combinations

- Last week we found all the results that could occur if we flip a coin three times.

HHH	HHT
HTH	THH
THT	HTT
TTH	TTT

Combinations

- Last week we found all the results that could occur if we flip a coin three times.
- What are the odds of getting at least two heads in three coin flips?

HHH	HHT
HTH	THH
THT	HTT
TTH	TTT

Combinations

- Last week we found all the results that could occur if we flip a coin three times.
- What are the odds of getting at least two heads in three coin flips?
- ½, or 4 events in a sample space of 8

HHH	**HHT**
HTH	**THH**
THT	HTT
TTH	TTT

Combinations

- Last week we found all the results that could occur if we flip a coin three times.
- What are the odds of getting exactly three heads in three flips?

HHH	HHT
HTH	THH
THT	HTT
TTH	TTT

Combinations

- Last week we found all the results that could occur if we flip a coin three times.
- What are the odds of getting exactly three heads in three flips?
- ⅛, or 1 event in a sample space of 8

HHH	HHT
HTH	THH
THT	HTT
TTH	TTT

Combinations

- Last week we found all the results that could occur if we flip a coin three times.
- What are the odds of getting exactly three heads in three flips given that you are going to get at least two heads?

HHH	HHT
HTH	THH
THT	HTT
TTH	TTT

Combinations

- Last week we found all the results that could occur if we flip a coin three times.
- What are the odds of getting exactly three heads in three flips given that you are going to get at least two heads?
- ¼ - even though the event is the same, the sample space is reduced to four events instead of the original eight

HHH	HHT
HTH	THH

Conditional Probability

- **Conditional Probability** refers to the probability of something happening given that something else happens

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

Combinations

- Last week we found all the results that could occur if we flip a coin three times.
- What are the odds of getting exactly three heads in three flips given that you are going to get at least two heads?
- ¼ - even though the event is the same, the sample space is reduced to four events instead of the original eight
- Also equivalent to P(A & B), or ⅛ / P(B), or ½ = 1/4

HHH	HHT
HTH	THH

Conditional Probability

- Say my office is having a 'father-daughter' dinner for employees who have at least one daughter. If I'm invited to to the dinner, what is the probability that both of my two children are girls?
- In this case, what do each of the probabilities in the equation below represent?

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

Conditional Probability

- Say my office is having a ‘father-daughter’ dinner for employees who have at least one daughter. If I’m invited to to the dinner, what is the probability that both of my two children are girls?
- In this case, what do each of the probabilities in the equation below represent?
- P(B) - the probability I’m invited to the dinner
- P(A&B) - the probability that both of my two children are girls
- P(A&B) | P(B) - the probability that both of my two children are girls given that I’m invited to the dinner

Conditional Probability

- Say my office is having a ‘father-daughter’ dinner for employees who have at least one daughter. If I’m invited to to the dinner, what is the probability that both of my two children are girls?

Scenarios	**Probability**
Two boys	
One girl, one boy	
Two girls	

Conditional Probability

- Say my office is having a ‘father-daughter’ dinner for employees who have at least one daughter. If I’m invited to to the dinner, what is the probability that both of my two children are girls?

Scenarios	Probability
Two boys	1/4
One girl, one boy	1/2
Two girls	1/4

Conditional Probability

- P(B) - the probability I'm invited to the dinner

Scenarios	**Probability**
Two boys	1/4
One girl, one boy	**1/2**
Two girls	**1/4**

Conditional Probability

- P(A&B) - the probability that both of my two children are girls

Scenarios	**Probability**
Two boys	1/4
One girl, one boy	**1/2**
Two girls	**1/4**

Conditional Probability

- P(B) - the probability I'm invited to the dinner = 3/4
- P(A&B) - the probability that both of my two children are girls = ¼
- P(A | B) - the probability that both of my children are girls given that I'm invited to the dinner: ¼ / ¾ = 1/3

Multiplication Rule

- The multiplication rule is simply a variation of the conditional probability equation

$$P(A \bigcup B) = P(A|B) * P(B)$$

$$P(A \bigcup B) = P(B|A) * P(A)$$

Law of Total Probability

- Say the probability of B has a sample space of two, B1 and B2 - there are only two possibilities for how B will occur
- Given that an experiment for B happens, the total probability for A is the sum of the probability of A occuring and the sum of each of the probabilities of B occuring.

