UNIVERSITY OF LINCOLN

SCHOOL OF MATHEMATICS AND PHYSICS

MTH 1005M
PROBABILITY AND STATISTICS

Semester B

(9/2/2023)

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 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,

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

From this function we could 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 idea of conditional probability allows us to calculate probabilities when we have partial information about the system, or to reassess likelihoods when we receive new information.

A conditional probability is the probability of some event *E*, given that another event *F* has occurred.

The probability of *E*, given that *F* has occurred is written

$$P\,(E \mid F\,)$$

CONDITIONAL PROBABILITY

Example: Consider rolling two dice.

We have a sample space

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

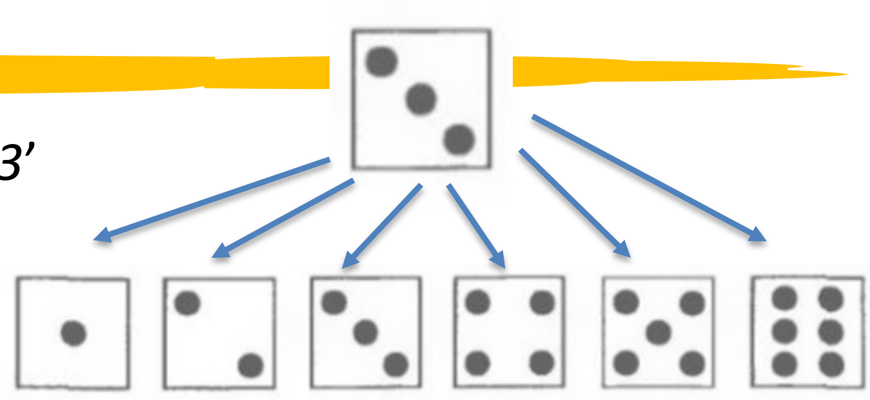
where say the outcome (i, j) is the first die lands with *i* dots up and the second die with j.
We assume the dice are fair - we assign a probability (distribution) function of

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

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)}.

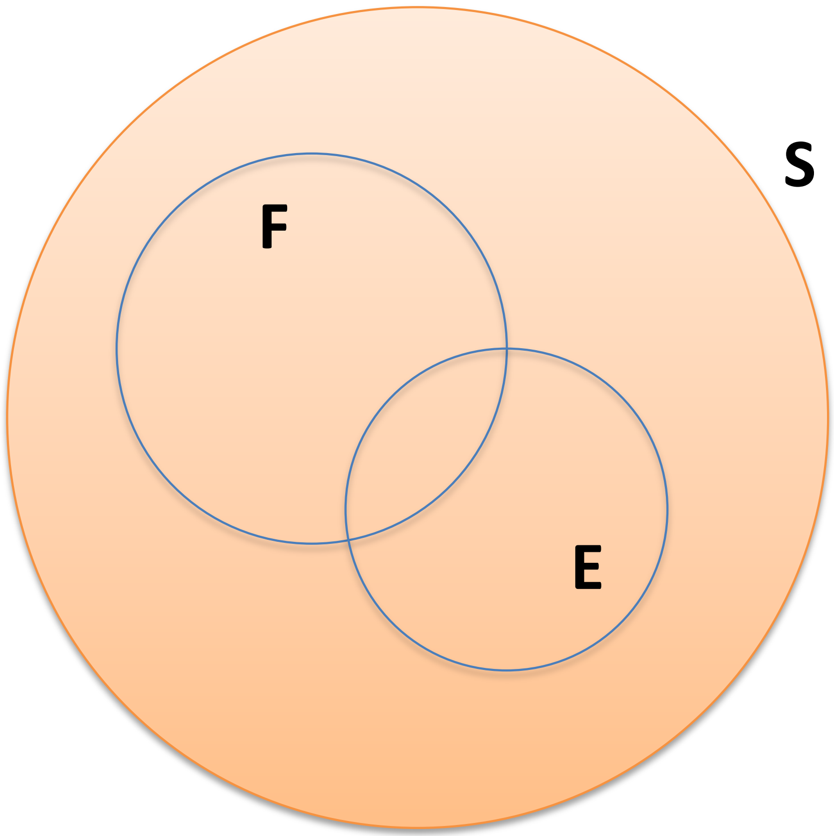
The only outcome that satisfies the criteria is {(3, 5)} and once we know that F has occurred, it has a one in 6 chance that {(3, 5)} will occur.

