UNIVERSITY OF LINCOLN

SCHOOL OF MATHEMATICS AND PHYSICS

MTH1005
PROBABILITY AND STATISTICS

Semester B

(13/2/2024)

Danilo Roccatano

Office: INB 3323
Email: droccatano@lincoln.ac.uk

CONDITIONAL PROBABILITY AND INDEPENDENCE

Learning outcomes:

- Conditional probability
- Definition of independence of events
- Bayes' Theorem
- Tree and Venn diagrams
- Law of total probability

PROBABILITY FUNCTION

In the previous lecture, we defined the probability function (sometimes also called a probability distribution function) for a sample space, S.

The sample space lists all possible outcomes of an experiment. For instance, for a single roll of a 6-sided die, a possible sample space is S = {1, 2, 3, 4, 5, 6}.

It is a function that, if you feed it an outcome of an experiment, produces a probability x,

$$P(i \in S\ ) = P(\{i\}) = 0 \leq x \leq 1$$

Using this function, we can calculate the probability of an event, E

$$P(E) = \sum_{i \in E} p(i)$$

I.e. we add together the probabilities of all the outcomes that are in the event, E .

PROBABILITY FUNCTION

EXAMPLE

For instance, the

'getting an even number larger than 3 when rolling a single dice' =

$$A = \{2, 4, 6\} \cap \{4, 5, 6\} = \{4, 6\}$$

is

$$P(A) = \sum_{i \in A} p(i) = p(4) + p(6) = \frac{2}{6}$$

CONDITIONAL PROBABILITY

The concept of conditional probability enables us to compute probabilities in situations where we possess only partial information about the system or to reevaluate probabilities upon receiving new information.

In simpler terms, conditional probability refers to the probability of an event E transpiring given that another event F has already occurred.

This is represented mathematically as

$$P(E \mid F)$$

CONDITIONAL PROBABILITY

Example: Consider rolling two dice. We have a sample space

$$S = \{(i, j), i = 1, 2, 3, 4, 5, 6, j = 1, 2, 3, 4, 5, 6\}$$

where say the outcome (i, j) is the first die lands with *i* dots up and the second die with j.

We assume the dice are fair - we assign a probability (distribution) function of

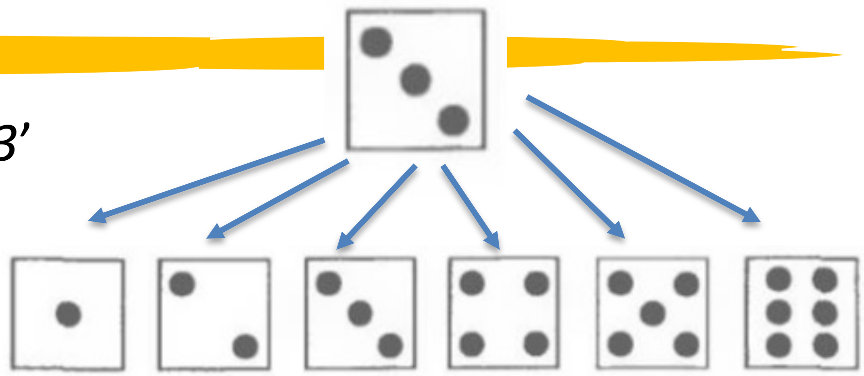
$$P(\{(i, j)\}) = p(i, j) = \frac{1}{36}$$

to all outcomes.

Suppose the first die comes down a three, 

what is the probability that the sum of the two dice equals eight?

CONDITIONAL PROBABILITY



Define F = *'the first die comes down showing a 3'*

F = {(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)}.

This extra information, allows us to cut down the number of possible outcomes. Effectively only 6 possible outcomes can now occur.
Each of these remaining outcomes is still equally likely (what do you think?).

So, to get the probability we are after we need to find the proportion of the remaining 6 outcomes that are in

E = 'the sum of the dice = 8' = {(2, 6), (3, 5), (4, 2), (5, 3), (6, 2)}.

The only outcome that satisfies the criteria is {(3, 5)} and once we know that F has occurred, it has a one in 6 chance that {(3, 5)} will occur.

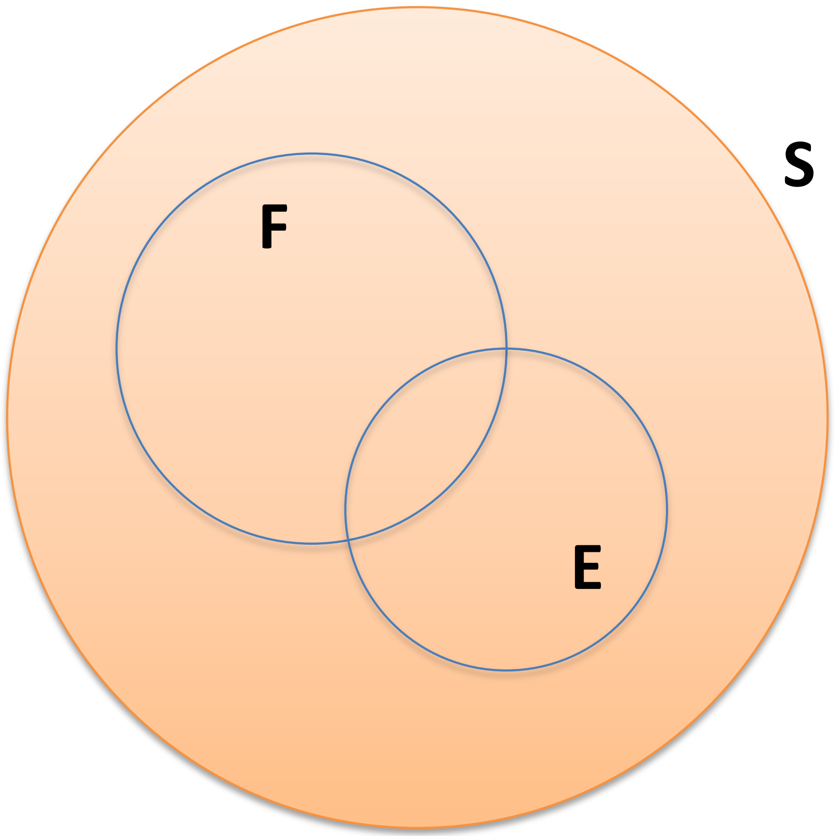
$$P(E \mid F) = \frac{1}{6}$$

Note also that {(3, 5)}= E ∩ F

CONDITIONAL PROBABILITY

The probability of the outcome (3,5) before rolling the first die was 1/36. Following the result of the first die showing a 3, we have asserted that the probability should be 1/6.

