Probability and Statistics

UG2, Core course, IIIT,H

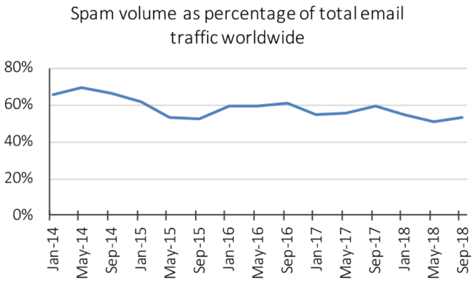
Pawan Kumar

IIIT, Hyderabad

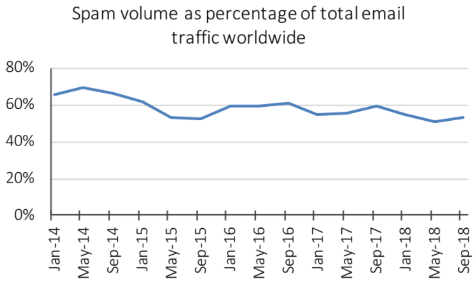
September 2, 2021

1 Outline

1 Bayes Theorem. Why?



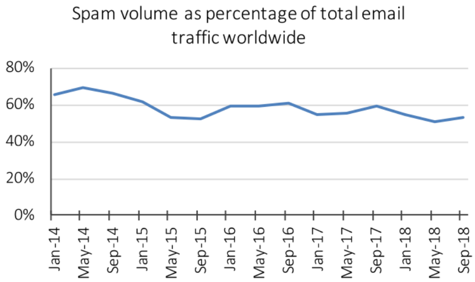
1 Bayes Theorem. Why?



- We can easily calculate how many spam emails contain "Dear":

$$P(E|F) = P(\text{Dear}|\text{Spam}) \; ?$$

1 Bayes Theorem. Why?



- We can easily calculate how many spam emails contain "Dear":

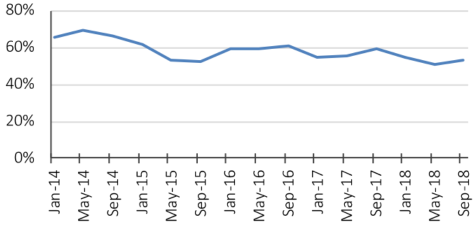
$$P(E|F) = P(\text{Dear}|\text{Spam})$$

- But what is the probability that an email containing "Dear" is spam?

$$P(F|E) = P(\text{Spam email}|\text{Dear})$$

1 Bayes Theorem. Why?

Spam volume as percentage of total email traffic worldwide



- We can easily calculate how many spam emails contain "Dear":

$$P(E|F) = P(\text{Dear}|\text{Spam})$$

- But what is the probability that an email containing "Dear" is spam?

$$P(F|E) = P(\text{Spam email}|\text{Dear})$$

1 Bayes Theorem

Bayes Theorem

For any events E and F where $P(E) > 0$ and $P(F) > 0$,

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

Proof of Bayes Theorem:

$$P(F|E) = \frac{P(F \cap E)}{P(E)}$$ (Using defn of conditional probab.)

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

$$\Rightarrow P(F|E)P(E) = P(E|F)P(F) \Rightarrow P(F|E) = \frac{P(E|F)P(F)}{P(E)}.$$

1 Bayes Theorem with Total Probability...

Bayes Theorem

For any events E and F where $P(E) > 0$ and $P(F) > 0$,

$$P(F|E) = \frac{P(E|F)P(F)}{\underline{P(E|F)P(F) + P(E|F^c)P(F^c)}}$$

$\| \; P(E) \leftarrow$ (Using total probab.)

Proof of Bayes Theorem:

1 Bayes Theorem Used in Spam Emails Example...

1 Bayes Theorem Used in Spam Emails Example...

Spam Email Example

Given the following:

- 60% of all email in 2016 is spam

Spam Email Example

Given the following:

- 60% of all email in 2016 is spam
- 20% of spam has the word “Dear”

Spam Email Example

Given the following:

- 60% of all email in 2016 is spam
- 20% of spam has the word "Dear"
- 1% of non-spam (aka ham) has the word "Dear"

E : Evidence, also use H
F : Fact

Spam Email Example

Given the following:

- 60% of all email in 2016 is spam
- 20% of spam has the word "Dear"
- 1% of non-spam (aka ham) has the word "Dear"

You get an email with the word "Dear" in it. What is the probability that the email is spam?

