Probability and Statistics

UG2, Core course, IIIT,H

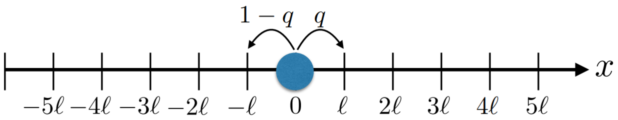
Pawan Kumar

IIIT, Hyderabad

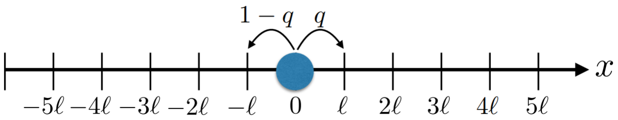
August 31, 2021

1 Outline

1 Random Walk, Probability

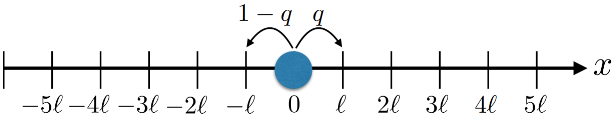


1 Random Walk, Probability



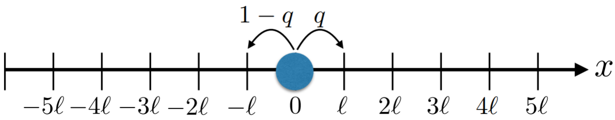
- Consider a person at $x = 0$, he can travel one step to the right or to the left

1 Random Walk, Probability

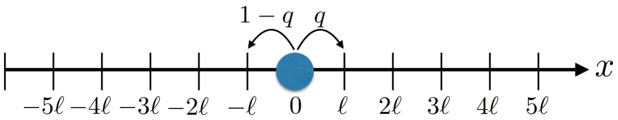


- Consider a person at $x = 0$, he can travel one step to the right or to the left
  - He can travel one step to the right with probability q

1 Random Walk, Probability



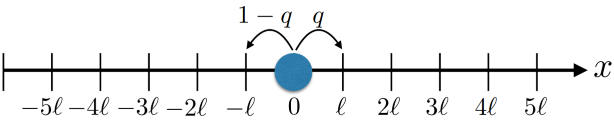
- Consider a person at $x = 0$, he can travel one step to the right or to the left
  - He can travel one step to the right with probability q
  - He can travel one step to the left with probability $(1 - q)$



- Consider a person at $x = 0$, he can travel one step to the right or to the left
  - He can travel one step to the right with probability q
  - He can travel one step to the left with probability $(1 - q)$

Simple Random Walk in 1D

A walk is called simple random walk in 1D if there is a equal probability of either going to right or going to the left. Above, we set $p = q = 1/2$



- Consider a person at $x = 0$, he can travel one step to the right or to the left
  - He can travel one step to the right with probability q
  - He can travel one step to the left with probability $(1 - q)$

Simple Random Walk in 1D

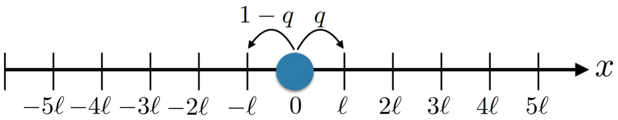
A walk is called simple random walk in 1D if there is a equal probability of either going to right or going to the left. Above, we set $p = q = 1/2$

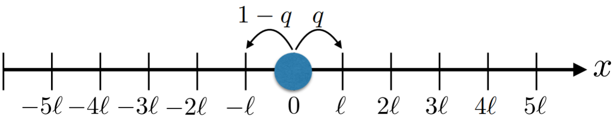
Question

What is the probability that the person after ith step is at $x = 0$?

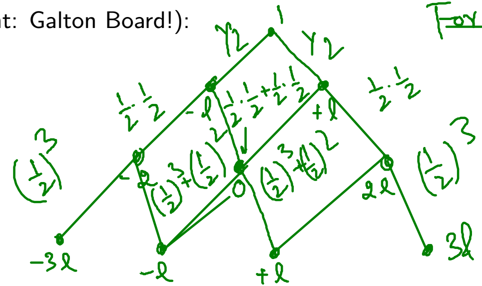
1 Analysis of Simple Random Walk in 1D

1 Analysis of Simple Random Walk in 1D





Draw the choice tree (Hint: Galton Board!):