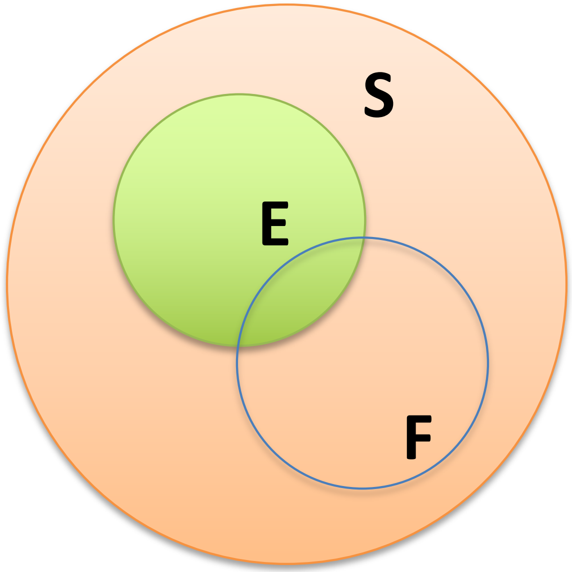
Now, let's visualize this scenario using a Venn diagram:



where S is the sample space, and we have labelled two events *E* and *F*.

CONDITIONAL PROBABILITY

Now the probability of E occurring is the probability of the outcomes in *E*.



If we are dealing with finite sets with equal likelihood, this probability is

$$\frac{n(E)}{n(S)}$$

where *n(E)* is the number of ways that the event *E* can occur, and *n(S)* the same for *S*.

CONDITIONAL PROBABILITY

If, however, we are seeking to calculate P(E | F), then the ratio of n(E) to n(S) becomes irrelevant:
Given that event F has already occurred, our focus shifts to determining the number of occurrences where both E and F happen simultaneously. However, in this context, our effective sample space is constrained to just F.

This gives

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

For discrete equal-likelihood sample spaces, this is:

$$P(E \mid F) = \frac{n(E \cap F)\, n(S)}{n(S) n(F)} = \frac{n(E \cap F)}{n(F)}$$

CONDITIONAL PROBABILITY

If we look back to our dice example, this is exactly what we did.

Example revisited. we started from

$$S = \{(i, j),\ i = 1, 2, 3, 4, 5, 6,\ j = 1, 2, 3, 4, 5, 6\},$$

and asked, "*Suppose the first die comes down with a three, what is the probability that the sum of the two dice equals eight*"?

In the language we've just used, our event E is

$E=\{(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)\}$,

the event

$F = \{(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)\}$.

So

$$E \cap F= \{(3, 5)\}.$$

CONDITIONAL PROBABILITY

Looking at the size of these events (all individual outcomes are equally likely), we have

$$P(E \mid F) = \frac{n(E \cap F)\, n(S)}{n(S) n(F)} = \frac{n(E \cap F)}{n(F)} = \frac{1}{6}$$

CONDITIONAL PROBABILITY

Definition: the conditional probability of E given F is

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

combined with the standard properties of probabilities this implies

- $P(\bar{E} \mid F) = 1 - \frac{P(E \cap F)}{P(F)}$

- if E and F are mutually exclusive, then $P(E|F) = 0$ because $E \cap F = \emptyset$

MULTIPLICATIVE RULE OF PROBABILITIES

The expression for conditional probabilities implies that

$$P(E \cap F) = P(E \mid F)P(F) \qquad (1)$$

In simple terms, this equation states that the probability of both events E and F occurring is equal to the probability of event F occurring multiplied by the probability of event E occurring given that F has already occurred.

This can be generalized to the case of many events E_i

$$P(E_1 \cap E_2 \cap E_3 \cap \cdots \cap E_n) = P(E_1) \cdot P(E_2 \mid E_1) \cdot P(E_3 \mid E_1 \cap E_2) \cdots P(E_n \mid E_1 \cap \cdots \cap E_{n-1})$$

MULTIPLICATIVE RULE OF PROBABILITIES

To demonstrate this relation, let's substitute the definition of conditional probability into the right side of the generalized equation:

$$P\,(E_1 \cap E_2 \cap E_3 \cap \cdots \cap E_n\,) = P\,(E_1\,) \cdot P\,(E_2 \mid E_1\,) \cdot P\,(E_3 \mid E_1 \cap E_2\,)$$
$$\cdots P\,(E_n \mid E_1 \cap \cdots \cap E_{n-1}\,)$$

Using the definition of conditional probability, $P(E \cap F) = P(E \mid F)\;\; P(F)$, we can rewrite each conditional probability term in the equation above:

$$\cancel{P(E_1)}\frac{\cancel{P(E_1 \cap E_2)}}{\cancel{P(E_1)}} \cdot \frac{\cancel{P(E_1 \cap E_2 \cap E_3)}}{\cancel{P(E_1 \cap E_2)}} \cdots \frac{P(E_1 \cap E_2 \cdots \cap E_n)}{\cancel{P(E_1 \cap E_2 \cdots \cap E_{n-1})}}$$

each internal term of the numerator cancels with the next term in the denominator.

MULTIPLICATIVE RULE OF PROBABILITIES

EXAMPLE: Let's revisit the scenario of selecting balls from a bowl, as discussed in lecture 1, where there were 6 black balls and 5 white ones.

How many ways exist to select all the balls from the bowl?

In this case, the number of ways can be calculated using permutations, resulting in N=11! (11 factorial).

To elaborate, we can approach the calculation as follows:
Since each permutation is equally likely, we can compute the probability of obtaining a specific selection, which would be 1/N.

Therefore, we determine N by taking the inverse of the probability.

MULTIPLICATIVE RULE OF PROBABILITIES

If we focus on the specific selection:

{(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)}

where each ball is numbered accordingly, we can define

E_n='the nth ball is ball n'

$$P(E_1 \cap E_2 \cap E_3 \cap \cdots \cap E_n) = P(E_1) \cdot P(E_2 \mid E_1) \cdot P(E_3 \mid E_1 \cap E_2) \\ \cdots P(E_n \mid E_1 \cap \cdots \cap E_{n-1})$$

And we'd get

$$P(E_1 \cap E_2 \cap E_3 \cap \cdots \cap E_n) = 1/11 \cdot 1/10 \cdot 1/9 \cdots 1/3 \cdot 1/2 \cdot 1/1) = 1/11! = |E_1 \cap E_2 \\ \cap E_3 \cap \cdots \cap E_n|/N = 1/N \implies N = 11!$$

TREE DIAGRAMS

- Tree diagrams are a useful way of keeping track of the progress of compound experiments.

- They can also be seen to work using the multiplicative rule of probabilities.

- Let's analyse throwing 3 coins sequentially.

