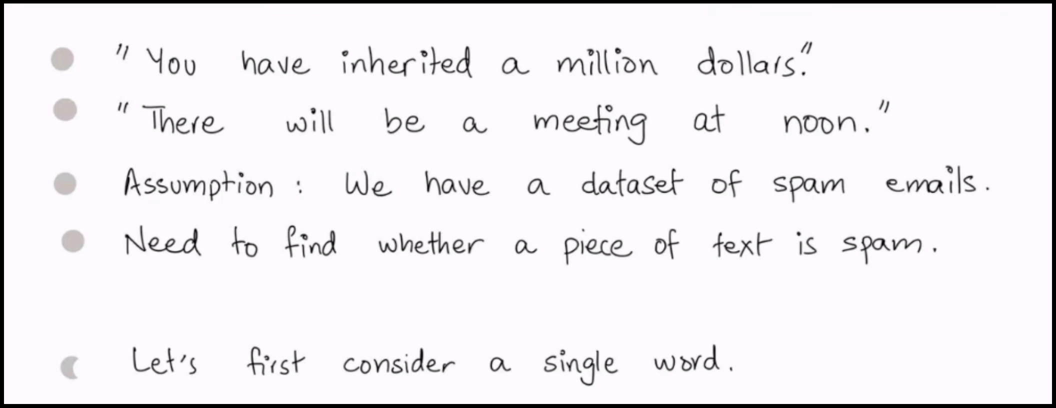
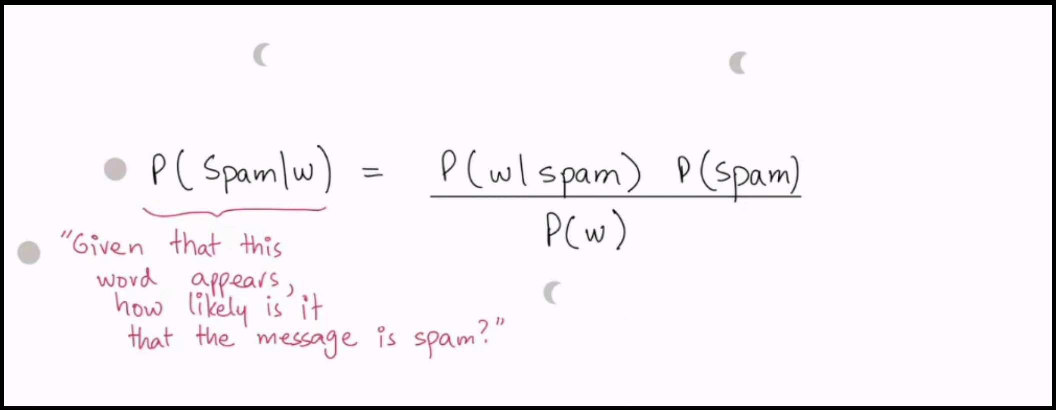
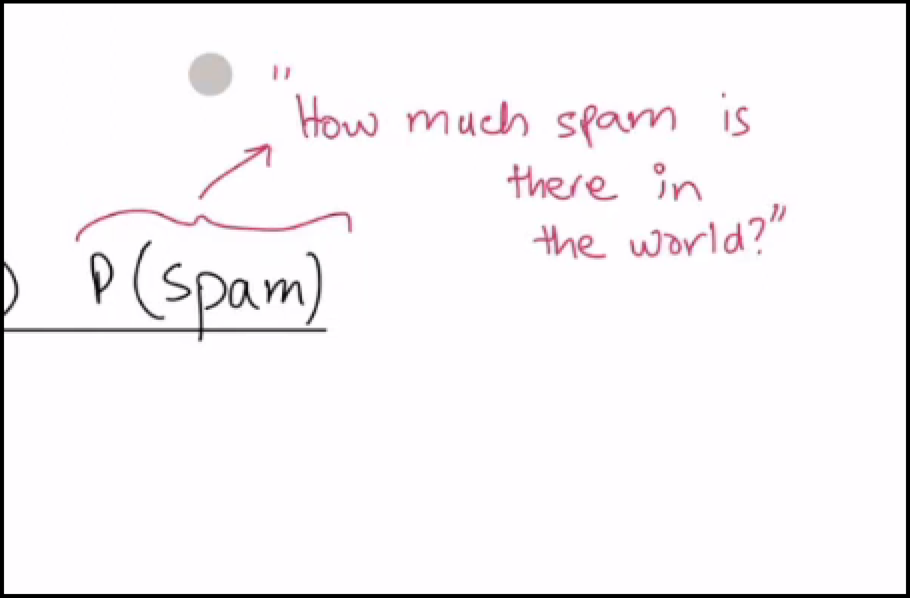
Probability And Statistics Lecture : 12

- Spam Detection
  - 

    - "You have inherited a million dollars."
    - "There will be a meeting at noon."
    - Assumption : We have a dataset of spam emails.
    - Need to find whether a piece of text is spam.