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

Note also that {(3, 5)}= E ∩ F

CONDITIONAL PROBABILITY

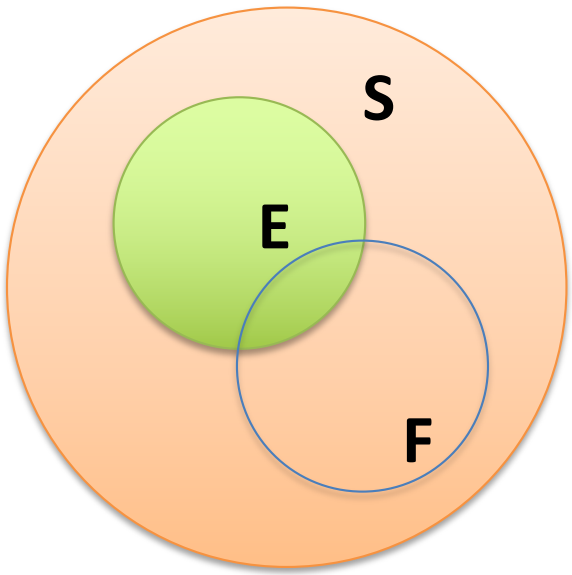
The probability of the outcome (3,5) before the first die was rolled was 1/36. After the first die comes up a 3, we've argued that the probability should be 1/6. Lets look at this as a Venn diagram:



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 instead we are looking for P (E | F) then it is not the ratio n(E) to n(S) that matters

we know that event F has occurred - so now we want the number of ways that both E and F can occur, but now our effective sample space is 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 words: the probability that E and F occur is the probability that F occurs multiplied by the probability that E occurs given that F has occurred.

This can be generalised to the case of many events E_i

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

To prove this relation, we can inserte the definition (1) of conditional probability in the right side of the equation, giving

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

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

MULTIPLICATIVE RULE OF PROBABILITIES

EXAMPLE: If you remember the examples of taking balls out of a bowl in lecture 1 - there were 6 black and 5 white ones.

How many ways are there to select all the balls from the bowl?

In this case there are $11 \cdot 10 \cdot 9 \cdots 3 \cdot 2 \cdot 1 = 11!$

We could calculate this as follows:
Each permutation is equally likely, so we can calculate the probability of getting a particular selection, which would be 1/N , so we find N by the inverse of the probability.

MULTIPLICATIVE RULE OF PROBABILITIES

If we look for the particular selection
{(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)} ,
where we've numbered the balls, then