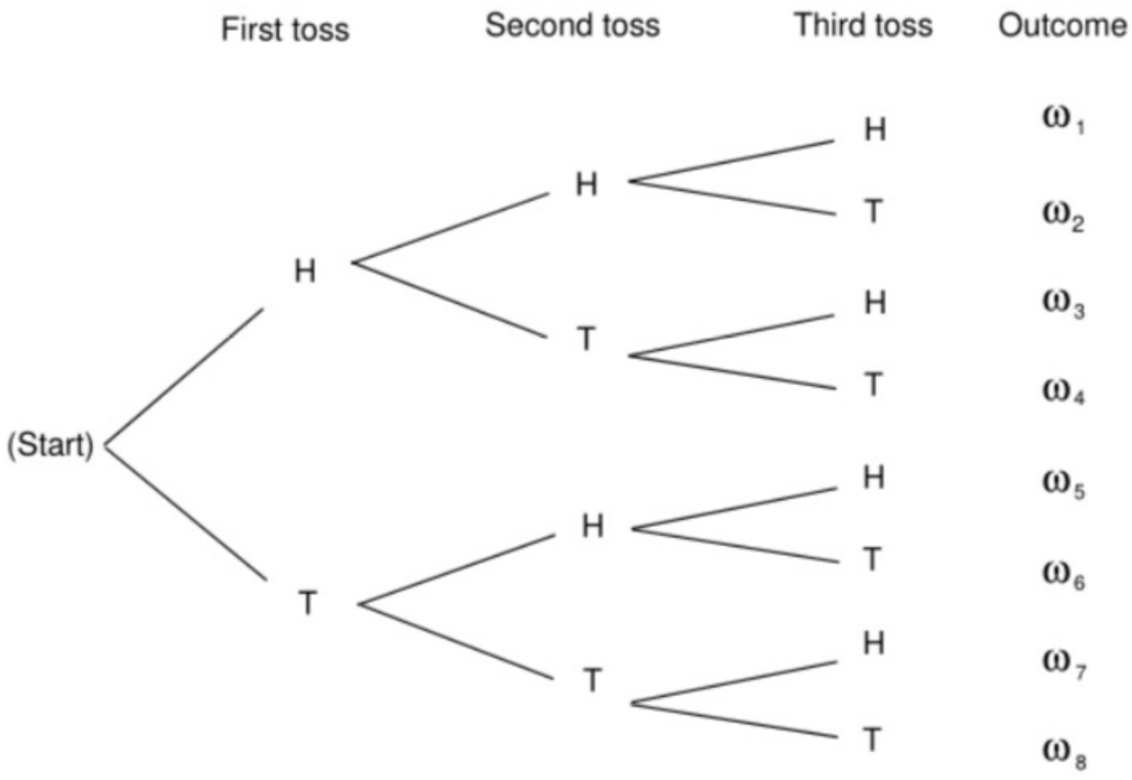
TREE DIAGRAMS

Tree diagrams are visual representations that help organize and illustrate the outcomes of compound experiments, especially when these experiments involve multiple stages or steps. They provide a clear and structured way of understanding the possible outcomes and the probabilities associated with each outcome.

In the context of probability theory, the multiplicative rule states that the probability of two independent events occurring together is the product of their individual probabilities. Tree diagrams are particularly useful for applying this rule because they help break down the compound experiment into its individual stages, making it easier to calculate probabilities at each step and then multiply them together to find the overall probability of a specific outcome.

For example, let's consider the experiment of tossing 3 coins sequentially.

TREE DIAGRAMS



Tree diagram for three tosses of a coin.

A tree diagram for this experiment would have three levels, each corresponding to a toss of a coin. At each level, there are two branches representing the two possible outcomes of a coin toss: heads (H) or tails (T). By following the branches of the tree diagram, we can systematically enumerate all possible sequences of coin toss outcomes.

TREE DIAGRAMS

Elaborating on how tree diagrams work with the multiplicative rule, at each stage of the experiment represented by a level in the tree diagram, we calculate the probability of each outcome (heads or tails) based on the probability of a single coin toss (which is 1/2 for a fair coin).

Then, we multiply these probabilities along the branches to find the probability of each sequence of outcomes. Finally, we sum up the probabilities of all sequences that lead to the event of interest to determine its overall probability.

TREE DIAGRAMS

The overall probabilities to arrive at one outcome is obtained by multiplying the probabilities at each stage of the experiment.

This works because at each stage in the compound experiment, the result doesn't depend on the result further down the chain. So, to get

$$P(\omega_1) = P(\{(H,H,H)\}) = P(\{H\}) \cdot P(\{(H,H)\} \mid \{H\}) \cdot P(\{(H,H,H)\} \mid \{(H,H)\})$$
$$= P(\{\{H\}\}) \cdot P(\{\{H\}\}) \cdot P(\{\{H\}\}) = p^3$$

because each probability P ({(H, H, H)} | {(H, H)}) only depends on the current coin flip, not anything else.

In other words, the individual parts of the compound experiment are independent.

EXAMPLE

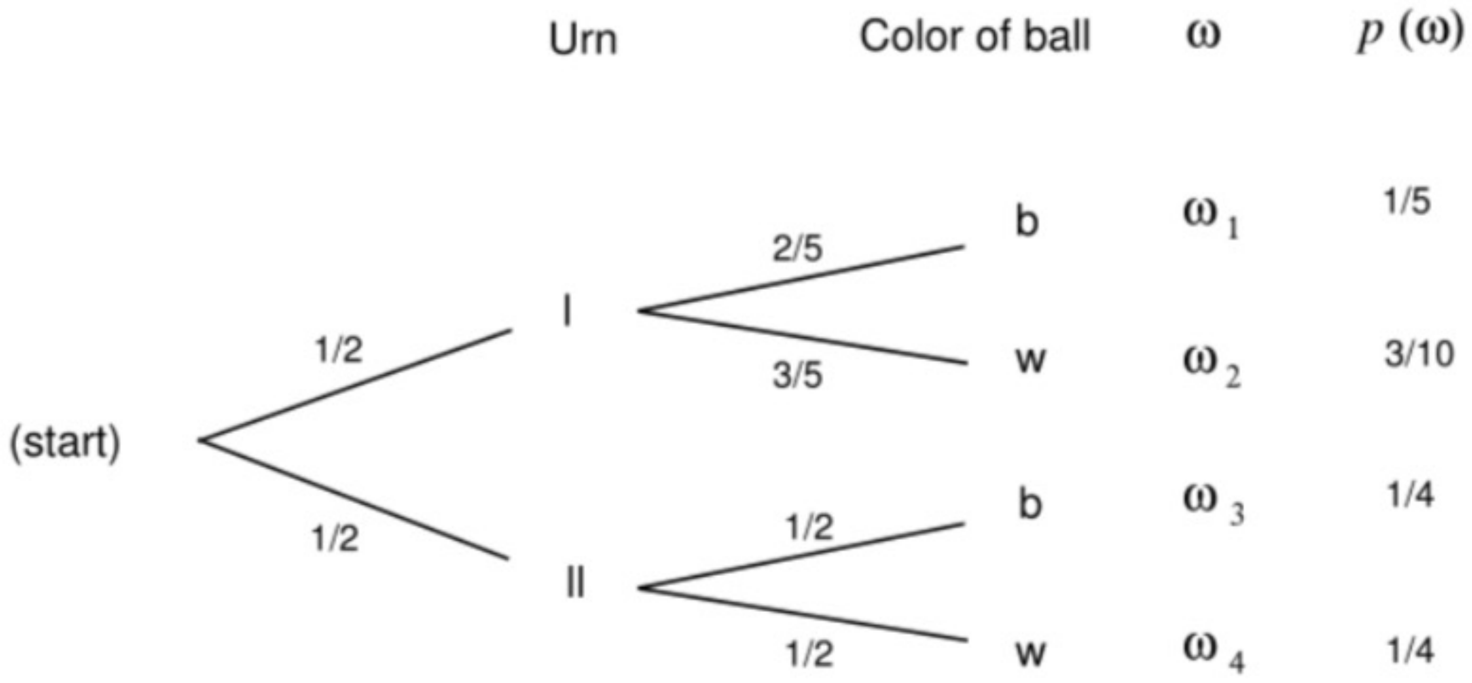
Consider two urns, labeled Urn I and Urn II.