$$P(A) = P(A) \bigcup P(B_1) + P(A) \bigcup P(B_2)$$

Independence

- Two events are independent if and only if knowing that one event occured doesn’t change the probability that another event occured
- In order to be independent, the two events must satisfy the equation below
- If two events aren’t independent, they are **dependent**

$$P(A \bigcup B) = P(A) * P(B)$$

Independence

- What is the probability of getting heads on the first flip and the second flip? Does the probability of getting heads on the first flip affect the probability of getting heads on the second flip?
- What is the probability of getting 2 on your first dice roll and 5 on the sum of your dice rolls? Does the probability of the former affect the probability of the latter?
- The notion of independence is one of the most abused tenets of statistics. Are you *sure* that the two events are independent??

$$P(A \bigcup B) = P(A) * P(B)$$

Independence


HuffPost
@HuffPost
Follow

Our @pollsterpolls model gives @HillaryClinton a 98.1% chance of winning the presidency elections.huffingtonpost.com/2016/forecast/…

RETWEETS	LIKES
11,660	9,592

11:25 AM - 7 Nov 2016

2.6K 12K 9.6K


Nate Cohn
@Nate_Cohn
Follow

Sleep well.

RETWEETS	LIKES
2,868	4,139

11:14 PM - 23 Oct 2016 from Manhattan, NY

367 2.9K 4.1K

Independence

- Toss a coin three times. A is the event ‘heads on the first toss’ and B is the event ‘two heads total’. Are A and B independent?
- Find the probability of event A and event B and see if P(A&B) = P(A) * P(B)

Independence

- Toss a coin three times. A is the event ‘heads on the first toss’ and B is the event ‘two heads total’. Are A and B independent?
- P(A) = ½ -[HHH, HHT, HTH, HTT]
- P(B) = ⅜ - [HHT, HTH, THH]
- P(A & B) = ¼ - [HHT, HTH]
- ½ * ⅜ = 3/16 != 1/4

Bayes' Theorem

- Bayes' Theorem is a simple algebraic transformation of the conditional probability equation using the multiplication rule
- It is notable because it allows you to find the conditional probability of one variable given the conditional probability of the other variable

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

Bayes' Theorem

Conditional Probability: $$P(A|B) = \frac{P(A \bigcup B)}{P(B)}$$

Multiplication Rule: $$P(A \bigcup B) = P(B|A) * P(A)$$

Bayes' Theorem: $$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

Bayes Theorem

- Say I want to test a new software that allows me to detect if a student has cheated on a test
- 0.0001% (or 1 in 10,000) of students cheat on tests (what angels!)
- When a student cheats on a test, the software will pick up on it 99% of the time
- When a student doesn't cheat on a test, the software will say they've cheated 1% of the time
- If the software tells me a student has cheated on a test, what is the probability that they've actually cheated?

Bayes Theorem

- Let’s draw this out on a table with 1,000,000 simulated students

	Cheated	Not Cheated	**Total**
Alert			
No Alert			
Total			1,000,000

Bayes Theorem

- 1 in 10,000 students cheat

	Cheated	Not Cheated	**Total**
Alert			
No Alert			
Total	100	999,900	1,000,000

Bayes Theorem

- When a student cheats on a test, the software will pick up on it 99% of the time

	Cheated	Not Cheated	**Total**
Alert	99		
No Alert	1		
Total	100	999,900	1,000,000

Bayes Theorem

- When a student doesn't cheat on a test, the software will say they've cheated 1% of the time

	Cheated	Not Cheated	**Total**
Alert	99	9,999	
No Alert	1	989,901	
Total	100	999,900	1,000,000

Bayes Theorem

- You can then fill out the rest of the table accordingly

	Cheated	Not Cheated	**Total**
Alert	99	9,999	10,098
No Alert	1	989,901	989,902
Total	100	999,900	1,000,000

Bayes Theorem

- If the software tells me a student has cheated on a test, what is the probability that they've actually cheated?
- P(Cheated | Alert) = P(Cheated & Alert) / P(Alert) = 99/10,098 = 0.0098 or less than 1%
- The software will identify a cheater 1% of the time!!

	Cheated	Not Cheated	**Total**
Alert	99	9,999	10,098
No Alert	1	989,901	989,902
Total	100	999,900	1,000,000

Bayes Theorem

- If the software tells me a student has cheated on a test, what is the probability that they've actually cheated?
- P(Cheated) = 0.0001
- P(Not Cheated) = 0.9999
- P(Alert | Cheated) = 0.99
- P(Alert | Not Cheated) = 0.01
- P(Cheated | Alert) = ?