. Assume : 30% of all emails have the word "Dear"

$\frac{1}{100} \times \frac{4}{10} = \frac{4}{1000}$ 0.004

$P(F^c) = 0.4$

Solution: Define suitable events :

E : "Dear"

F : "Spam"

Recall Baye's thm

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

Note : E is the event of having word "Dear" in all emails, this is not given!

$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$
$= 0.2 \times 0.6 + 0.01 \times 0.4 = 0.12 + 0.004 = 0.124$

$P(E|F) = 20\% = 0.2$

$P(F) = 60\% = 0.6$

$P(E) = ? = 30\% = 0.3$

$P(F|E) = \frac{0.2 \times 0.6}{0.3} = \frac{0.12}{0.3} = 0.4$

0.01

Example

A test is 98% effective at detecting a disease ("true positive"). However, the test has a "false positive" rate of 1%. The 0.5% of the US population has disease. What is the likelihood you have the disease, if you test positive?

Solution: Define events

E : you test positive

F : you actually have the disease

Want to find : $P(F|E)$

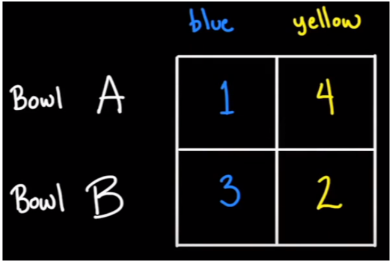
Confusion Matrix

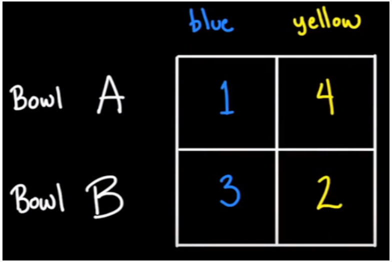
	F	F^c
E	$P(E\mid F)$ 98%	$P(E\mid F^c)$ 1%
E^c	$P(E^c\mid F)$ 2%	$P(E^c\mid F^c)$ 99%

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)} = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^c)P(F^c)}$$

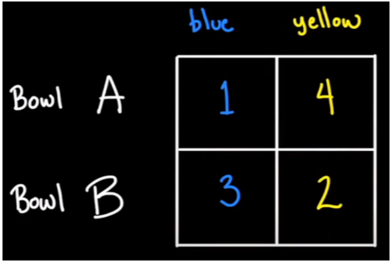
$$= \frac{0.98 \times 0.005}{0.98 \times 0.005 + 0.01 \times 0.995}$$

1 Conditional Probability...

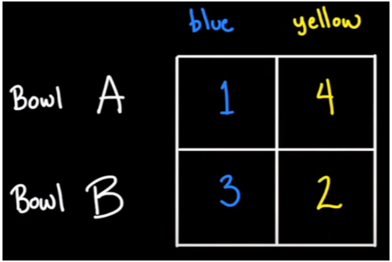




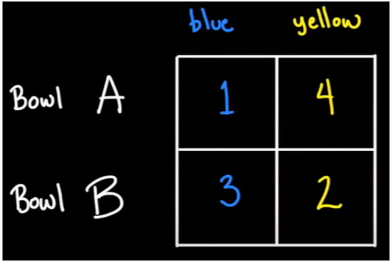
- consider two bowls A and B



- consider two bowls A and B
- bowl A contains 1 blue and 4 yellow marbles

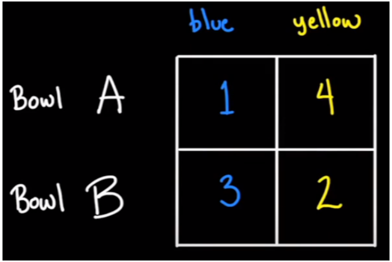


- consider two bowls A and B
- bowl A contains 1 blue and 4 yellow marbles
- bowl B contains 3 blue and 2 yellow marbles



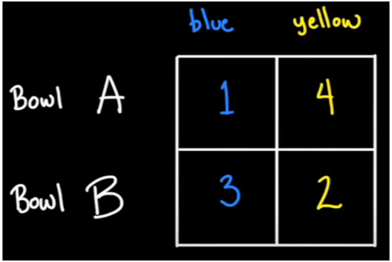
- consider two bowls A and B
- bowl A contains 1 blue and 4 yellow marbles
- bowl B contains 3 blue and 2 yellow marbles

