Probability Basics

Random Process

- It is a function of time
- We know all the possible outcomes
- But we don't know which outcome will occur exactly
- Also known as ***Stochastic Process***

Probability

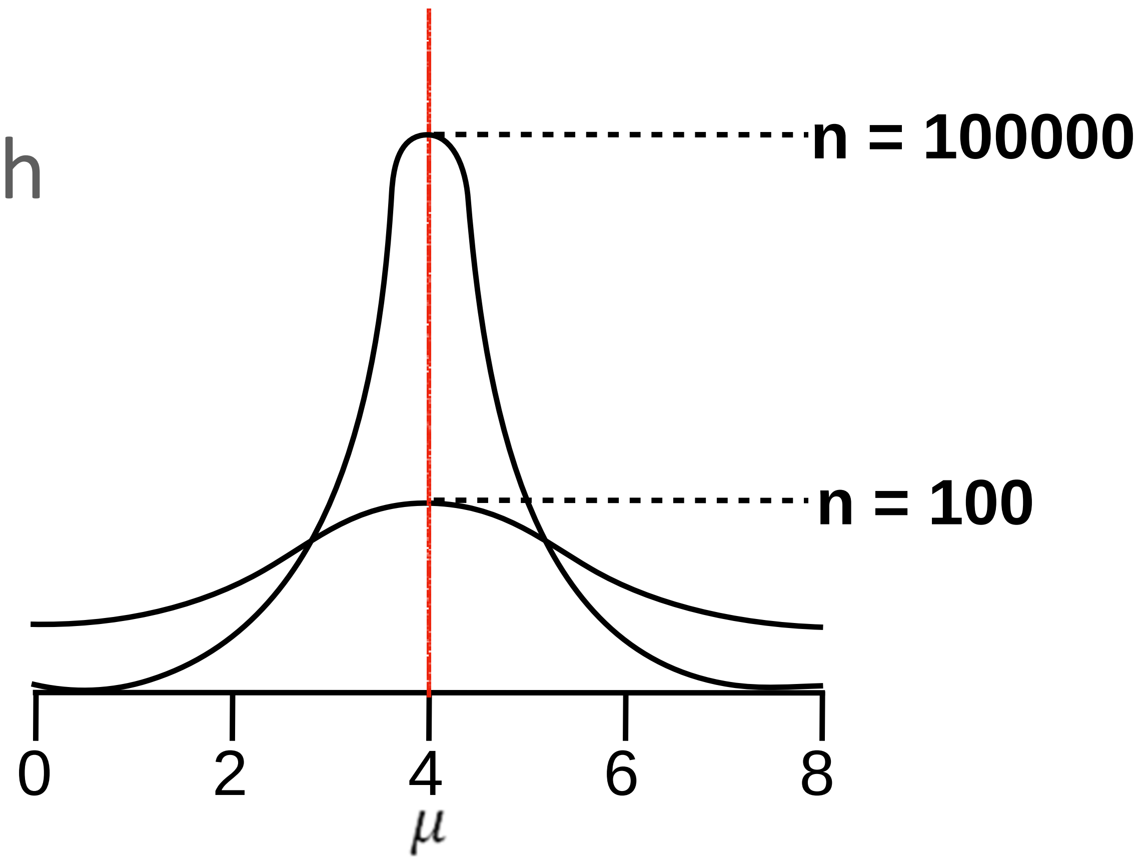
- Quantifiable likelihood (chance) of the occurrence of an event expressed as odds, or a fraction of 1.
- Notation : P(A) = 0.3, read as Probability of Event A is 0.3 or 30%