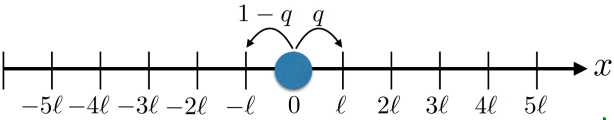
For a simple random walk

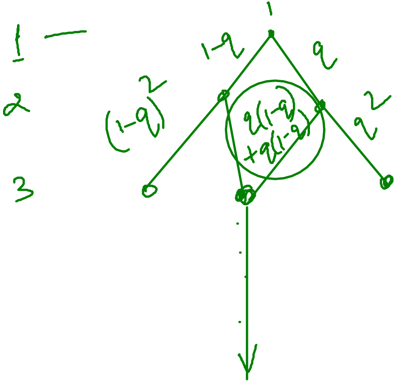
$p = q = 1/2$

$\frac{1}{2} \cdot \frac{1}{2}$

1 Analysis of Biased Random Walk in 1D, Binomial Distribution

1 Analysis of Biased Random Walk in 1D, Binomial Distribution

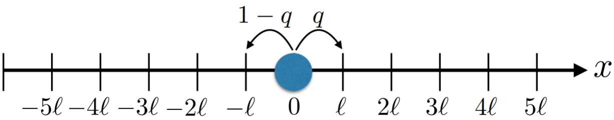




Q Write down the expression for prob. at i^{th} node of j^{th} level.

Q Prob to come back at 0 at i^{th} step.

1 Analysis of Biased Random Walk in 1D, Binomial Distribution

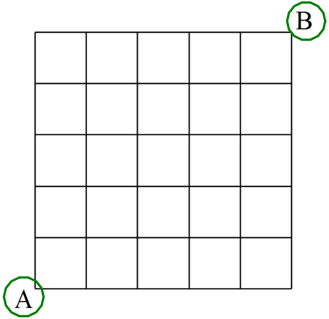


Draw the choice tree for unbiased random walk, derive binomial distribution:

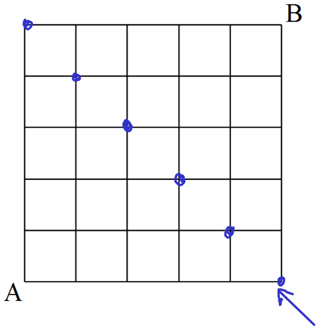
1 Random Walkers on 2D Grid...

1 Random Walkers on 2D Grid...

Alice moves = {right, up}
Bob moves = {left, down}

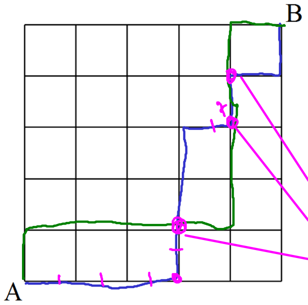


- Alice starts at point A, and each second, walks one edge right or up (if a point has two options, each direction has a 50



- Alice starts at point A, and each second, walks one edge right or up (if a point has two options, each direction has a 50
- At the same time, Bob starts at point B, and each second he walks one edge left or down (if a point has two options, each direction has a 50

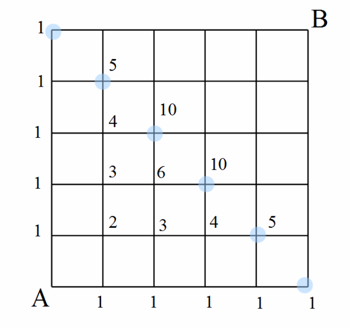
- Point where they can meet at same time (i.e., in this case 5 seconds).



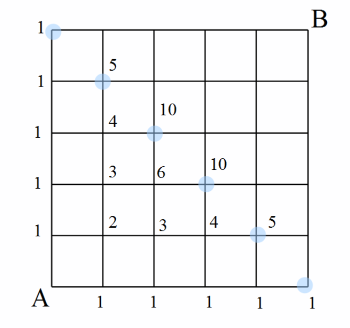
- Alice starts at point A, and each second, walks one edge right or up (if a point has two options, each direction has a 50
- At the same time, Bob starts at point B, and each second he walks one edge left or down (if a point has two options, each direction has a 50
- What is the probability Alice and Bob meet during their random walks?

The paths crossed but they did not meet due to time difference.

1 Solution: Random Walkers on 2D Grid...



- If Alice and Bob meet, they can meet only at the blue circles, which is after both have taken 5 steps



- If Alice and Bob meet, they can meet only at the blue circles, which is after both have taken 5 steps
- Number of ways they could have taken the 5 steps is 2^5 each, so combined, by product rule they can take $\underline{2^5 2^5} = 4^5 = \underline{1024}$ total paths!