- Blue: event of picking blue



- consider two bowls A and B
- bowl A contains 1 blue and 4 yellow marbles
- bowl B contains 3 blue and 2 yellow marbles

- Blue: event of picking blue
- Yellow: event of picking yellow



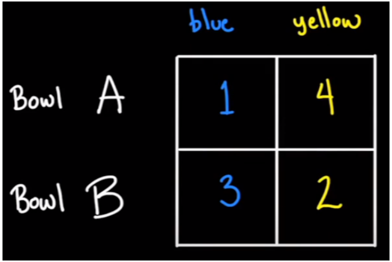
- consider two bowls A and B
- bowl A contains 1 blue and 4 yellow marbles
- bowl B contains 3 blue and 2 yellow marbles

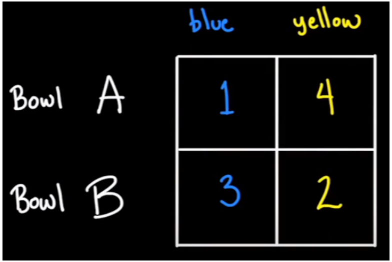
- Blue: event of picking blue
- Yellow: event of picking yellow
- What is $Pr(Blue), Pr(Yellow)$?

Answer:

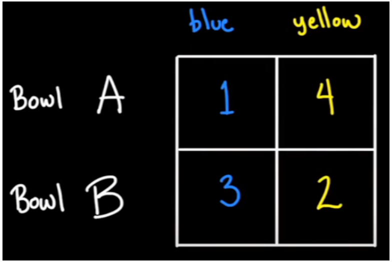
Try!

1 Conditional Probability...





- What is Pr(Blue) given that only bowl A is allowed?



- What is Pr(Blue) given that only bowl A is allowed?

Answer: Pr(Blue | A) = probability to choose blue given that bowl A is fixed

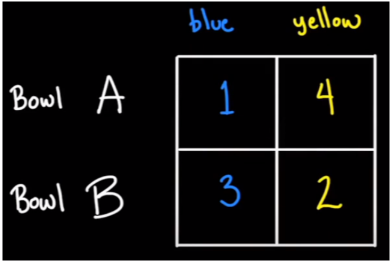
1 Conditional Probability and Choice Tree...

	blue	yellow
Bowl A	1	4
Bowl B	3	2

	blue	yellow
Bowl A	1	4
Bowl B	3	2

- Question: What is $P(A|blue)$?

1 Conditional Probability and Choice Tree...



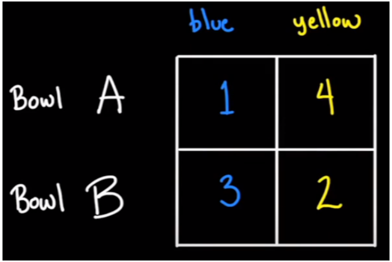
- Question: What is $P(A|blue)$?
- Question: Is $P(A|blue) = P(blue|A)$?

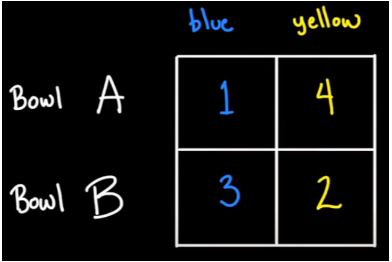
Answer:

	blue	yellow
Bowl A	1	4
Bowl B	3	2

- Question: What is $P(A|blue)$?
- Question: Is $P(A|blue) = P(blue|A)$?

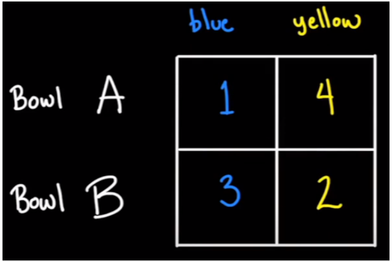
1 Conditional Probability and Choice Tree...





- Draw choice tree, given that, after picking, the ball is not placed back in bowl

1 Conditional Probability and Choice Tree...



- Draw choice tree, given that, after picking, the ball is not placed back in bowl

Movie Monty Hall Brooklyn Video Here!

2 Conditional Probability and Game of Chance Movie...

Movie Monty Hall Movie 21 Video Clip Here!