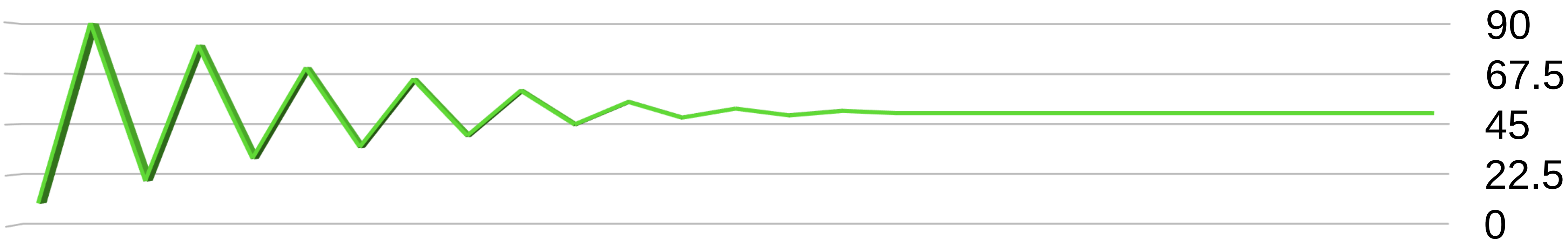
- Example : What is the probability that we get a Head after a coin toss?

P(H) = 0.5

Law of Large Numbers

- The **law of large numbers** is a principle of probability according to which the frequencies of events with the same likelihood of occurrence even out, given enough trials or instances.
- As the **number** of experiments increases, the actual ratio of outcomes will converge on the theoretical, or expected, ratio of outcomes.

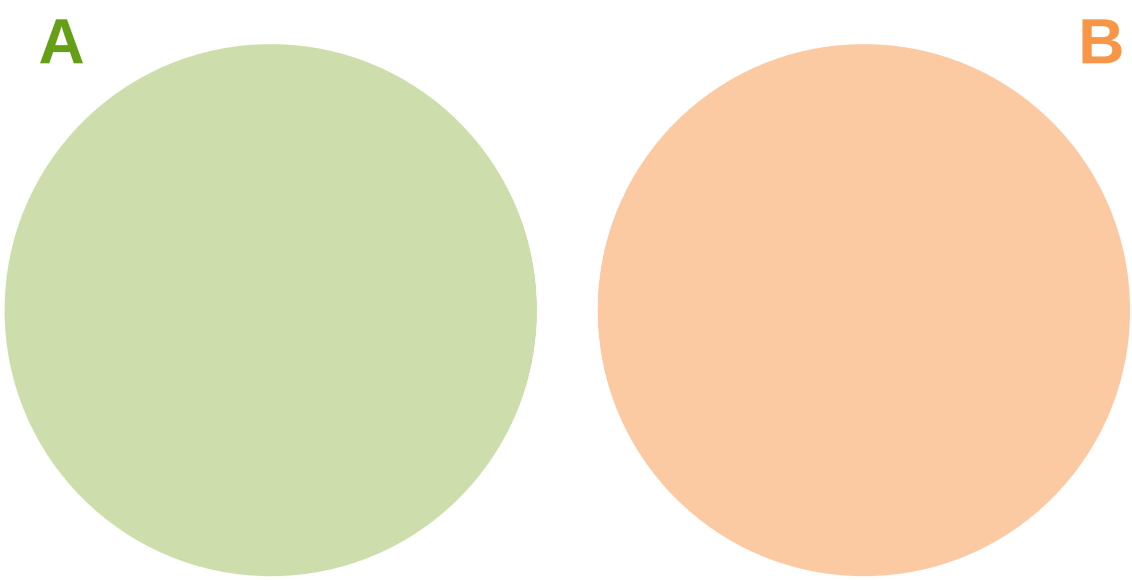




Disjoint Events

- Events that are *mutually exclusive* and cannot happen at the same time.

- Example : Result of a coin toss can either be Head or Tail, not both.