We define E_n=*'the nth ball is ball n'*

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

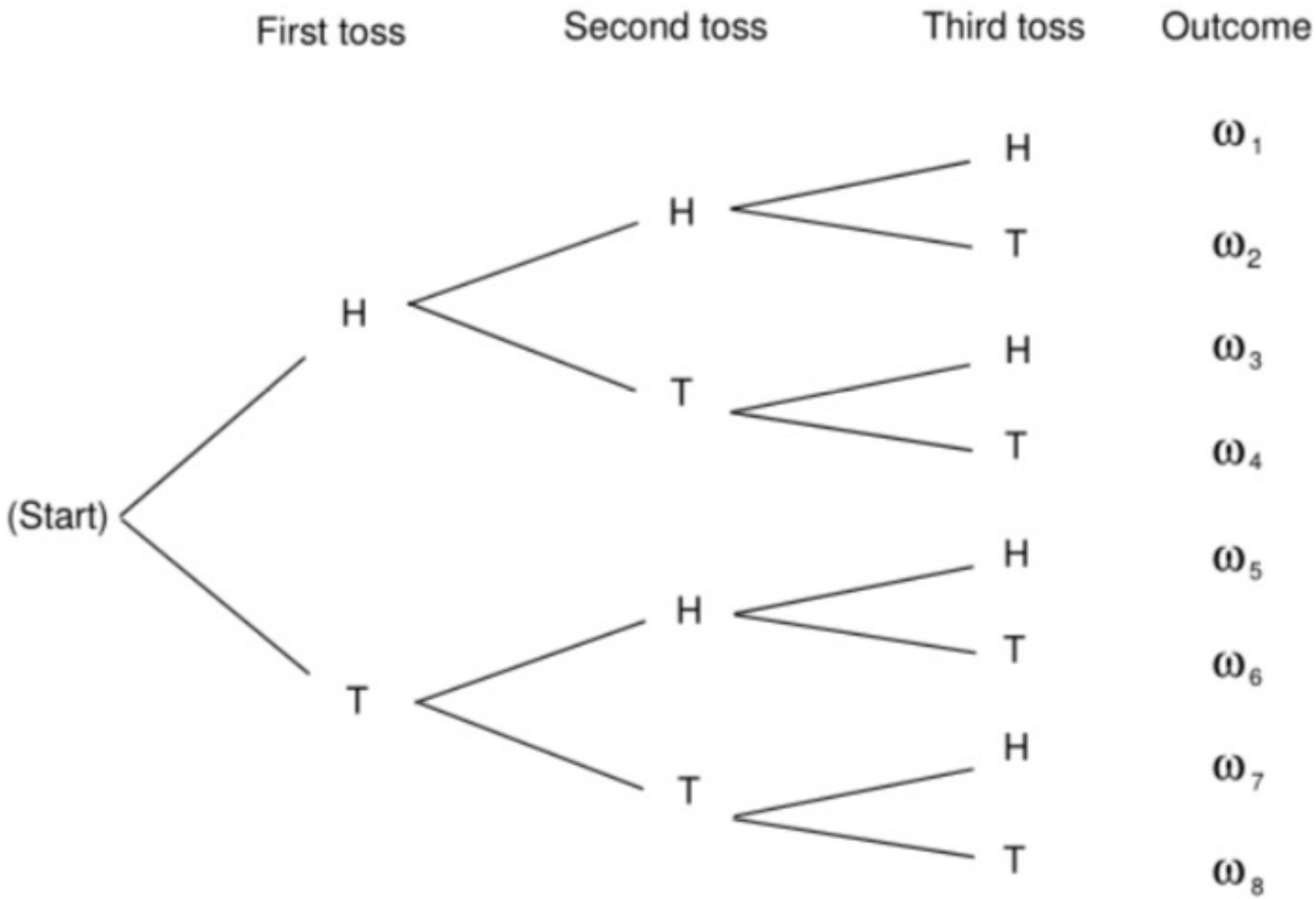
And we'd get

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

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 diagram for three tosses of a coin.