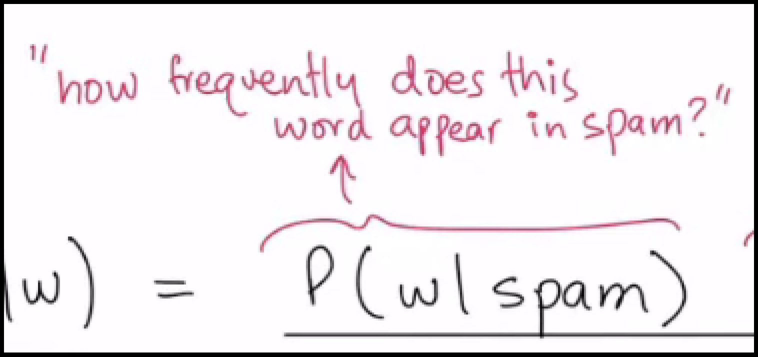
    - Let's first consider a single word.
  - First we know that we must have a dataset to work on and secondly we guess or consider some word to be spam before moving forward to work on it
  - 

    - $$\underbrace{P(Spam|w)}_{} = \frac{P(w|spam)\ P(spam)}{P(w)}$$
    - "Given that this word appears, how likely is it that the message is spam?"
  - With the help of bayesian rule
  - Likelihood into prior divided by normalizing factor

- 

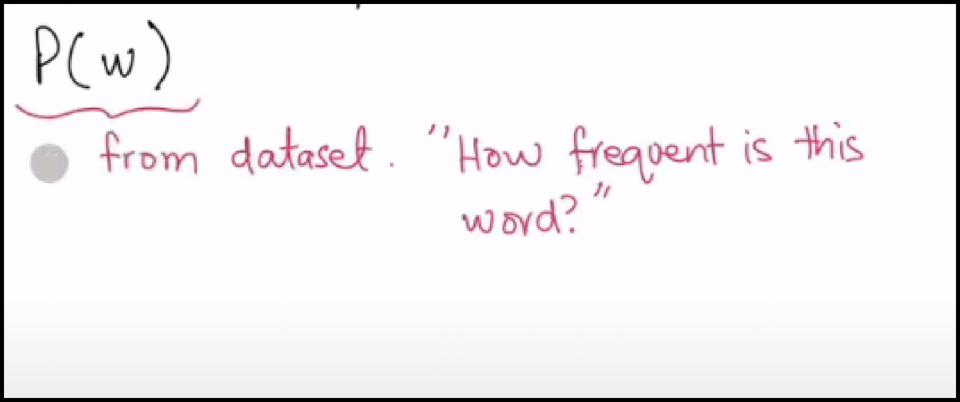
  "How much spam is there in the world?"

  $) \quad \overbrace{P(\text{spam})}$

- So we first find the total population and find the spams messages in it
- 

  "how frequently does this word appear in spam?"

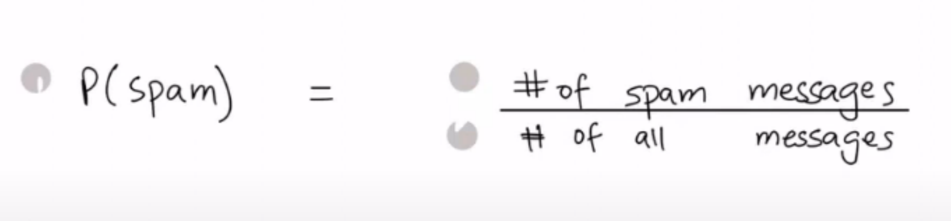
  $|w) = \overbrace{P(w|\text{spam})}$

- In the total spam we find the frequency of the specific word
- 

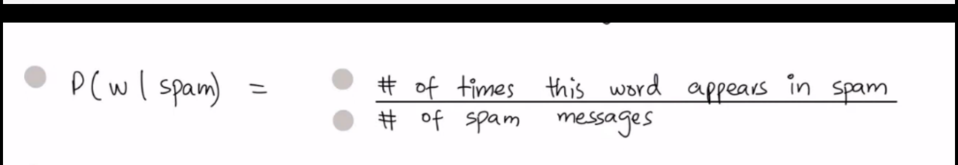
  $\underbrace{P(w)}$

  from dataset. "How frequent is this word?"