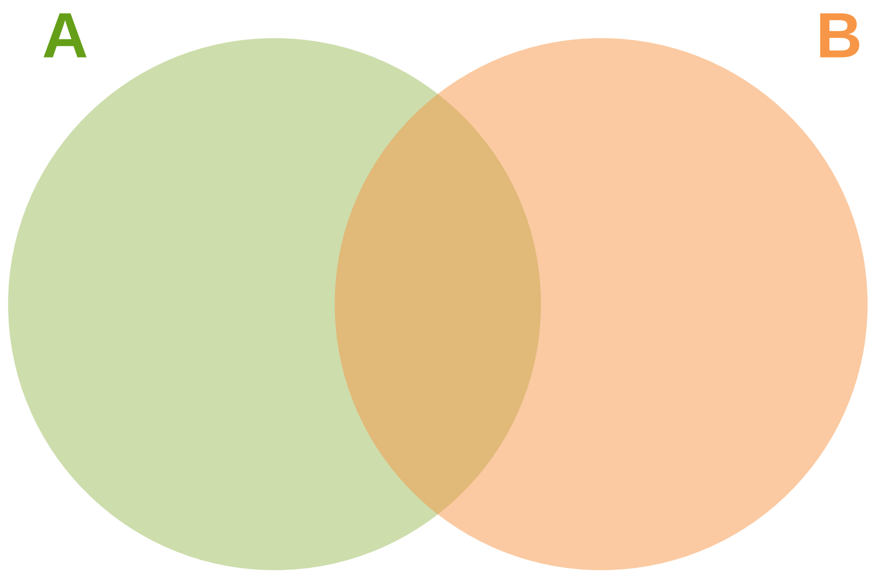


P (A and B) = 0

Non-Disjoint Events

- Events that are not *mutually exclusive* and can happen at the same time.

- Example : A person can like both Biryani and Pizza



P (A and B) ≠ 0

Union of Events

- Disjoint Events : P(A or B) = P(A) + P(B)

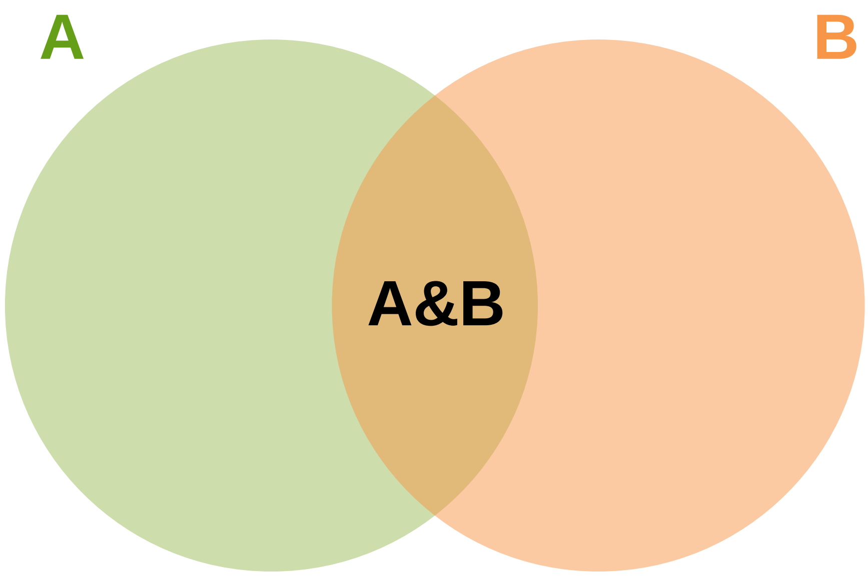
Because P(A and B) = 0

- Non-Disjoint Events : P(A or B) = P(A) + P(B) - P(A and B)

Because P(A and B) ≠ 0

General Addition Rule

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$



Sample Space

- Collection of all possible outcomes of a trial
- Example : What is the sample space of an experiment where two coins are tossed?

S = {HH, HT, TH, TT}

Probability Distributions

- It lists all possible outcomes in the sample space, and the probabilities with which they occur.

One Toss	H	T
Probability	0.5	0.5

Two Tosses	HH	HT	TH	TT
Probability	0.25	0.25	0.25	0.25

Rules :

1. All events must be disjoint
2. Each probability must be between 0 and 1
3. All probabilities must add up to 1

Complimentary Events

- Two Events that are *mutually exclusive* and their probabilities add up to 1.

One Toss	H	T
Probability	0.5	0.5

Two Tosses	HH	HT	TH	TT
Probability	0.25	0.25	0.25	0.25

Complimentary

Complimentary

DICE ANALYTICS

Sensitivity: Internal

Independent Process

- Two processes are said to be independent if knowing the outcome of one provides no useful information about the other.