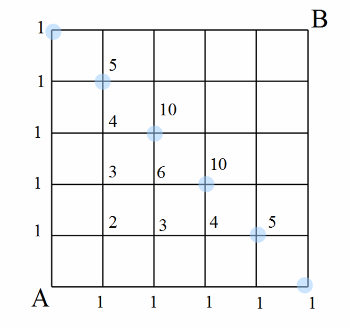
$A = \{R, U\}$

$\underline{2}\ \underline{2}\ \underline{2}\ \underline{2}\ \underline{2} = 2^5$

$B = \{L, D\}$

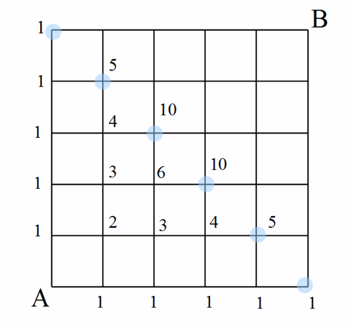
$\underline{2}\ \underline{2}\ \underline{2}\ \underline{2}\ \underline{2} = 2^5$

1 Solution: Random Walkers on 2D Grid...

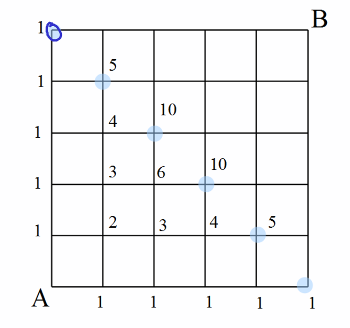


- If Alice and Bob meet, they can meet only at the blue circles, which is after both have taken 5 steps
- Number of ways they could have taken the 5 steps is 2^5 each, so combined, by product rule they can take $2^5 2^5 = 4^5 = 1024$ total paths!
- The number of ways Bob can reach the blue dots is given by binomial coefficients or Pascal's triangle! Same for Alice

1 Solution: Random Walkers on 2D Grid...



- If Alice and Bob meet, they can meet only at the blue circles, which is after both have taken 5 steps
- Number of ways they could have taken the 5 steps is 2^5 each, so combined, by product rule they can take $2^5 2^5 = 4^5 = 1024$ total paths!
- The number of ways Bob can reach the blue dots is given by binomial coefficients or Pascal's triangle! Same for Alice
  - Total ways Bob and Alice could meet at blue dots is square of binomial cofficients



- If Alice and Bob meet, they can meet only at the blue circles, which is after both have taken 5 steps
- Number of ways they could have taken the 5 steps is 2^5 each, so combined, by product rule they can take $\underbrace{2^5 2^5 = 4^5 = 1024}$ total paths!
- The number of ways Bob can reach the blue dots is given by binomial coefficients or Pascal's triangle! Same for Alice
  - Total ways Bob and Alice could meet at blue dots is square of binomial cofficients
- Hence, the total number of ways Bob and Alice could meet

$$1^2 + 5^2 + 10^2 + 5^2 + 1^2 = 252$$

- The probability that they meet is

$$252/1024 = 24.6\%$$

- Experiment: Throw two dice A and B simultaneously

2 Motivation for Conditional Probability with an Example ...

- Experiment: Throw two dice A and B simultaneously
- Event: Odd number on first die

- Experiment: Throw two dice A and B simultaneously
- Event: Odd number on first die
- Question: What is the probability of the event E?

$$S = \{(1,1),(1,2),(1,3),\ldots,(1,6)\\ (2,1),(2,2),\ldots\ldots,(2,6)\\ \vdots\\ (6,1),(6,2),\ldots,(6,6)\}$$

$$|S| = 36$$

$$E = \{(1,1),(1,2),\ldots\ldots,(1,6)\\ (3,1),(3,2),\ldots\ldots,(3,6)\\ (5,1),(5,2),\ldots\ldots,(5,6)\}$$

$$|E| = 18$$

$$P(E) = \frac{18}{36} = \frac{1}{2}$$

- Experiment: Throw two dice A and B simultaneously
- Event: Odd number on first die
- Question: What is the probability of the event E?

- Event: Odd number on first die A, given that even shows on die B

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

- Experiment: Throw two dice A and B simultaneously
- Event: Odd number on first die

$E = \{(1,1), (1,2), \ldots (1,6), (3,1), (3,2) \ldots (3,6), (5,1), (5,2), \ldots (5,6)\}$

6×3

- Question: What is the probability of the event E?