- Urn I contains 2 black balls and 3 white balls.
- Urn II contains 1 black ball and 1 white ball.

Now, let's suppose that an urn is selected at random from the two available urns, and then a ball is chosen randomly from the selected urn.

TREE DIAGRAMS

We can present the sample space of this experiment as the paths through a tree as shown

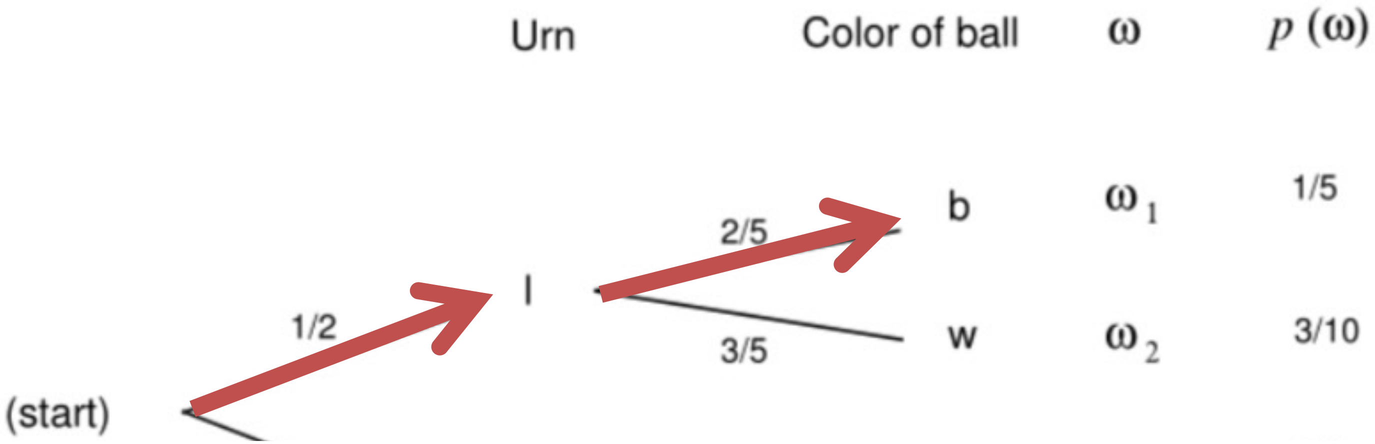


Tree diagram.

The probabilities assigned to the paths are also shown.

TREE DIAGRAMS

Let B = "a black ball is drawn," and *I= "urn I is chosen."*



Then the branch weight 2/5, which is shown on one branch in the figure, can now be interpreted as the conditional probability

P (B | I) .

If there were more branching's, we would have to use the multiplicative rule of probability.

EXAMPLE - BACK TO THE URNS AND BALLS

Suppose that we wanted to work out ***P (I / B)*** *?*

In words
"What is the probability that the ball came from urn I, given that the ball is black."

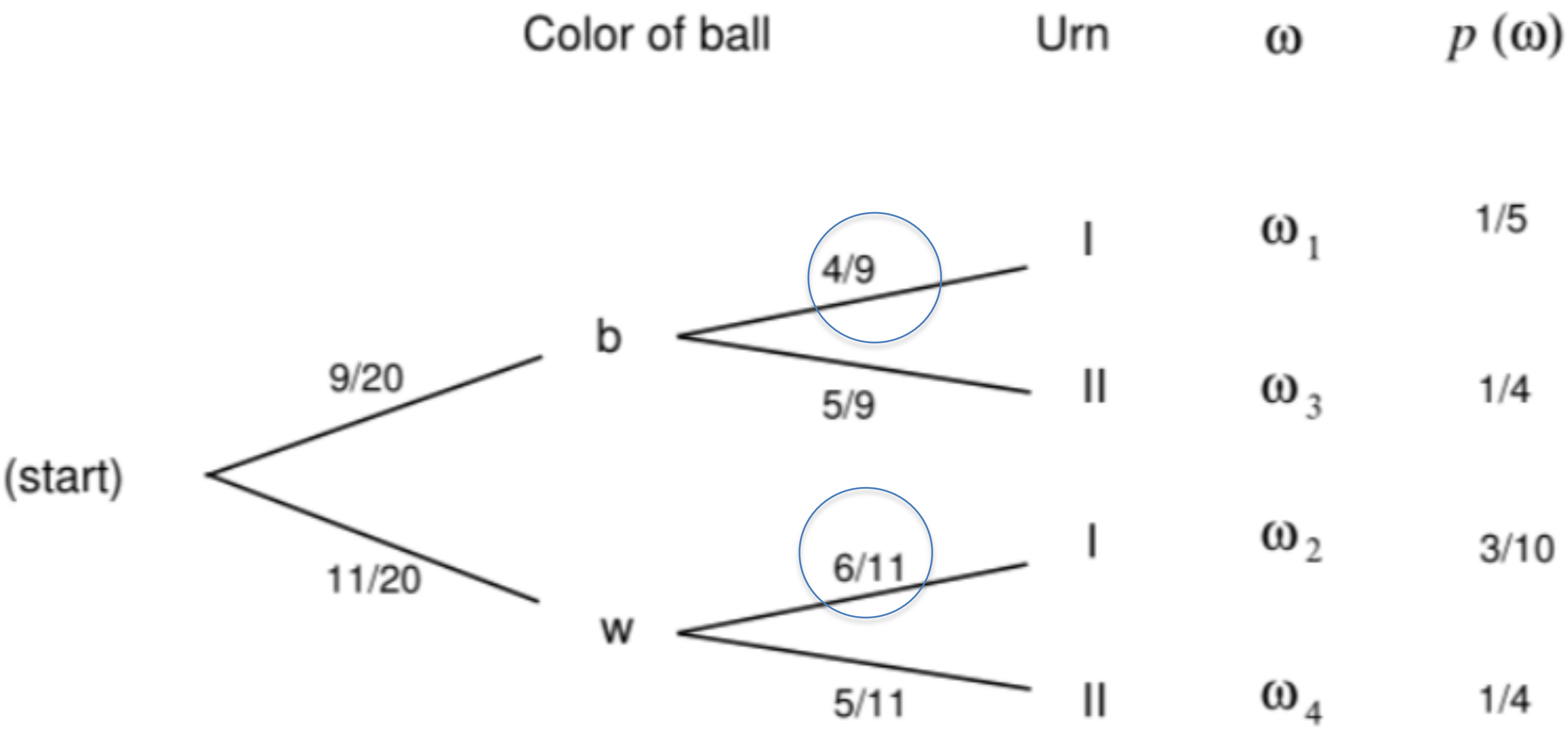
We can do this as follows

$$P(I \mid B) = \frac{P(I \cap B)}{P(B)} = \frac{P(I \cap B)}{P(B \cap I) + P(B \cap II)} = \frac{\frac{1}{5}}{\frac{1}{5} + \frac{1}{4}} = \frac{4}{9}$$

$$P(I|W) = \frac{P(I \cap W)}{P(W)} = \frac{P(I \cap W)}{P(W \cap I) + P(W \cap II)} = \frac{\frac{3}{10}}{\frac{3}{10} + \frac{1}{4}} = \frac{6}{11}$$