Product Rule

- If A and B are independent, then

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

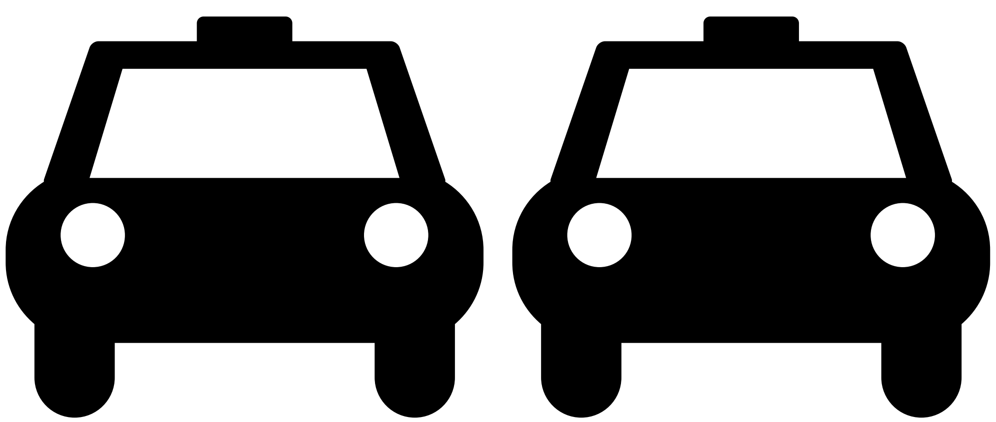
- If A_1 , A_2 ,……, A_k are independent, then

$P\ (A_1 \text{ and } A_2 \text{ and } \ldots\ A_k) = P(A_1) \times P(A_2) \times \ldots. \times P(A_k)$

DICE ANALYTICS

Disjoint vs Independence

Disjoint & Dependent



Non-Disjoint & Independent

Study of Students

	Basketball	No basketball	Total
Cereal	2000	1750	3750
No Cereal	1000	250	1250
Total	3000	2000	5000

Marginal Probability

What is the probability that a student play basketball?

P (basketball) = 3000 / 5000 = 0.60

Note that the counts used to calculate Marginal Probability came from the *margins* of Contingency Table.

Joint Probability

What is the probability that a student play basketball and eat cereal in breakfast?

P (basketball & cereal) = 2000 / 5000 = 0.40

Note that the counts used to calculate Joint Probability came from the *intersection* of Contingency Table.

Conditional Probability

What is the probability that a student play basketball after eating cereal in breakfast?

P (breakfast | cereal) = 2000 / 3750 = 0.53

Note that we first *conditioned* on the cereal and then calculated probability using counts only in this column.

Bayes' Theorem

$$P(A \mid B) = \frac{P(A \,\&\, B)}{P(B)} \quad \rightarrow \quad P(A \,\&\, B) = P(A \mid B) \times P(B)$$

$$P(B \mid A) = \frac{P(A \,\&\, B)}{P(A)} \quad \rightarrow \quad P(A \,\&\, B) = P(B \mid A) \times P(A)$$

