Lesson - Probability: experiments, Venn and tree diagrams

Year: Year 11

Probability is on both the Foundation and the Higher papers. Every probability is a number from 0 (impossible) to 1 (certain). Give answers as fractions, decimals or percentages, and simplify fractions unless told otherwise.

Starter activity

Roll a die 30 times and write down how many times you get a six. Work out the fraction of rolls that gave a six. Compare your result with the probability $\frac{1}{6}$. Collect the results of the whole class together. What happens to the fraction as the number of rolls gets bigger? Explain in one sentence why more trials give a better estimate.

Basic probability and experiments

Rule: equally likely outcomes

If all outcomes are equally likely, $P(\text{event}) = \dfrac{\text{number of successful outcomes}}{\text{total number of outcomes}}$. The probabilities of all the outcomes add up to 1, so $P(\text{not } A) = 1 - P(A)$.

Worked example 1: a spinner

A spinner is red with probability 0.3 and blue with probability 0.5. Find $P(\text{green})$.

$1 - 0.3 - 0.5 = 0.2$.

Definition: relative frequency

When outcomes are not equally likely (a biased dice, drawing pins), estimate the probability from an experiment: $\text{relative frequency} = \dfrac{\text{number of times it happens}}{\text{total number of trials}}$. The more trials, the more reliable the estimate. The **expected number** is probability $\times$ number of trials.

Worked example 2: drawing pins and fairness

320 out of 500 drawing pins land point up. Estimate the probability and the expected number in 150 drops. A spinner lands on red 28 times in 80 spins; should red have probability 0.25?

The estimate is $\frac{320}{500} = 0.64$, so in 150 drops we expect $0.64 \times 150 = 96$. For the spinner, the relative frequency is $\frac{28}{80} = 0.35$ and we would expect $0.25 \times 80 = 20$ reds. The spinner may be biased, but 80 trials is not many, so do more trials before deciding. Give the numbers, then a conclusion.

Sample space diagrams

Definition: sample space

List all the outcomes systematically in a table or a list. Two dice have 36 equally likely outcomes. Two coins have outcomes HH, HT, TH, TT.

Worked example 3: dice and coins

Two dice are added: find $P(7)$ and $P(8)$. Two coins are tossed: find $P(\text{two heads})$ and $P(\text{at least one head})$.

A sum of 7 occurs in 6 ways, so $P(7) = \frac{6}{36} = \frac{1}{6}$. A sum of 8 occurs in 5 ways $(2,6), (3,5), (4,4), (5,3), (6,2)$, so $P(8) = \frac{5}{36}$. For the coins, $P(\text{two heads}) = \frac{1}{4}$ and $P(\text{at least one head}) = \frac{3}{4}$.

Venn diagrams

Definition: Venn diagram notation

A Venn diagram sorts items into overlapping sets. The overlap is "both" and the area outside the circles is "neither". $A \cup B$ means "A or B (or both)", and $A \cap B$ means "A and B".

Rule: addition rule

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$. **Mutually exclusive** events cannot happen together, so $P(A \text{ or } B) = P(A) + P(B)$.

Method: filling in a Venn diagram

1. Fill in the overlap first.
2. Fill in the other parts of each circle by subtracting the overlap.
3. Fill in the outside from the total, then read off probabilities.

Worked example 4: football and tennis

30 students: 18 play football (F), 12 play tennis (T), 5 play both. Find $P(F \cup T)$, $P(F \cap T)$ and $P(F \text{ only})$.

Both $= 5$. F only $= 13$. T only $= 7$. Neither $= 30 - 25 = 5$. Then $P(F \cup T) = \frac{25}{30} = \frac{5}{6}$, $P(F \cap T) = \frac{5}{30} = \frac{1}{6}$ and $P(F \text{ only}) = \frac{13}{30}$.

Independent events and tree diagrams

Definition: independent events

Independent events do not affect each other, so $P(A \text{ and } B) = P(A) \times P(B)$. If $P(A) = 0.3$ and $P(B) = 0.5$ are independent then $P(A \text{ and } B) = 0.15$.

Method: tree diagrams

1. Draw a branch for each outcome and write the probability on it (the branches from one point add up to 1).
2. **Multiply along** the branches to get the probability of a combination.
3. **Add** the results of the combinations you want.
4. For "at least one", use the complement: $1 - P(\text{none})$.

Worked example 5: with replacement

A bag has 5 red and 3 blue counters. Two are taken with replacement. Find $P(RR)$ and $P(\text{one of each})$.

$P(RR) = \frac{5}{8} \times \frac{5}{8} = \frac{25}{64}$. $P(\text{one of each}) = 2 \times \frac{5}{8} \times \frac{3}{8} = \frac{30}{64} = \frac{15}{32}$.

Worked example 6: without replacement

Same bag, but the first counter is not put back. Find $P(RR)$, $P(BB)$, $P(\text{different})$ and $P(\text{at least one blue})$.

After one is removed there are 7 left. $P(RR) = \frac{5}{8} \times \frac{4}{7} = \frac{5}{14}$. $P(BB) = \frac{3}{8} \times \frac{2}{7} = \frac{3}{28}$. $P(\text{different}) = 1 - \frac{5}{14} - \frac{3}{28} = \frac{15}{28}$. $P(\text{at least one blue}) = 1 - \frac{5}{14} = \frac{9}{14}$.