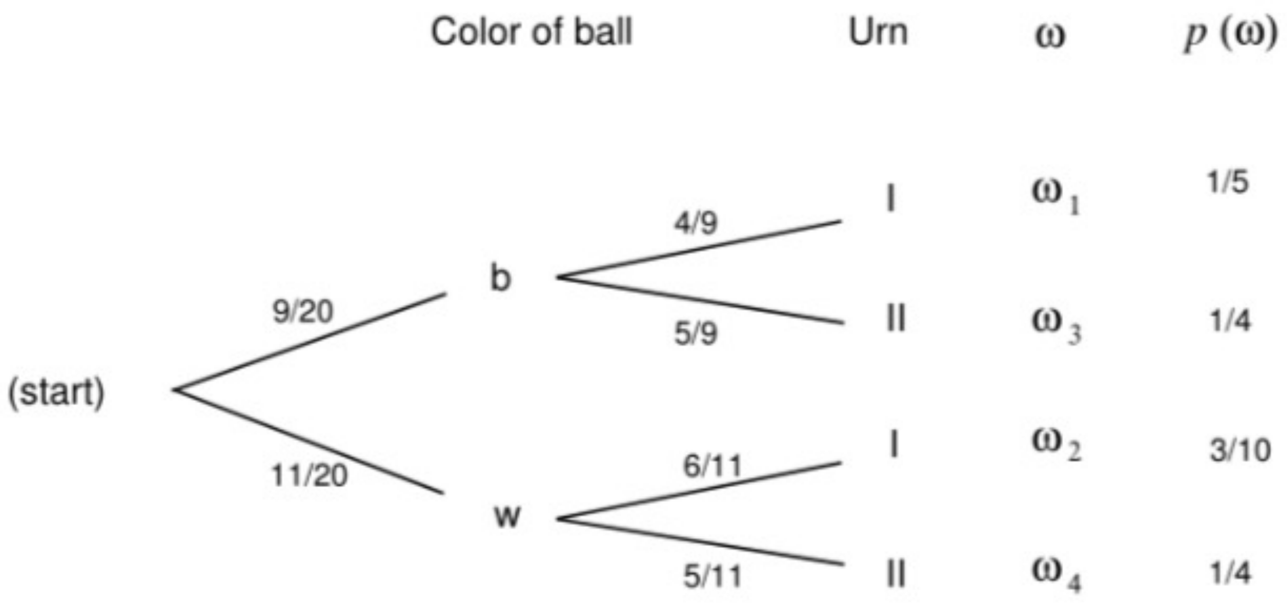
EXAMPLE - BACK TO THE URNS AND BALLS

We can repeat this for other conditional probabilities and use it to construct a reverse tree diagram



Reverse tree diagram.

EXAMPLE - BACK TO THE URNS AND BALLS



Reverse tree diagram.

These '*reversed*' conditional probabilities are often associated with the name **Bayes.**

We'll look at a formula that bears his name shortly that can be used to systematically calculate these reversed probabilities.

LAW OF TOTAL PROBABILITY

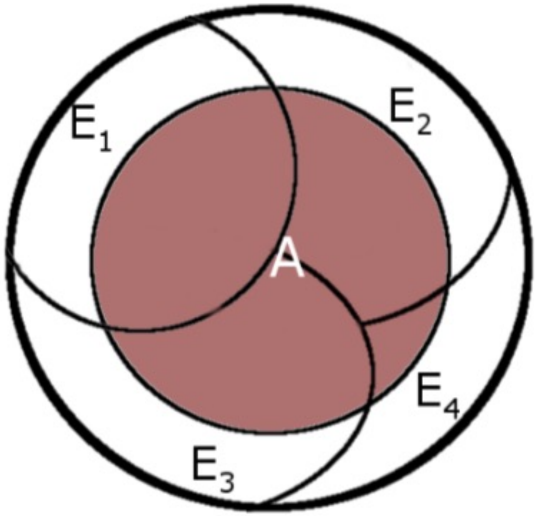
We can split S up into an exhaustive collection of mutually exclusive events E_i such that

$$\bigcup_{i=1}^{n} E_i = S$$

remembering that mutually exclusive means that

$$E_i \cap E_j = \emptyset, i \neq j.$$

In Venn diagram form this collection of events could look like this

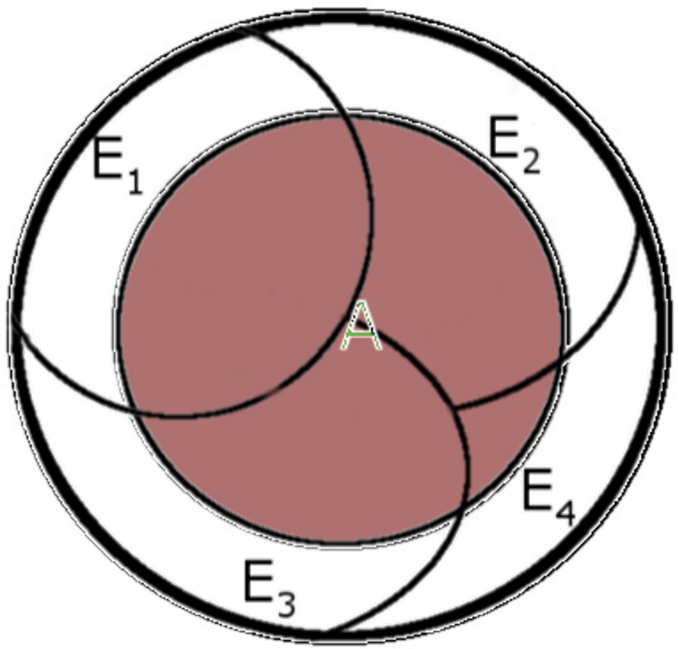


where $A \subset S$.

LAW OF TOTAL PROBABILITY

we can then use the E_i (a bit like basis vectors in linear algebra) to decompose an event A into its intersections with the E_i.

$$A = \bigcup_{i=1}^{n} A \cap E_i$$



LAW OF TOTAL PROBABILITY

From the conditional probability formula

$$P(A \mid E) = \frac{P(A \cap E)}{P(E)}$$

Then, it follows that the general form of the law of total probability is given by the following expression

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(A \mid E_i) P(E_i)$$

BAYES' FORMULA

Using the law of total probability in the previous slides

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(A \mid E_i) P(E_i)$$

Rev. Thomas Bayes (1701-1761)

We can now define the so-called **Bayes'formula**

$$P(E_i \mid A) = \frac{P(A \cap E_i)}{P(A)} = \frac{P(A \mid E_i)P(E_i)}{\sum_{i=j}^{n} P(A \mid E_j)P(E_j)}$$

Pierre-Simon Laplace (1749-1827)

Bayes' solution to the problem of inverse probability was presented in "*An Essay Towards Solving a Problem in the Doctrine of Chances*," which was read to the Royal Society in 1763 following Bayes' death. This solution was further developed by Pierre-Simon Laplace, who first published the modern formulation in 1812. For a detailed account of the history of the Bayes formula, one can refer to the essay by E.T. Jaynes, "Bayesian Methods: General Background." (https://bayes.wustl.edu/etj/articles/general.background.pdf)