$$P(A \,\&\, B) = P(A \mid B) \times P(B) = P(B \mid A) \times P(A)$$

$$P(A \mid B) = \frac{P(B \mid A) \times P(A)}{P(B)}$$

DICE ANALYTICS

General Product Rule

$$P\ (A\ \&\ B) = P\ (A \mid B) \times P\ (B)$$

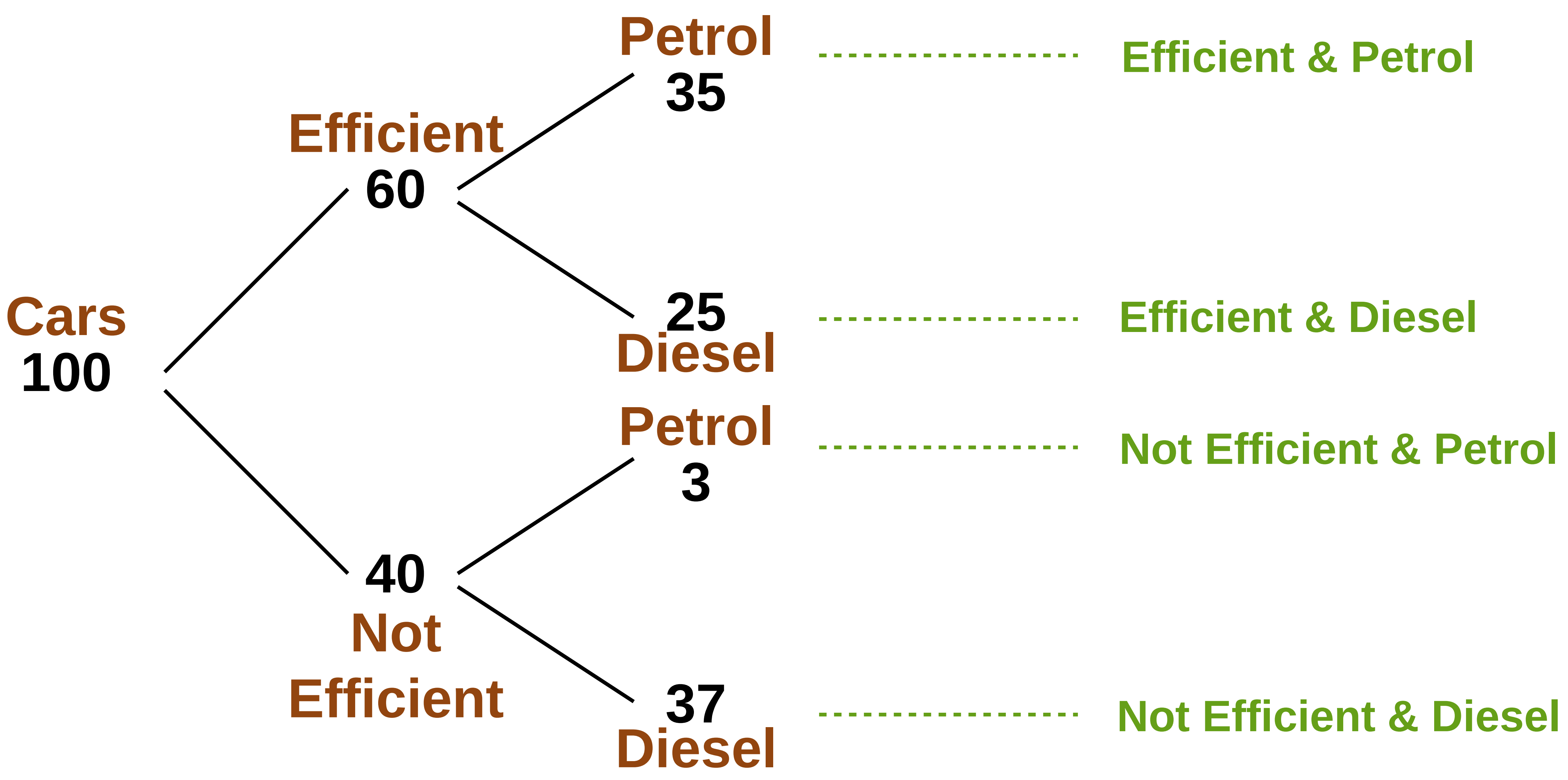
Useful when we don't know whether events are independent or dependent

Probability Trees

Study

There are 100 cars. 60 are fuel efficient while 40 are not fuel efficient. Of the 60 fuel efficient cars, 35 are Petrol while 25 are Diesel. Of the rest, 3 are Petrol while 37 are Diesel.

Probability Tree



Probability Tree (Examples)

If the car is Petrol, what is the probability that it is Fuel Efficient?