Two-way tables and frequency trees

Note: two-way tables and frequency trees

A two-way table sorts data by two features. Complete the row and column totals first, then read off probabilities as a fraction of the relevant total. A frequency tree uses numbers of people on the branches: from 200 people, if 40% are members, the branches show 80 members and 120 non-members; if 25% of members and 10% of non-members are under 18, the next branches show 20 and 60 members and 12 and 108 non-members.

Watch out: common mistakes

- Adding along the branches instead of multiplying.
- Not changing the second set of probabilities when there is no replacement.
- Writing probabilities greater than 1 or as ratios such as "3:5".
- Forgetting that "one of each" has two routes (RB and BR).
- Writing a Venn diagram total that does not match the number of people.

Worked example 7: Using relative frequency to predict

A bag contains red and blue counters. In 300 trials a red counter was taken 123 times. Estimate the probability of red and the number of red counters in a bag of 50.

The relative frequency of red is $\frac{123}{300} = 0.41$. This is the best estimate of $P(\text{red})$. In a bag of 50 counters we expect about $50 \times 0.41 = 20.5$, so about 20 or 21 red counters. This is an estimate: with more trials it would be more reliable.

Worked example 8: The addition rule for mutually exclusive events

A card is chosen from a pack where $P(\text{red}) = 0.3$, $P(\text{blue}) = 0.5$ and the rest are green. Find $P(\text{red or blue})$ and $P(\text{green})$.

The events are mutually exclusive because a card cannot be two colours at once. So $P(\text{red or blue}) = 0.3 + 0.5 = 0.8$. The probabilities add to 1, so $P(\text{green}) = 1 - 0.8 = 0.2$. The word "or" usually means add and the word "and" usually means multiply.

Worked example 9: The general addition rule with a Venn diagram

$P(A) = 0.5$, $P(B) = 0.4$ and $P(A \cap B) = 0.15$. Find $P(A \cup B)$ and $P(\text{neither})$.

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.5 + 0.4 - 0.15 = 0.75$. We subtract the overlap because it was counted twice. $P(\text{neither}) = 1 - 0.75 = 0.25$. On a Venn diagram: A only 0.35, both 0.15, B only 0.25, neither 0.25.

Worked example 10: A tree diagram with three branches

A fair spinner is red, blue or green with equal probability. It is spun twice.

Find the probability of getting two different colours.

The probability of getting the same colour twice is $\frac{1}{3} \times \frac{1}{3} = \frac{1}{9}$ for each colour, so $3 \times \frac{1}{9} = \frac{1}{3}$. Therefore different colours has probability $1 - \frac{1}{3} = \frac{2}{3}$. Using the complement is quicker than listing all six different pairs.

Worked example 11: Expected frequency

The probability that a train is late is 0.15. How many late trains are expected in 40 journeys? Is it possible for none to be late?

Expected number $= 40 \times 0.15 = 6$. This is an average and the real number varies. It is possible for none to be late: the probability is $0.85^{40} = 0.0015$, which is very small but not zero. Expected value does not mean "exactly".

Worked example 12: Sample space for a coin and a die

A fair coin is tossed and a fair die is rolled. Find the probability of getting heads and an even number.

There are $2 \times 6 = 12$ equally likely outcomes. Heads and even: (H,2), (H,4), (H,6), which is 3 outcomes. The probability is $\frac{3}{12} = \frac{1}{4}$. Because the events are independent, you could also multiply: $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

Worked example 13: Fairness and bias

A spinner is spun 80 times and lands on red 30 times. Is it fair if it is supposed to land on red with probability $\frac{1}{3}$?

A fair spinner would give about $80 \times \frac{1}{3} = 26.7$ reds. The experiment gave 30, which is 3.3 more. This is a small difference for 80 trials, so there is no strong evidence that the spinner is biased. More trials would give a clearer answer.

Method: choosing a probability tool

1. Use a **sample space table** for two events with a small number of outcomes, such as two dice or two spinners.
2. Use a **Venn diagram** to sort items into overlapping groups, filling in the overlap first.
3. Use a **tree diagram** for two or more events in sequence, multiplying along branches and adding between branches.
4. Check every answer: probabilities must be between 0 and 1 and all possible outcomes must add up to 1.

Note: mutually exclusive and independent are different ideas

Two events are **mutually exclusive** if they cannot happen at the same time, so $P(A \text{ and } B) = 0$. They are **independent** if one event does not change the probability of the other, so $P(A \text{ and } B) = P(A) \times P(B)$. Two events that both have a non-zero probability cannot be both mutually exclusive and independent. Students often confuse the two words, so write the definition you are using in your answer.

Watch out: probability mistakes

- Do not add the probabilities along the branches of a tree; multiply along the branches and add the results for different routes.
- When items are not replaced, the denominators change on the second stage.
- Relative frequency is only an estimate; it is more reliable with more trials.
- Write probabilities as fractions, decimals or percentages, never as a ratio such as $3 : 5$.

Did you know?

The study of probability began in the 17th century when a French gambler asked the mathematician Blaise Pascal why he lost money in a dice game. Pascal wrote letters about the problem to Pierre de Fermat, and these letters started the mathematics of chance. Today probability is used in weather forecasts, insurance, medical trials, games and the way computers recommend films and music.

Otto's tip

Check that every set of branches adds to 1 and the final probabilities of the whole tree add to 1. If either fails, there is a slip somewhere.