$A|B = \{(1,2), (1,4), (1,6), (3,2), (3,4), (3,6), (5,2), (5,4), (5,6)\}$

$B =$ even shows on die B

$= \{(1,2), (2,2), (3,2), \ldots (6,2), (1,4), (2,4), (3,4), \ldots (6,4), (1,6), (2,6), (3,6), \ldots (6,6)\}$

- Event: Odd number on first die A, given that even shows on die B
- Question: What is the probability of the event E?

$P(E) = \frac{9}{18} = \frac{1}{2}$

{ Since given that B happens, B becomes sample space !

Question

Roll two 6-sided dice, yielding values D_1 and D_2. Let E be the event $D_1 + D_2 = 4$. What is $P(E)$?

$E: D_1 + D_2 = 4.$

$S = \{(1,1), (1,2), \dots, (1,6)$
$(2,1), (2,2), \dots, (2,6)$
$\vdots$
$(6,1), (6,2), \dots, (6,6)\}$

$|S| = 36$

$E = \{(1,3), (2,2), (3,1)\}$

$|E| = 3$

$Pr(E) = \frac{3}{36} = \frac{1}{12}$

Question

Roll two 6-sided dice, yielding values D_1 and D_2. Let E be the event $D_1 + D_2 = 4$. What is $P(E)$?

- $|S| =$

Question

Roll two 6-sided dice, yielding values D_1 and D_2. Let E be the event $D_1 + D_2 = 4$. What is $P(E)$?

- $|S| =$
- $E =$

Question

Roll two 6-sided dice, yielding values D_1 and D_2. Let E be the event $D_1 + D_2 = 4$. What is $P(E)$?

- $|S| =$
- $E =$

$P(E) =$

Question

Roll two 6-sided dice, yielding values D_1 and D_2. Let E be the event $D_1 + D_2 = 4$. What is $P(E)$?

- $|S| =$
- $E =$

$P(E) =$

Question

Roll two 6-sided dice, yielding values D_1 and D_2. Let E be the event $D_1 + D_2 = 4$. Let F be the event $D_1 = 2$. What is $P(E,$ given F already observed)?

$E|F = \{(2,2)\}$

$S_F = \{(2,1), (2,2), \ldots, (2,6)\}$

$F : D_1 = 2$

New sample space, because F is observed!

$P(E|F)$

$Pr(E|F) = \frac{1}{6}$

Question

Roll two 6-sided dice, yielding values D_1 and D_2. Let E be the event $D_1 + D_2 = 4$. What is $P(E)$?

Question

Roll two 6-sided dice, yielding values D_1 and D_2. Let E be the event $D_1 + D_2 = 4$. Let F be the event $D_1 = 2$. What is $P(E, \text{given } F \text{ already observed})$?

- $|S| =$
- $E =$

$P(E) =$
$P(E|F) =$

3 Define Conditional Probability...

3 Define Conditional Probability...

Definition of Conditional Probability

The conditional probability of E given F is the probability that E occurs given that F has already occurred. This is known as conditioning on F.

Definition of Conditional Probability

The conditional probability of E given F is the probability that E occurs given that F has already occurred. This is known as conditioning on F.

- It is denoted by $P(E \mid F)$

Definition of Conditional Probability

The conditional probability of E given F is the probability that E occurs given that F has already occurred. This is known as conditioning on F.

- It is denoted by $P(E \mid F)$
- It means $P(E, \text{given } F \text{ already observed})$

Definition of Conditional Probability

The conditional probability of E given F is the probability that E occurs given that F has already occurred. This is known as conditioning on F.

- It is denoted by $P(E \mid F)$
- It means $P(E, \text{given } F \text{ already observed})$
- Sample space is all possible outcomes consistent with F (i.e., $S \cap F$)

Definition of Conditional Probability

The conditional probability of E given F is the probability that E occurs given that F has already occurred. This is known as conditioning on F.

- It is denoted by $P(E \mid F)$
- It means $P(E, \text{given } F \text{ already observed})$
- Sample space is all possible outcomes consistent with F (i.e., $S \cap F$)
- Event is all outcomes in E consistent with F (i.e., $E \cap F$)

3 Define Conditional Probability...

Definition of Conditional Probability

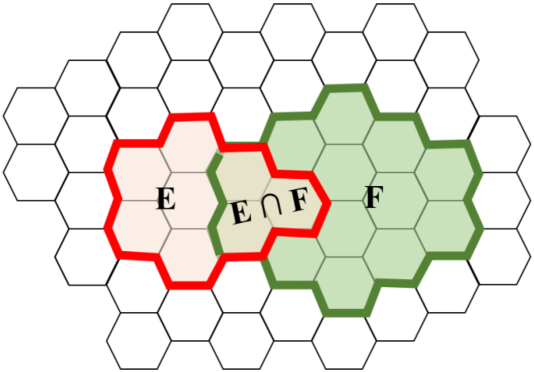
The conditional probability of E given F is the probability that E occurs given that F has already occurred. This is known as conditioning on F.