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

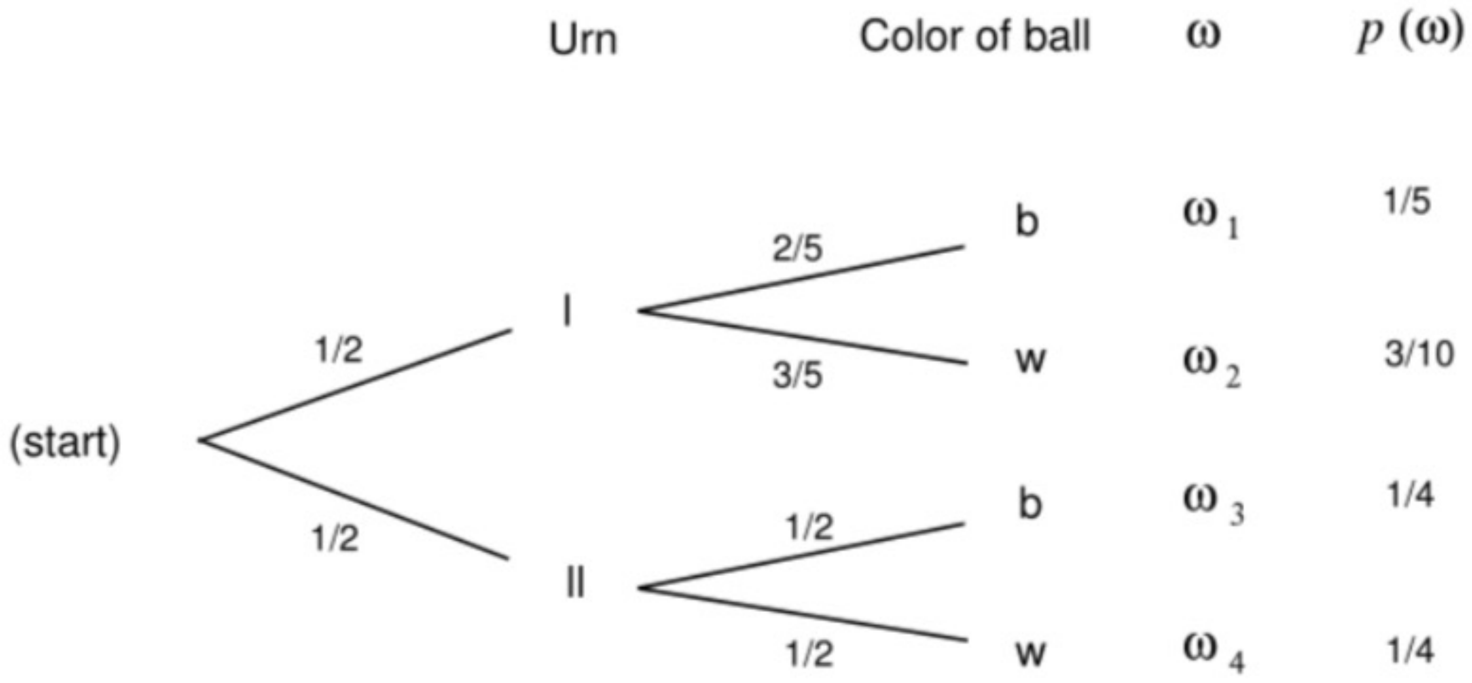
We have two urns, I and II.

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

An urn is selected at random and a ball is chosen at random from it.

TREE DIAGRAMS

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

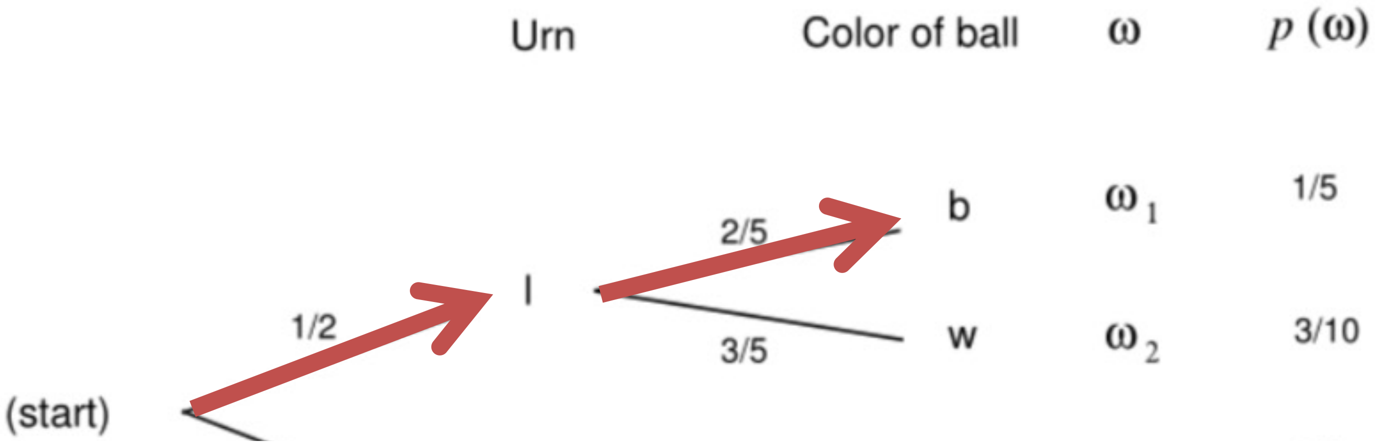


Tree diagram.

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.