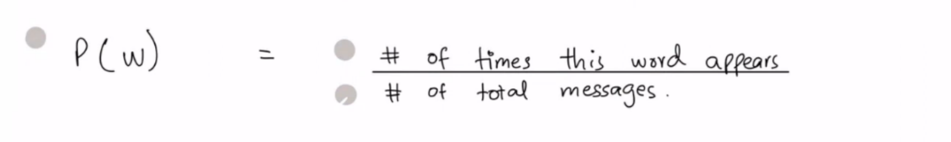
- How frequent is this word

- 

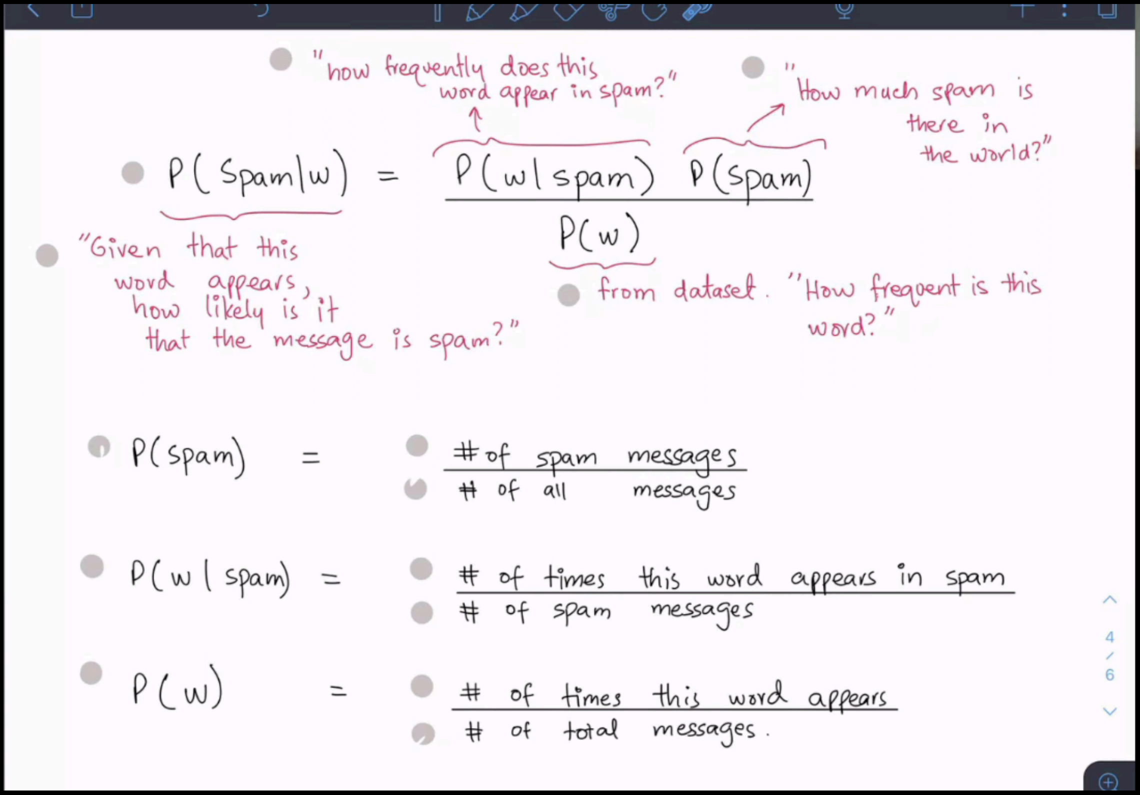
  $$P(spam) = \frac{\#\text{ of spam messages}}{\#\text{ of all messages}}$$

  $$P(w \mid spam) = \frac{\#\text{ of times this word appears in spam}}{\#\text{ of spam messages}}$$

- 

  $$P(w) = \frac{\#\text{ of times this word appears}}{\#\text{ of total messages.}}$$



"how frequently does this word appear in spam?"