Bayes Theorem

- If the software tells me a student has cheated on a test, what is the probability that they've actually cheated?
- P(Cheated) = 0.0001
- P(Not Cheated) = 0.9999
- P(Alert | Cheated) = 0.99
- P(Alert | Not Cheated) = 0.01
- P(Cheated | Alert) = (P(Alert | Cheated) * P(Cheated)) / P(Alert)

Bayes Theorem

- If the software tells me a student has cheated on a test, what is the probability that they've actually cheated?
- P(Cheated) = 0.0001
- P(Not Cheated) = 0.9999
- P(Alert | Cheated) = 0.99
- P(Alert | Not Cheated) = 0.01
- P(Cheated | Alert) = (P(Alert | Cheated) * P(Cheated)) / P(Alert)
- P(Cheated | Alert) = (0.99 * 0.0001) / P(Alert)
- **Law of Total Probability**: P(Alert) = (P(Alert) & P(Cheated)) * (P(Alert) & P(Not Cheated)

Bayes Theorem

- **Law of Total Probability**: P(Alert) = (P(Alert) & P(Cheated)) * (P(Alert) & P(Not Cheated)
- **Multiplication Rule**: P(Alert) & P(Cheated) = P(Alert | Cheated) * P(Cheated)
- P(Alert | Cheated) = 0.99
- P(Cheated) = 0.0001
- P(Alert) & P(Cheated) = 0.99 * 0.0001 = 0.000099
- **Multiplication Rule**: P(Alert) & P(Not Cheated) = P(Alert | Not Cheated) * P(Not Cheated)
- P(Alert | Not Cheated) = 0.01
- P(Not Cheated) = 0.9999
- P(Alert) * P(Not Cheated) = 0.01 * 0.999 = 0.009999
- P(Alert) = 0.000099 + 0.009999 = 0.010098

Bayes Theorem

- If the software tells me a student has cheated on a test, what is the probability that they've actually cheated?
- P(Cheated) = 0.0001
- P(Not Cheated) = 0.9999
- P(Alert | Cheated) = 0.99
- P(Alert | Not Cheated) = 0.01
- P(Cheated | Alert) = (P(Alert | Cheated) * P(Cheated)) / P(Alert)
- P(Cheated | Alert) = (0.99 * 0.0001) / P(Alert)
- P(Alert) = 0.010098
- **P(Cheated | Alert) = (0.99 * 0.0001) / 0.010098 = 0.0098**
- This confirms the answer we found earlier via creating a table

Bayes Theorem

- If the software tells me a student has cheated on a test, what is the probability that they've actually cheated?
- True Positive - Cheated & Alert
- False Positive - Type I Error - Not Cheated & Alert
- True Negative - Not Cheated & No Alert
- False Negative -Type II Error - Cheated & No Alert

	Cheated	Not Cheated	**Total**
Alert	99	9,999	10,098
No Alert	1	989,901	989,902
Total	100	999,900	1,000,000

Bayes Theorem

- If the software tells me a student has cheated on a test, what is the probability that they've actually cheated?
- Assessing this tradeoff is a huge part of **machine learning** for classification
- In this scenario, there's no way this software would be feasible given the high false positive rate

	Cheated	Not Cheated	**Total**
Alert	99	9,999	10,098
No Alert	1	989,901	989,902
Total	100	999,900	1,000,000

Bayes Theorem

- Accuracy: (989901 + 99) / 1000000 = 0.99
- Precision: 99 / (99 + 9,999) = 0.0098
- Recall: 99 / 100 = 0.99

	Cheated	Not Cheated	**Total**
Alert	99	9,999	10,098
No Alert	1	989,901	989,902
Total	100	999,900	1,000,000

Bayes Theorem

- Accuracy: (989901 + 99) / 1000000 = 0.99
- Precision: 99 / (99 + 9,999) = 0.0098
- Recall: 99 / 100 = 0.99

	Cancer	Not Cancer	**Total**
Alert	99	9,999	10,098
No Alert	1	989,901	989,902
Total	100	999,900	1,000,000