BAYES’ FORMULA

If we look at the formula, we see that the two types of conditional probabilities have swapped events on the left and right hand sides.

$$P(E_i \mid A) = \frac{P(A \mid E_i)P(E_i)}{P(A)}$$

BAYES' FORMULA

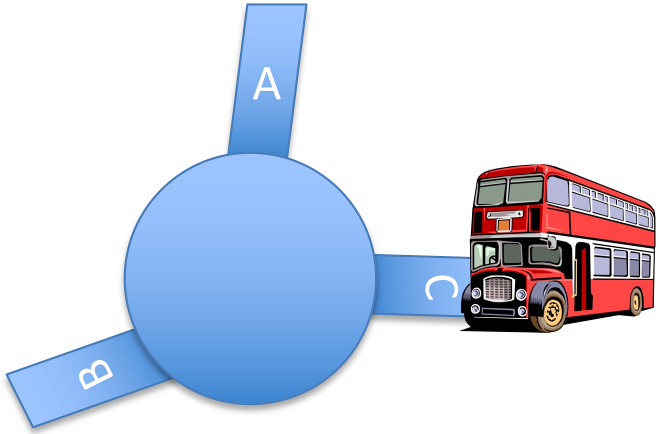
Bayes' Formula can be regarded as a method to revise our initial beliefs about a problem as new data becomes available. From this perspective:

- The E_i represents our prior beliefs or opinions about the likelihood of various events.
- These beliefs are then adjusted or updated based on the occurrence of event A.

This approach can also be applied to data-fitting tasks. In this context, Bayesian inference stands as the other primary approach, known for its subjective interpretation within probability theory.

BAYES' FORMULA: AN EXAMPLE

A town has three bus routes, A,B and C. *During rush hour there are twice as many buses on the A route as on B or C.*



Over a period of time, it has been observed that at a crossroads, where the routes converge, the buses run more than 5 minutes late 1/2 , 1/5 , 1/10 of the time.

If an inspector at the crossroads finds that the first bus, he sees is more than five minutes late, what is the chance that it is a route B bus?

BAYES' FORMULA: AN EXAMPLE

We don't have complete information about this problem (like all possible outcomes etc).

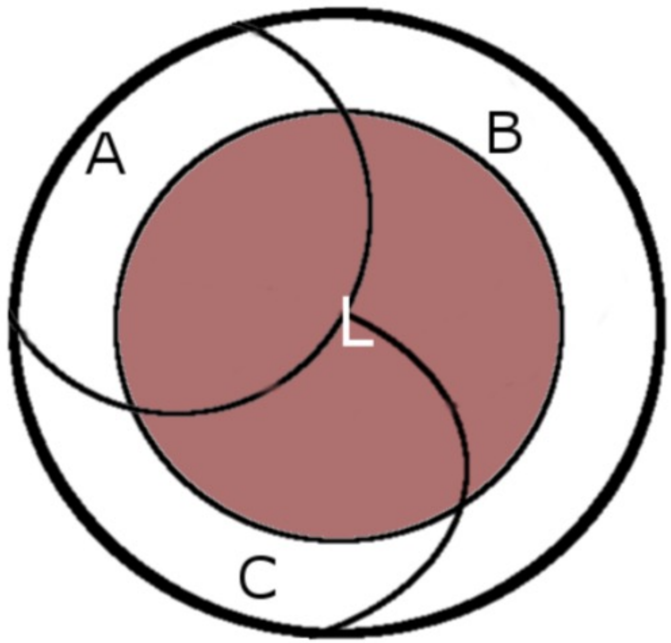
We can form an exhaustive partition of the experiment into the mutually exhaustive events A, B, C that the inspector sees a bus on that route first. And we have the event <u>**L** that the bus is late.</u>

We see that the problem looks like our Bayes' Formula situation: we know the probabilities

P (L | A), P (L | B), P (L | C) ,

but we want to know

P (B | L) .



BAYES' FORMULA: AN EXAMPLE

We require **P (B | L).** Using Bayes' formula and the law of the total probability, we get that

$$P(B \mid L) = \frac{P(B) \cdot P(L \mid B)}{P(A) \cdot P(L \mid A) + P(B) \cdot P(L \mid B) + P(C) \cdot P(L \mid C)}$$

From the information we have ("*during rush hour there are twice as many buses on the A route as on B or C*") we can work out that

P(A) = 1/2 and **P(B)=P(C)=1/4**

BAYES' FORMULA: AN EXAMPLE

Also, we are given that :

Over some time, it has been observed that at a crossroads, where the routes converge, the buses run more than 5 minutes late 1/2 , 1/5 , 1/10 of the time.

That means:

$P(L \mid A) = 1/2$,
$P(L \mid B) = 1/5$,
$P(L \mid C) = 1/10$

Plugging these numbers in the formula gives $P(B \mid L) = 2/13$

BAYES’ FORMULA: ANOTHER EXAMPLE

Bayes' rule applies not only to probabilistic scenarios but also to situations involving uncertainty regarding the value of a measured quantity.

For instance, let's consider an example from J.V. Stone's work, "*Bayes’ Rule: A Tutorial Introduction to Bayesian Analysis*," which aims to determine the probability of correctly understanding the meaning of a spoken conversation.

Before delving into this example, it might be helpful to provide some context, such as the famous gag from The Two Ronnies...

BAYES' FORMULA: ANOTHER EXAMPLE

In any spoken language, regional variations in pronunciation can lead to humorous misunderstandings, as famously depicted in this sketch of *The Two Ronnies*.

Now, let's explore how Bayes' Rule can be employed to calculate the probability of providing an incorrect answer in such situations.

BAYES' FORMULA: ANOTHER EXAMPLE

For candles?
or
Fork handles?

Fork 'andles

BAYES' FORMULA: ANOTHER EXAMPLE

Likelihood

P(acoustic data| four candles)=0.6

P(acoustic data| fork handles)=0.7

This give the probabilities of acoustic data given the two possible phrases.
In this case it is almost identical.

Each likelihood answer to the question:

What is the probability of the observed acoustic data given that each of two possible phrases was spoken?

Posterior probability

P(four candles|data)

P(fork handles|data)

This give the probability of each phrase given the acoustic data

BAYES' FORMULA: ANOTHER EXAMPLE

Prior probability

Let's suppose that the assistant has been asked for four candle a total of 90 times
In the past, whereas he has been asked for fork handles only 10 times.