"How much spam is there in the world?"

$$P(Spam \mid w) = \frac{P(w \mid spam)\ P(spam)}{P(w)}$$

"Given that this word appears, how likely is it that the message is spam?"

from dataset. "How frequent is this word?"

$$P(spam) = \frac{\#\text{ of spam messages}}{\#\text{ of all messages}}$$

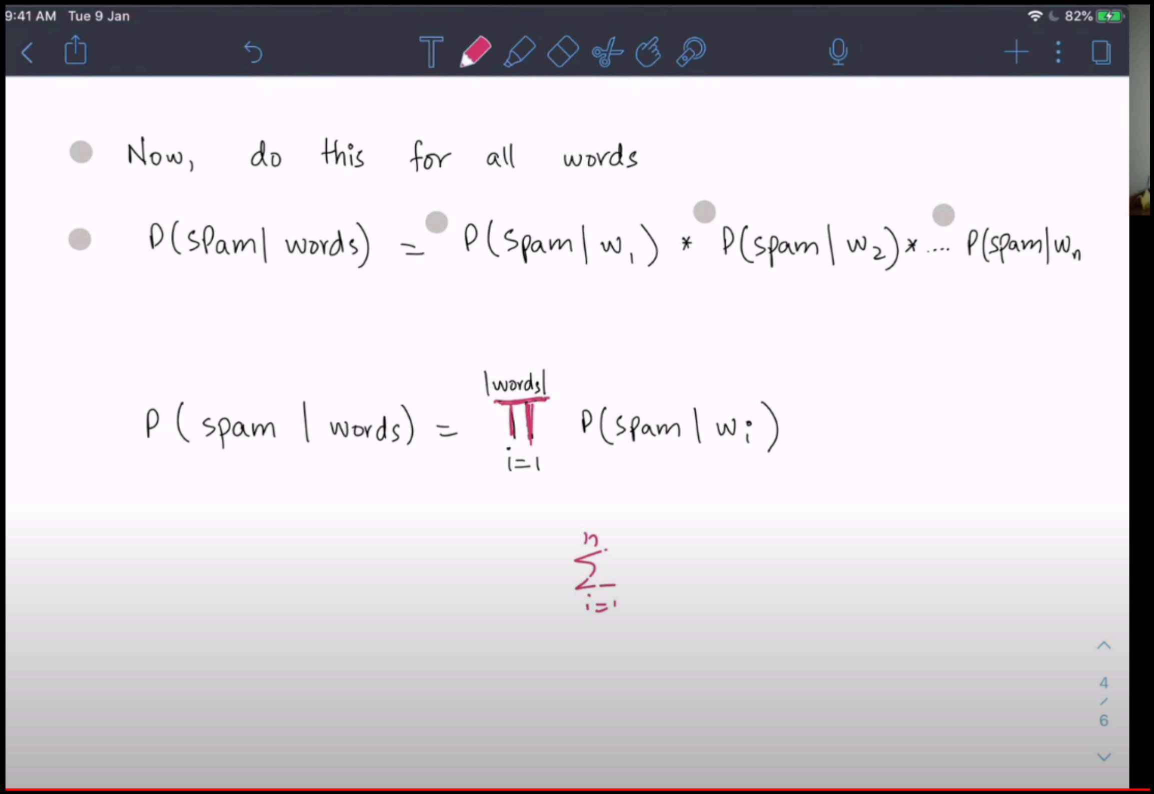
$$P(w \mid spam) = \frac{\#\text{ of times this word appears in spam}}{\#\text{ of spam messages}}$$

$$P(w) = \frac{\#\text{ of times this word appears}}{\#\text{ of total messages.}}$$

- Prior : If there is alot of spam in this world then there new event happening or message then there is a larger possibility or prob of it being a spam
- Likelihood : The more frequently this word appears in messages the greater the prob of it being spam

- The above are all directly proportional

- Now we do this for all the words



- Now, do this for all words
- $P(spam|\ words) = P(spam|w_1) * P(spam|w_2) * \dots P(spam|w_n$

$$P(spam \mid words) = \prod_{i=1}^{|words|} P(spam \mid w_i)$$

$$\sum_{i=1}^{n}$$

-
- Product of all the prob of a message being spam such that the spam word was in it