Bayes Theorem

- Accuracy: (989901 + 99) / 1000000 = 0.99
- Precision: 99 / (99 + 9,999) = 0.0098
- Recall: 99 / 100 = 0.99

	Criminal	Not Criminal	**Total**
Arrest	99	9,999	10,098
Don’t Arrest	1	989,901	989,902
Total	100	999,900	1,000,000

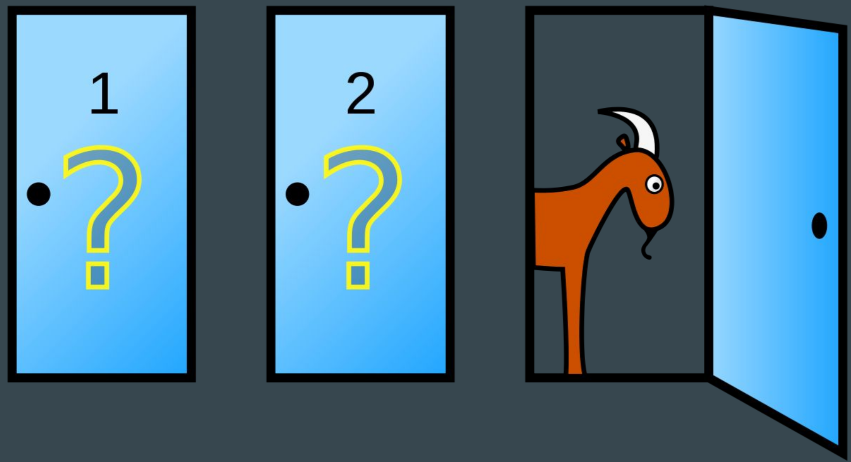
Bayes Theorem

- Accuracy: (989901 + 99) / 1000000 = 0.99
- Precision: 99 / (99 + 9,999) = 0.0098
- Recall: 99 / 100 = 0.99

	'Wire Transfer'	-	**Total**
Junk Mail	99	9,999	10,098
Not Junk Mail	1	989,901	989,902
Total	100	999,900	1,000,000

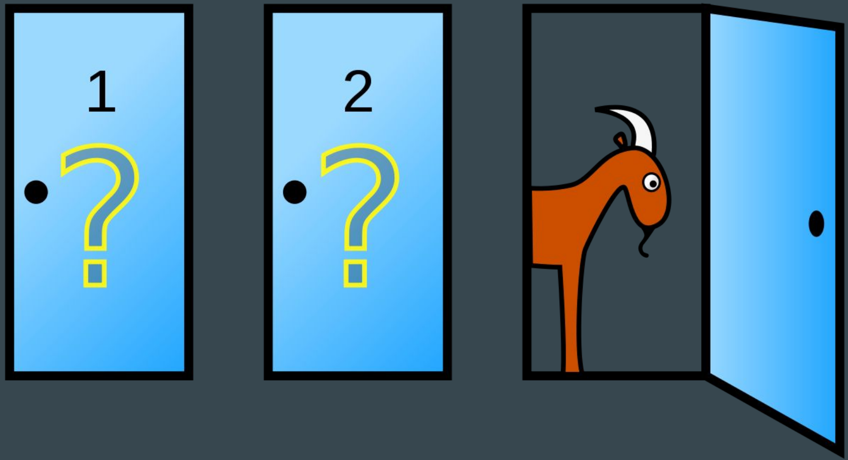
Monty Hall Problem

- Say we find ourselves on a game show.
- There are three doors available - one has a car behind it and the other two have goats behind it.
- Once we pick a door, the host - Monty Hall - opens another door to reveal a goat behind it and gives us the offer to switch doors to the final available door.
- **Assume we pick Door 1.** Should we switch doors? Does it make a difference?