2 Conditional Probability and Game of Chance Movie...

Movie Monty Hall Youtube Video Here

Another Monty Hall Youtube Movie Here!

2 Conditional Probability and Game of Chance...

2 Conditional Probability and Game of Chance...

- There is a game show.

2 Conditional Probability and Game of Chance...

- There is a game show. The show host shows you three doors

- There is a game show. The show host shows you three doors
- Behind one of the doors, there is a car,

2 Conditional Probability and Game of Chance...

- There is a game show. The show host shows you three doors
- Behind one of the doors, there is a car, and in other two there are goats

2 Conditional Probability and Game of Chance...

- There is a game show. The show host shows you three doors
- Behind one of the doors, there is a car, and in other two there are goats
- Rules of the Game Show:

2 Conditional Probability and Game of Chance...

- There is a game show. The show host shows you three doors
- Behind one of the doors, there is a car, and in other two there are goats
- Rules of the Game Show:
  - you are allowed to pick a door without opening

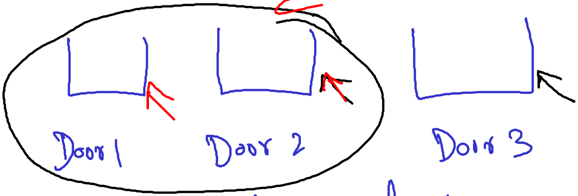
- There is a game show. The show host shows you three doors
- Behind one of the doors, there is a car, and in other two there are goats
- Rules of the Game Show:
  - you are allowed to pick a door without opening
  - then the host opens a door

2 Conditional Probability and Game of Chance...

- There is a game show. The show host shows you three doors
- Behind one of the doors, there is a car, and in other two there are goats
- Rules of the Game Show:
  - you are allowed to pick a door without opening
  - then the host opens a door
- Question: if the host always opens goat door,

- There is a game show. The show host shows you three doors
- Behind one of the doors, there is a car, and in other two there are goats
- Rules of the Game Show:
  - you are allowed to pick a door without opening
  - then the host opens a door
- Question: if the host always opens goat door, is it wise to change your door?

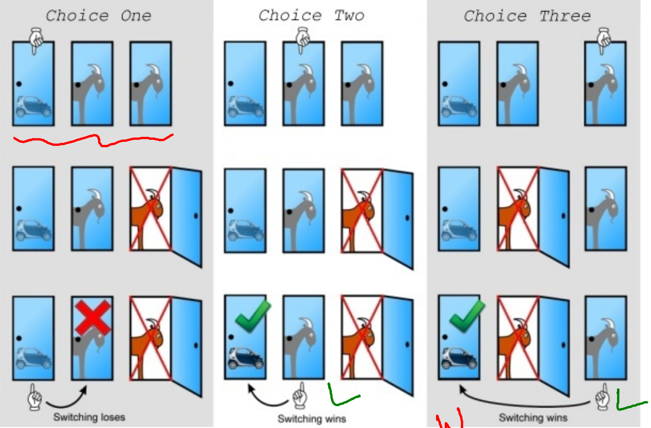
step 3 :



Step 1 : you pick one door

Step 2 : host opens one of the other two door. He definitly opens door with the goat.

2 Solution to Monty Hall Problem with Graphical Illustration



W L W L W L

$\frac{2}{3}$ If switch

$\frac{1}{3}$ If not switch

$\frac{1}{3}$

Figure: Graphical illustration of Monty hall problem. Source: Google

2 Solution to Game Show: Choice Tree, Conditional Probability

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose
1	2	3	loose	win

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose
1	2	3	loose	win
1	3	2	loose	win

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose
1	2	3	loose	win
1	3	2	loose	win
2	1	3	loose	win

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	$2/3$	win	loose
1	2	3	loose	win
1	3	2	loose	win
2	1	3	loose	win
2	2	$1/3$	win	loose

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose
1	2	3	loose	win
1	3	2	loose	win
2	1	3	loose	win
2	2	1/3	win	loose
2	3	1	loose	win

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose
1	2	3	loose	win
1	3	2	loose	win
2	1	3	loose	win
2	2	1/3	win	loose
2	3	1	loose	win
3	1	2	loose	win

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	$2/3$	win	loose
1	2	3	loose	win
1	3	2	loose	win
2	1	3	loose	win
2	2	$1/3$	win	loose
2	3	1	loose	win
3	1	2	loose	win
3	2	1	loose	win