Therefore, the assistant estimates that the probability for a next client asking for one of the two items are

$$\left.\begin{array}{l} P(\text{four candles})=\frac{90}{100}=0.9 \\ P(\text{fork handles})=\frac{10}{100}=0.1 \end{array}\right\}$$

This give the prior knowledge of the shop assistant
Based on his previous experience of what customers say

BAYES' FORMULA: ANOTHER EXAMPLE

INFERENCE : *A conclusion reached on the basis of evidence and reasoning.*

Bayes' inference engine

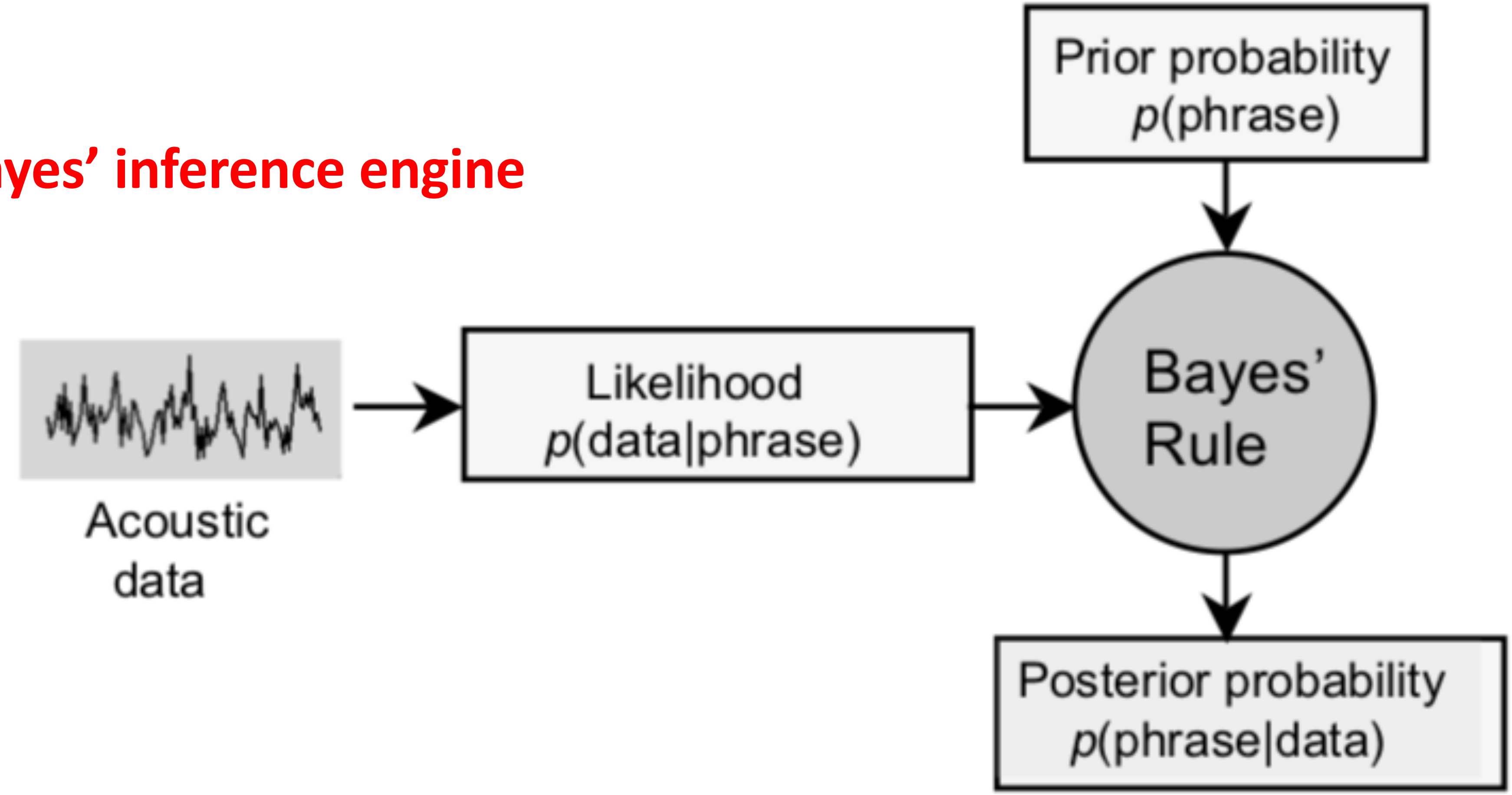


Figure source: J.V. Stone. Bayes' Rule: a tutorial introduction to Bayesian analysis. Sebtel Press.

BAYES' FORMULA: ANOTHER EXAMPLE

likelihood

Prior Probability

$$P(E_i \mid A) = \frac{P(A \mid E_i)P(E_i)}{P(A)}$$

Posterior probability

Marginal probability

$$p(\text{four candles}|\text{data}) = \frac{p(\text{data}|\text{four candles})p(\text{four candles})}{p(\text{data})}$$

$$p(\text{fork handles}|\text{data}) = \frac{p(\text{data}|\text{fork handles})p(\text{fork handles})}{p(\text{data})},$$

BAYES' FORMULA: ANOTHER EXAMPLE

Using the calculate probability we can estimate the posterior probability

$$
\begin{aligned}
p(\text{four candles}|\text{data}) &= p(\text{data}|\text{four candles})p(\text{four candles})/p(\text{data}) \\
&= 0.6 \times 0.9/0.61 = 0.885, \\
p(\text{fork handles}|\text{data}) &= p(\text{data}|\text{fork handles})p(\text{fork handles})/p(\text{data}) \\
&= 0.7 \times 0.1/0.61 = 0.115.
\end{aligned}
$$

$$
\begin{aligned}
p(\theta_c|x) &= p(x|\theta_c)p(\theta_c)/p(x) = 0.885 \\
p(\theta_h|x) &= p(x|\theta_h)p(\theta_h)/p(x) = 0.115.
\end{aligned}
$$

$$
\begin{aligned}
x &= \text{acoustic data}, \\
\theta_c &= \text{four candles}, \\
\theta_h &= \text{fork handles},
\end{aligned}
$$

