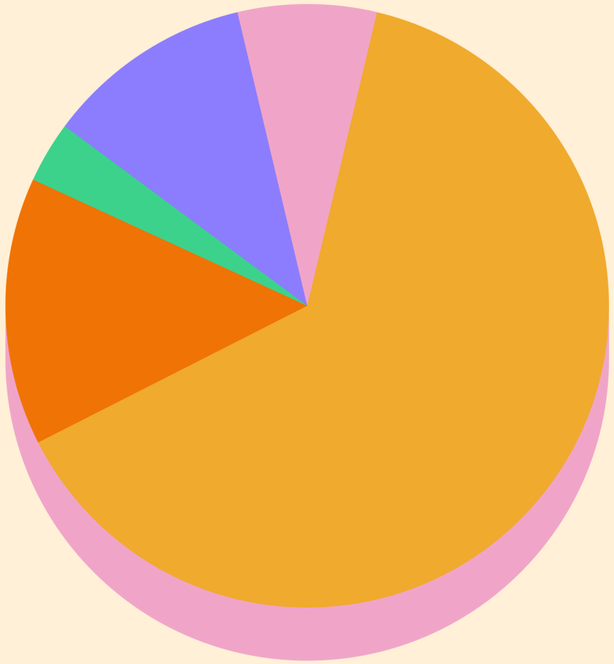
PROBABILITY FOR MACHINE LEARNING



WHY THE PROBABILITY IS IMPORTANT TO ML?

- ML is all about making predictions in uncertain situations. And probability is the math of uncertainty.

classify emails (model)

Is this email spam or not spam?

Won't say → This is 100% spam

Will Say → There's a 90% chance this is Spam

That's Probability!

IN MACHINE LEARNING:

- Real world data → Noisy & uncertain, so that the probability helps:
  - Make decisions with incomplete data
  - Deal with randomness in sampling, predictions, etc.

BASIC PROBABILITY

probability = No. of favourable outcomes / total outcomes.

Coin – H & T

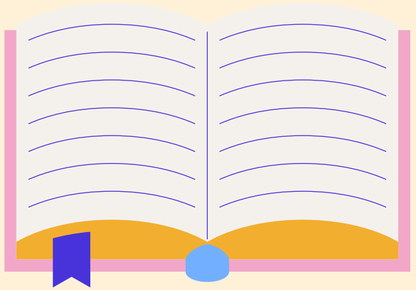
P(H) = ½ = 50%

P(H or T) = ½

Dice = {1,2,3,4,5,6}

P(2) = ⅙

P(even) = n(2,4,6)/n(1,2,3,4,5,6) = 3/6 = ½



COMPLEMENT RULE

$P(A^c) = 1 - P(A)$

eg:

$P(\text{No Rain}) = 1 - P(\text{Rain})$

$= 1 - 0.7$

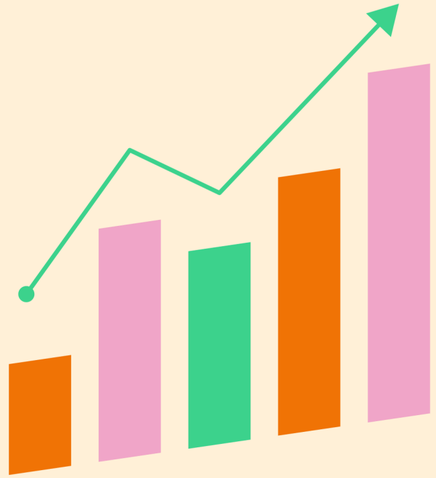
$= 0.3$

if there's a 70% chance of rain today, then there's a 30% chance that it won't rain.

→ Naïve Bayes

→ Logistics Regression

→ Any binary Classification task



INDEPENDENT EVENTS

Two Events are independent if the outcome of one does not affect the outcome of the other.

Coin A

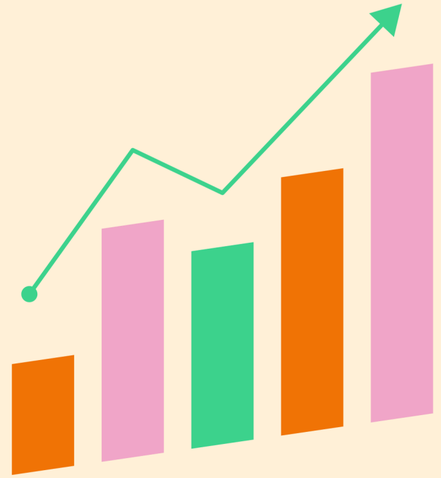
P(H on A) = 0.5 | ½

Coin B

P(H on B) = 0.5 | ½

Combined : P(H on A and B) = 0.5 * 0.5

= 0.25



This is Independence: Both outcomes are unrelated.

INDEPENDENT EVENTS IN ML

In Naïve Bayes, we assume feature independence.
Let's say you're building a spam filter:
-Feature 1: Email has the word "Buy"
-Feature 2: Email has the word "Offer"

The algorithm assumes:
The chance that both "Buy" and "Offer" appear together in a spam email is just:

$$P(\text{Buy}|\text{Spam}) * P(\text{Offer}|\text{Spam})$$

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

Imagine your model looks at different words in an email – like 'Buy', 'offer', 'free'. It acts like each word's presence doesn't depend on the other – Naïve assumption.

CONDITIONAL PROBABILITY

dice = {1,2,3,4,5,6}

If the die shows an odd number, then what us the probability that it is greater than 2?

1.shows an odd number – {1,3,5}

2.greater than 2 in it – {3,5}

P = n(3,5)/n(1,3,5) = ⅔

BAYES' THEOREM

Bayes' Theorem helps you answer this question:

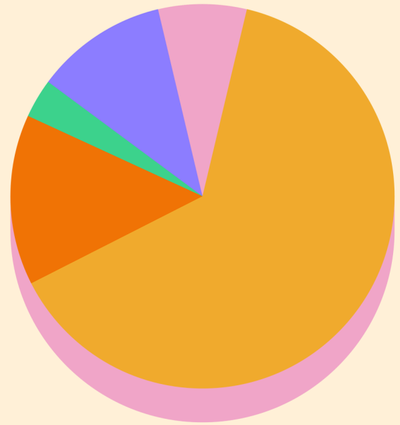
→ If something happened, what's the real reason behind it?

Example: College ID Card Story

College
- CS (100) → Black Hoodies –wear – 30
- Bio (50) → Black Hoodies –wear – 5

you see a student wearing a black hoodie.

→ What's the chance they are from computer science?



WHAT WE KNOW

Department	Total Students	Black hoodie wearers
CS	100	30
Bio	50	5
Total	150	35

P(CS) = 100/150 = 0.67 (Chance anyone is from CS)

P(Black Hoodie|CS) = 30/100 = 0.3

P(Black Hoodie) = 35/150 = 0.233

Bayes' Theorem → P(CS|BlackHoodie) = P(BH|CS) * P(CS)/P(BH)

= 0.3*0.67/0.233 = 0.86 → 86%

PROBABILITY DISTRIBUTION

- A Probability distribution tells us “How the values of a variable are spread” – and how likely each value is,

In ML:

- Understand data behaviour
- make predictions
- calculate loss functions
- detect anomalies (outliers)

DISTRIBUTION

DISCRETE

Used when thevariable can take “Specific countable values (like Yes/No)”

- Bernouli Distribution
- Binomial Distribution

CONTINUOUS

Used when the variable can take “any value in a range” – like height, weight, temperature.

- Uniform Distribution
- Normal Distribution

THANK YOU