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose
1	2	3	loose	win
1	3	2	loose	win
2	1	3	loose	win
2	2	1/3	win	loose
2	3	1	loose	win
3	1	2	loose	win
3	2	1	loose	win
3	3	1/2	win	loose

Table: Exhaustive list of possibilities

$\frac{6}{9}$ $\frac{2}{3}$

$\frac{3}{9} = \frac{1}{3}$

Conclusion

If you switch, the probability that you win a car is 2/3,

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose
1	2	3	loose	win
1	3	2	loose	win
2	1	3	loose	win
2	2	1/3	win	loose
2	3	1	loose	win
3	1	2	loose	win
3	2	1	loose	win
3	3	1/2	win	loose

Table: Exhaustive list of possibilities

Conclusion

If you switch, the probability that you win a car is 2/3, and if you switch,

2 Solution to Game Show: Choice Tree, Conditional Probability

Let us look into all possible (exhaustive) cases:

Door You Choose	Prize in Door	Host Opens	Stay	Switch
1	1	2/3	win	loose
1	2	3	loose	win
1	3	2	loose	win
2	1	3	loose	win
2	2	1/3	win	loose
2	3	1	loose	win
3	1	2	loose	win
3	2	1	loose	win
3	3	1/2	win	loose

Table: Exhaustive list of possibilities

Conclusion

If you switch, the probability that you win a car is $2/3$, and if you switch, the probability that you win goat is $1/3$.

2 Solution to Monty Hall Problem with Choice/Decision Tree

2 Solution to Monty Hall Problem with Choice/Decision Tree

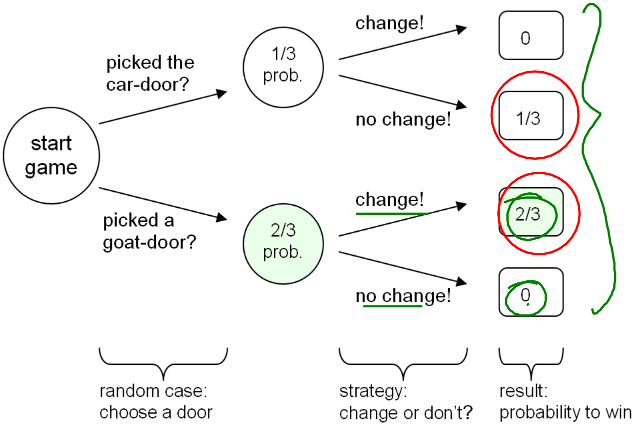


Figure: Graphical illustration of Choice Tree of Monty hall problem. Source: Google

2 Using Bayes Theorem in Monty Hall's Problem...

2 Using Bayes Theorem in Monty Hall's Problem...

- Let H be the hypothesis "door 1 has a car behind it," and E be the evidence that Monty has revealed a door with a goat behind it

- Let H be the hypothesis "door 1 has a car behind it," and E be the evidence that Monty has revealed a door with a goat behind it
- Then the problem can be restated as calculating $P(H \mid E)$, the conditional probability of H given E

2 Using Bayes Theorem in Monty Hall's Problem...

- Let H be the hypothesis "door 1 has a car behind it," and E be the evidence that Monty has revealed a door with a goat behind it
- Then the problem can be restated as calculating $P(H \mid E)$, the conditional probability of H given E
- Since every door either has a car or a goat behind it, the hypothesis "H^c" is the same as "door 1 has a goat behind it"

- Let H be the hypothesis "door 1 has a car behind it," and E be the evidence that Monty has revealed a door with a goat behind it
- Then the problem can be restated as calculating $P(H \mid E)$, the conditional probability of H given E
- Since every door either has a car or a goat behind it, the hypothesis "H^c" is the same as "door 1 has a goat behind it"

Write the following in words:

- $P(H) =$ 1/3
- $P(H^c) =$ $1 - 1/3 = 2/3$
- $P(E|H) =$ P(monty opens goat door | door 1 has car behind it) = 1
- $P(E|H^c) =$

$P(H^c|E) = 2/3$ $\quad P(E|H^c) = 1$

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} = \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|H^c)P(H^c)} = \frac{1 \cdot 1/3}{1 \cdot 1/3 + 1 \cdot 2/3} = \frac{1}{3}$$