BAYES' FORMULA: ANOTHER EXAMPLE

SUMMARY

x = acoustic data,

θ_c = four candles,

θ_h = fork handles,

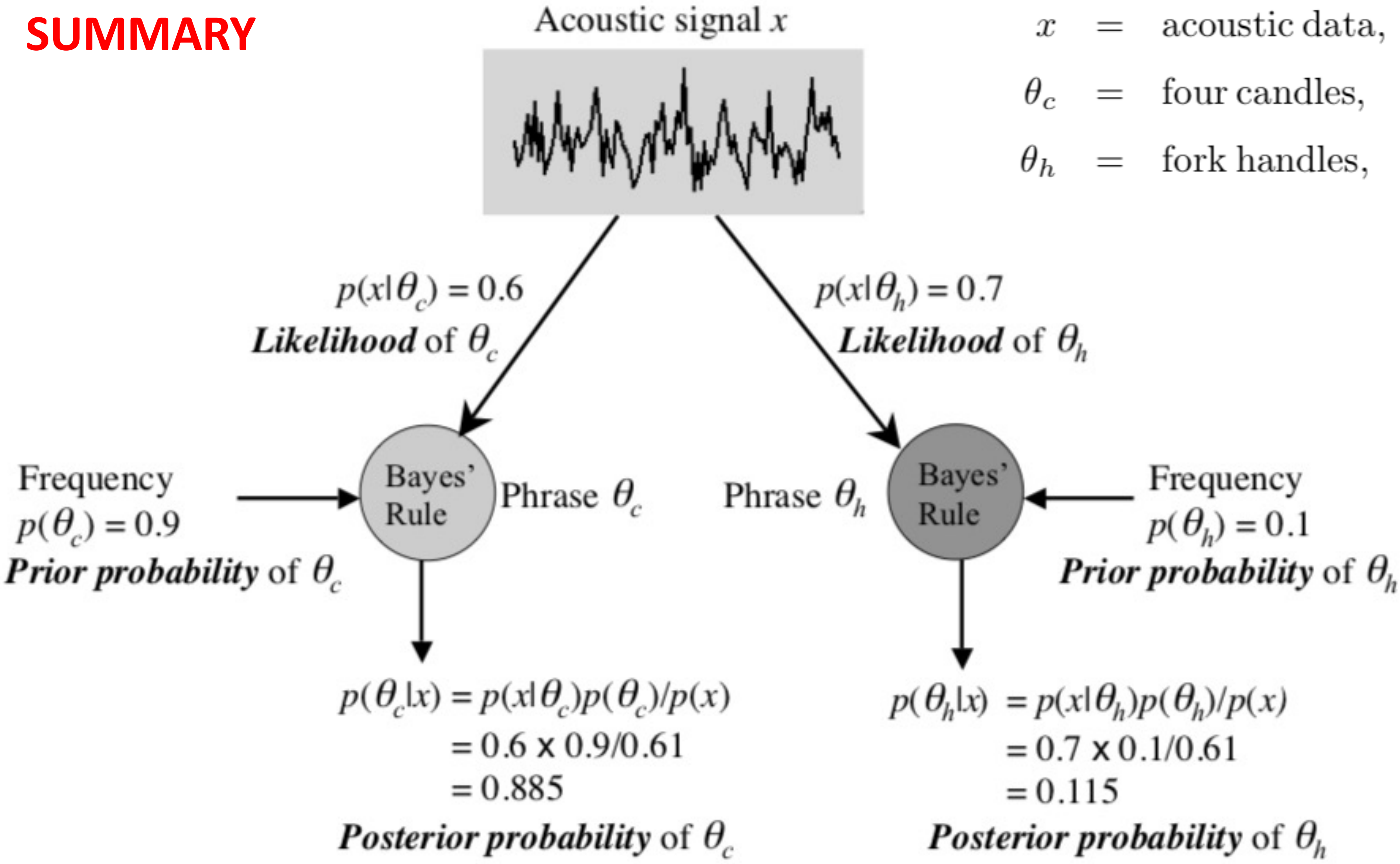


Figure source: J.V. Stone. Bayes' Rule: a tutorial introduction to Bayesian analysis. Sebtel Press.

CONDITIONAL PROBABILITIES ARE PROBABILITIES

It is important to note that conditional probabilities obey all the axioms we established to call P (E) a probability.

For the statement above to be true our conditional probabilities must satisfy the axioms of probabilities:

- $0 \leq P(E \mid F) \leq 1$
- $P(S) = 1$
- if $E_i, i = 1 \ldots n$ are mutually exclusive events then

$$P\left(\bigcup_{i-1}^{n} E_i \mid F\right) = \sum_{i=1}^{n} P(E_i \mid F)$$

We are not going to prove these results here, but you can find a proof in textbook.

SUMMARY OF CONDITIONAL PROBABILITIES

- **Definition of conditional probabilities**

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

- **Multiplicative rule**

$$P(E_1 E_2 E_3 \ldots E_n) = P(E_1) \frac{P(E_1 \cap E_2)}{P(E_1)} \frac{P(E_1 \cap E_2 \cap E_3)}{P(E_1 \cap P(E_2)} \cdots \frac{P(E_1 \cap E_2 \cdots \cap E_n)}{P(E_1 \cap E_2 \cdots \cap E_{n-1})}$$

SUMMARY OF CONDITIONAL PROBABILITIES

- Law of total probability

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(A \mid E_i) P(E_i)$$

- Bayes' Formula

$$P(E_i \mid A) = \frac{P(A \cap E_i)}{P(A)} = \frac{P(A \mid E_i) P(E_i)}{\sum_{j=1}^{n} P(A \mid E_j) P(E_j)}$$

INDEPENDENT EVENTS

In the previous lesson, we discussed examples of experiments involving independent events.

Now that we have introduced conditional probability, let's delve deeper into this concept.

INDEPENDENT EVENTS

Two experiments are considered independent if the outcome of one experiment has no impact on the possible outcomes of the other.

Similarly, two events (E, F) are termed independent if the occurrence of one event does not influence the probability of the other event happening.

INDEPENDENT EVENTS

Mathematically, this independence can be expressed through conditional probabilities:

- For event E, the probability of E occurring is equal to the conditional probability of E given that event F has occurred, and also equal to the conditional probability of E given that event F has not occurred. This is represented as: $P(E) = P(E|F) = P(E|\neg F)$

- Similarly, for event F, the probability of F occurring is equal to the conditional probability of F given that event E has occurred, and also equal to the conditional probability of F given that event E has not occurred. This is represented as: $P(F) = P(F|E) = P(F|\neg E)$

INDEPENDENT EVENTS

Another way to interpret this is that the probability of both events E and F occurring (denoted as $P(E \cap F)$) is simply the product of the probabilities of each event occurring:

$$P(E \cap F) = P(E) \ P(F)$$

In simpler terms, if events E and F are independent, then the likelihood of both events happening simultaneously is just the product of the likelihood of each event happening independently.

INDEPENDENT EVENTS

Theorem: *Two events (E, F) are independent if and only if of their intersection (E ∩ F) equals the product of their probabilities:*

$$P(E \cap F) = P(E)P(F)$$

INDEPENDENT EVENTS

However, when dealing with more than two events, the notion of independence becomes more restrictive.

Theorem: *a set of events E_1 , E_2 , E_3 , $\cdots E_n$ are considered mutually independent if for every* subset if for every subset E_1', E_2', E_3',… E_r', $r \leq n$
the probability of the intersection of all events in the subset equals the product of the probabilities of each individual event in the subset:

$$P(E_1' \cap E_2' \cap E_3' \cap \cdots \cap E_r') = P(E_1') \cdot P(E_2') \cdot P(E_3') \cdots P(E_r')$$

INDEPENDENT EVENTS

The concept of independent processes holds significant importance as we progress further.

We frequently rely on the notion that repetitions of experiments constitute independent processes.

It's essential to note that this assumption aligns closely with the frequentist definition of probability.

INDEPENDENT EVENTS: EXAMPLE

Consider the compound experiment of throwing two fair coins.
The sample space is

S= {(H, H), (H, T), (T , H), (T , T)}.

Define two events

A = 'the first coin is a head' = {(H, H), (H, T)}
B = 'the second coin is a tail' = {(H, T), (T , T)}

And P (A)= |A|/|S| = 2/4 = ½ and P (B)= |B|/|S| = 2/4 = 1/2

Now P (A ∩ B)= P (H, T) = |A ∩ B|/|S| = 1/4 = P (A) ∗ P (B).

so A and B are independent.

INDEPENDENT EVENTS: EXAMPLE

But consider C = 'both coins are heads' = {(H, H)}

P (C) = |C|/|S| = 1/4

now

$P(B \cap C) = P(\emptyset) = |B \cap C|/|S| = 0/4 \neq P(A) * P(C) = 1/2 * 1/4$,

so B and C are not independent. A is also not independent of C .

In general, it is possible for all pairs of events to be independent, but the complete set of events not to be.

QUESTIONS

?