- It is denoted by $P(E \mid F)$
- It means $P(E, \text{given } F \text{ already observed})$
- Sample space is all possible outcomes consistent with F (i.e., $S \cap F$) $F \subseteq S$
- Event is all outcomes in E consistent with F (i.e., $E \cap F$)

With equally likely outcomes:

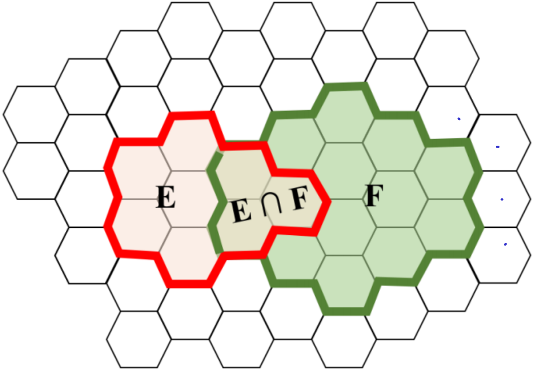
$$P(E|F) = \frac{|E \cap F|}{|S \cap F|} = \frac{|E \cap F|}{|F|}$$

3 Graphical Illustration/Example of Conditional Probability...



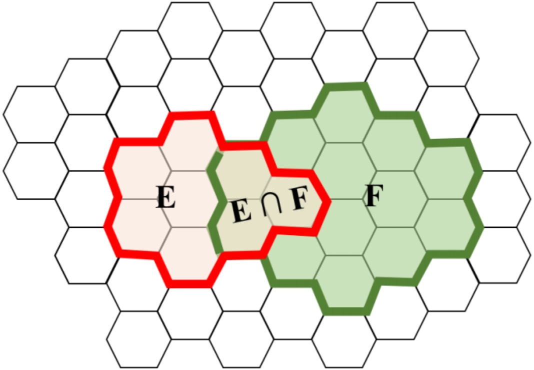
- Question: What is $P(E)$?

3 Graphical Illustration/Example of Conditional Probability...



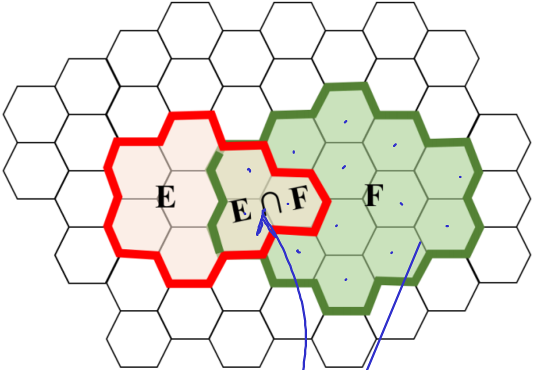
- Question: What is $P(E)$? Here $P(E) = \frac{8}{50} \approx 0.16$

3 Graphical Illustration/Example of Conditional Probability...



- Question: What is $P(E)$? Here $P(E) = \dfrac{8}{50} \approx 0.16$
- Question: What is $P(E \mid F)$?

3 Graphical Illustration/Example of Conditional Probability...



- Question: What is $P(E)$? Here $P(E) = \frac{8}{50} \approx 0.16$
- Question: What is $P(E \mid F)$? Here $P(E \mid F) = \frac{3}{14} \approx 0.21$

3 Probability of Receiving Spam Emails...

Email Spam Conditional Probability Problem

24 emails are sent, 6 each to 4 users.

- 10 of the 24 emails are spam.
- All possible outcomes are equally likely.

Question-1

Let event $E =$ user 1 receives 3 spam emails. What is $P(E)$?

3 Probability of Receiving Spam Emails...

6 6 6 6

3 spam emails

$P(E) = ?$

$$\binom{10}{3}\binom{14}{3} \Big/ \binom{24}{6}$$

Email Spam Conditional Probability Problem

24 emails are sent, 6 each to 4 users.