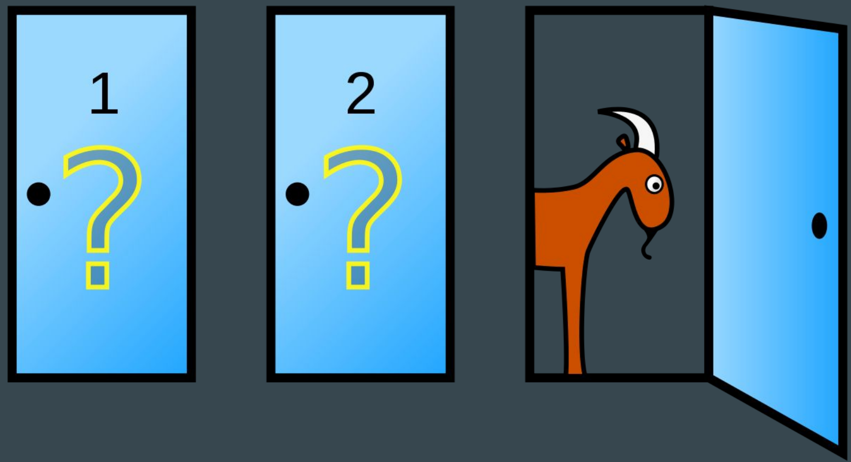
Monty Hall Problem

- **Assume we pick Door 1.** Should we switch doors? Does it make a difference?
- Let's start simple. With three doors and two outcomes, how many possible scenarios are there? (Order does not matter)



Monty Hall Problem

- **Assume we pick Door 1.** Should we switch doors? Does it make a difference?
- Let's start simple. With three doors and two outcomes, how many possible scenarios are there? (Order does not matter)
- This is a **combination** problem
  - 3! / (2!)(1!) = 3. There are **three** scenarios.



$$C(n,r) = \frac{n!}{(n-r)!r!}$$

Monty Hall Problem

	Door 1	Door 2	Door 3
Scenario #1	Car	Donkey	Donkey
Scenario #2	Donkey	Car	Donkey
Scenario #3	Donkey	Donkey	Car

Monty Hall Problem

	Door 1	Door 2	Door 3
Scenario #1	Car	Donkey	Donkey
Scenario #2	Donkey	Car	Donkey
Scenario #3	Donkey	Donkey	Car
Probability	1/3	1/3	1/3

Monty Hall Problem

	Opens Door 2	Opens Door 3	Total
Door 1			1/3
Door 2			1/3
Door 3			1/3
Total			1

Monty Hall Problem

	Opens Door 2	Opens Door 3	Total
Door 1			1/3
Door 2			1/3
Door 3			1/3
Total	1/2	1/2	1

Monty Hall Problem

	Opens Door 2	Opens Door 3	Total
Door 1			1/3
Door 2	0		1/3
Door 3		0	1/3
Total	1/2	1/2	1

Monty Hall Problem

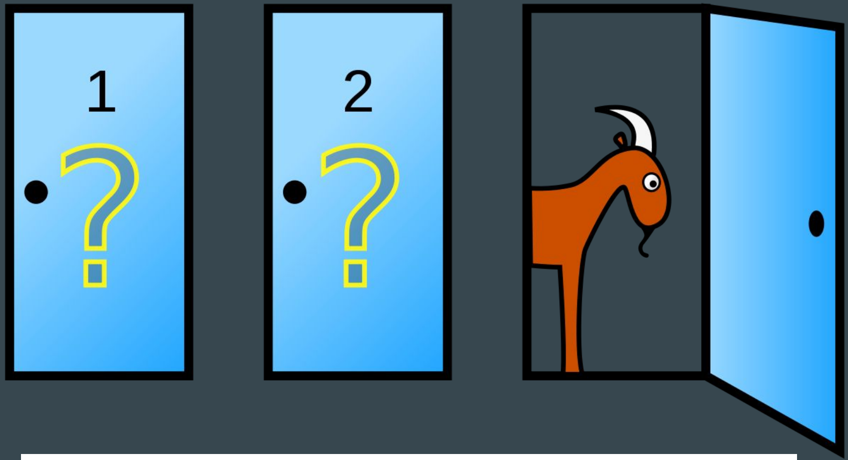
	Opens Door 2	Opens Door 3	Total
Door 1			1/3
Door 2	0	1/3	1/3
Door 3	1/3	0	1/3
Total	1/2	1/2	1

Monty Hall Problem

	Opens Door 2	Opens Door 3	Total
Door 1	1/6	1/6	1/3
Door 2	0	1/3	1/3
Door 3	1/3	0	1/3
Total	1/2	1/2	1

Monty Hall Problem

- **Assume we pick Door 1.** Should we switch doors? Does it make a difference?
- **Assume Monty picks Door 3.**
- The probability of Door 1 having the car and Monty pulling Door 3 is ⅙
- The probability of Monty pulling Door 3 is ½
- Thus the probability of Door 1 having the car given that Monty pulls Door 3 is ⅓!!
- **Thus you should switch.**



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