$$P\ (\text{Efficient} \mid \text{Petrol}) = \frac{35}{35 + 3} = 0.92$$

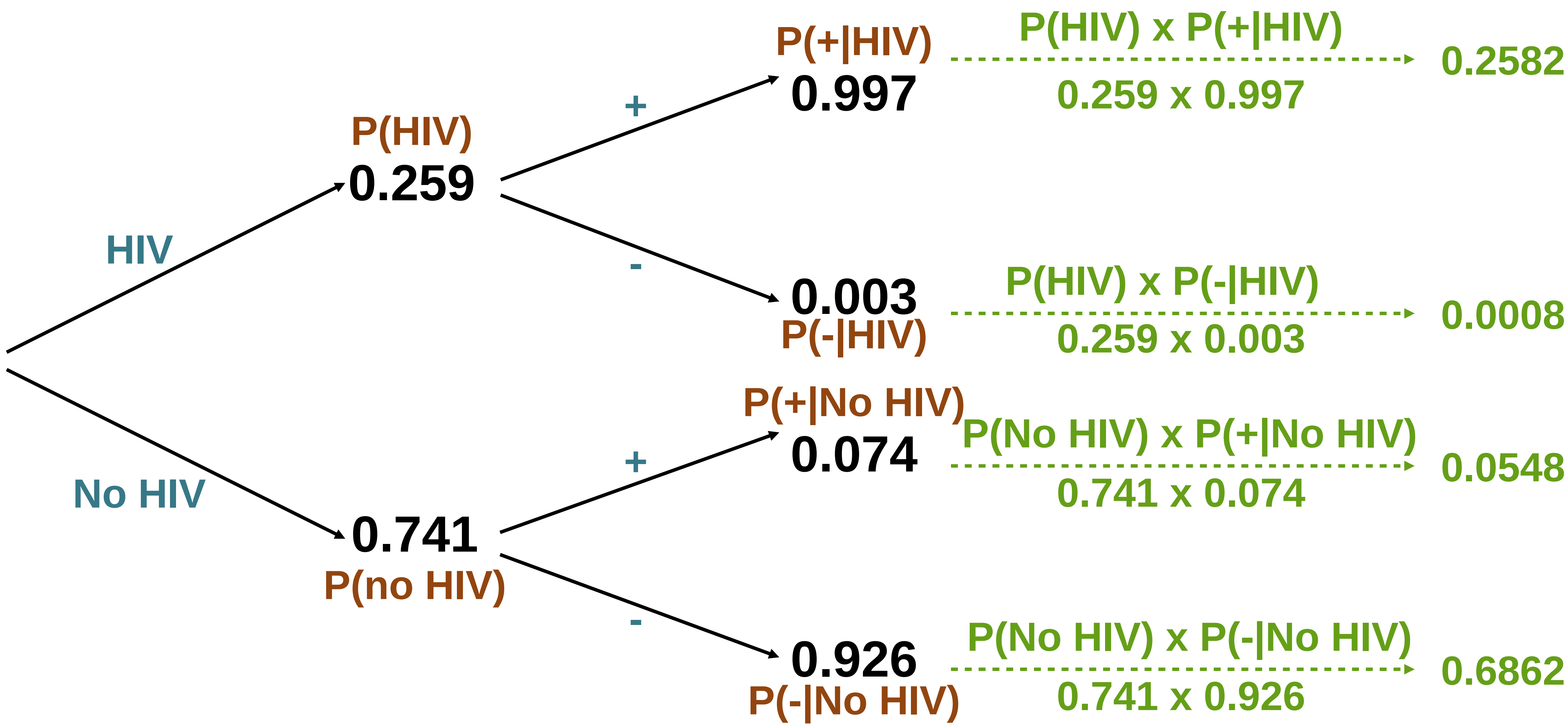
Note that this is a Conditional Probability

Have we made use of Bayes' Theorem here?

Study

As of 2009, Swaziland, has the highest HIV prevalence in the world. 25.9% of this country's population is infected with HIV. The ELISA test is one of the first and most accurate tests for HIV. For those who carry HIV, the ELISA test is 99.7% accurate. For those who do not carry HIV, the test is 92.6% accurate.

Probability Tree



Note that P(HIV) x P(+|HIV) means P(HIV & +), recall General Product Rule!

Probability Tree (Examples)

If an individual from Swaziland has tested positive, what is the probability that he carries HIV?

$$P\,(HIV \mid +) = \frac{P\,(HIV \,\&\, +)}{P(+)} = \frac{0.2582}{0.2582 + 0.0548} = 0.82$$

Note that this is a Conditional Probability

Have we made use of Bayes' Theorem here?