- 10 of the 24 emails are spam.
- All possible outcomes are equally likely.

Question-1

Let event E = user 1 receives 3 spam emails. What is $P(E)$?

Question-2

Let event F = user 2 receives 6 spam emails. What is $P(E|F)$?

Conditional Probability Implies Chain Rule...

$$\left\{ P(E|F) = \frac{P(E \cap F)}{P(F)} \right. \implies \underbrace{P(E \cap F)} = P(F)P(E|F)$$

These hold even when outcomes are not equally likely!

3 Law of Total Probability...

Conditional Probability Implies Chain Rule...

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

These hold even when outcomes are not equally likely!

Law of Total Probability (Theorem)

$$P(E) = P(E \mid F)P(F) + P(E|F^c)P(F^c)$$

Recall Axiom-3 $P(A \cup B) = P(A) + P(B)$

Conditional Probability Implies Chain Rule...

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

These hold even when outcomes are not equally likely!

Law of Total Probability (Theorem)

$$P(E) = P(E \mid F)P(F) + P(E|F^c)P(F^c)$$

Using these

Proof: $E = (E \cap F) \cup (E \cap F^c)$

are mutually exclusive.

From Axiom 3.

$P(E) = P(E \cap F) + P(E \cap F^c) = P(E|F)\,P(F) + P(E|F^c)\,P(F^c)$

3 Compute $P(E)$ from $P(E|F)$ Using Probability Tree...

3 Compute $P(E)$ from $P(E|F)$ Using Probability Tree...

Problem

Flips a fair coin.

3 Compute $P(E)$ from $P(E|F)$ Using Probability Tree...

Problem

Flips a fair coin.

- If heads: roll a fair 6-sided die.

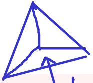
3 Compute $P(E)$ from $P(E|F)$ Using Probability Tree...

Problem

Flips a fair coin.

- If heads: roll a fair 6-sided die.
- Else: roll a fair 3-sided die.

3 Compute $P(E)$ from $P(E|F)$ Using Probability Tree...



bottom side fixed.

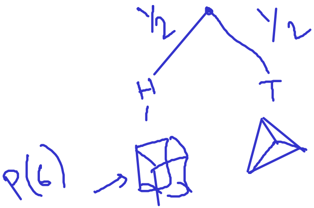
Problem

Flips a fair coin.

- If heads: roll a fair 6-sided die.
- Else: roll a fair 3-sided die.

You win if you roll a 6. What is P(winning)?

Solution using probability tree:



$$\frac{1}{2}\cdot\frac{1}{6}=\frac{1}{12}$$

3 Compute $P(E)$ from $P(E|F)$ Using Total Probability...

3 Compute $P(E)$ from $P(E|F)$ Using Total Probability...

Problem

Flip a fair coin.

Problem

Flip a fair coin.

- If heads: roll a fair 6-sided die.

3 Compute $P(E)$ from $P(E|F)$ Using Total Probability...

Problem

Flip a fair coin.

- If heads: roll a fair 6-sided die.
- Else: roll a fair 3-sided die.

E , $\boxed{F = \{H, T\}}$

$E = \{6\}$.

Problem

Flip a fair coin.

- If heads: roll a fair 6-sided die.
- Else: roll a fair 3-sided die.

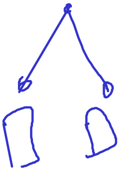
You win if you roll a 6. What is P(winning)?

$$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$$

$$= \frac{\cancel{2}}{12} \cdot \frac{1}{2} + 0 \cdot$$

$$= \frac{1}{12}$$

Solution using total probability:



$$\left\{ \begin{matrix} 1,H & 1,T \\ 2,H & 2,T \\ \vdots & \vdots \\ 6,H & 6,T \end{matrix} \right\}$$

12

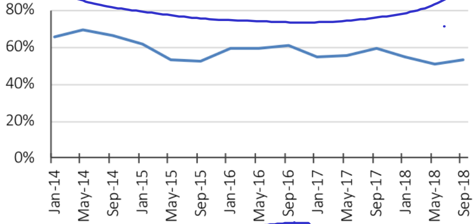
3 Bayes Theorem. Why?

Bayesian Inference

Machine learning

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

Spam volume as percentage of total email traffic worldwide



Data

~~Sample~~

Features → Labels

cat

dog

$P(\text{label} \mid \text{feature})$

want to know!

$P(\text{feature} \mid \text{label})$

Prediction Problem

$P(\text{Features} \mid \text{labels})$

known in a supervised ML.