L3 PROBABILITY CONCEPTS
CHEAT SHEET

STUDY TIME

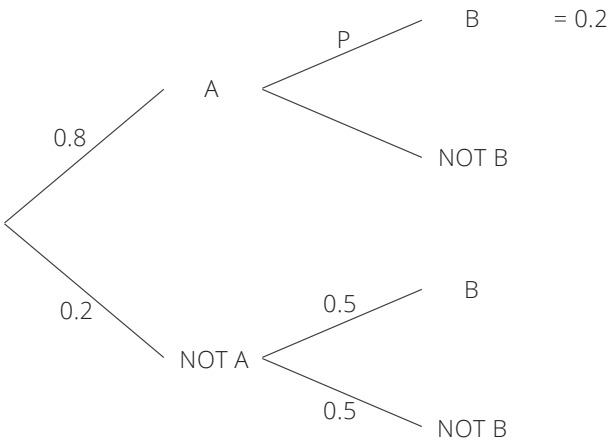
EXAM STRUCTURE

- The key skills in this paper are:
  - Probability trees
  - Tables
  - Venn diagrams
  - Mutually exclusive and independent events
  - Conditional probability
- Most of the questions will be calculation based, but you can also expect discussion questions about experimental, true and theoretical probabilities, and about applying answers in context.

BASIC THINGS TO KNOW

- Remember the **three key rules** to labelling and using probability trees!
  - Multiply if you're heading across the branches
  - Add if you're going down the outcomes
  - Branches coming off a particular point need to add up to 1
- You'll often have to **work backwards** to find the probability that goes along a branch, given the probability of the end event. Since you multiply along branches to get an end event, you can work backwards by dividing the end probability by the branch probability to get the missing number.
  - Let's say the probability of A is 0.8, and the end probability of A and B is 0.2, and we're looking for this missing probability 'p'. We know that the end probability is what you get when you multiply across the branches.

    Therefore, we know that $0.8 \times p = 0.2$



  - So, to get the probability of B "given" A – which is the missing branch – we would have to divide 0.2 by 0.8.

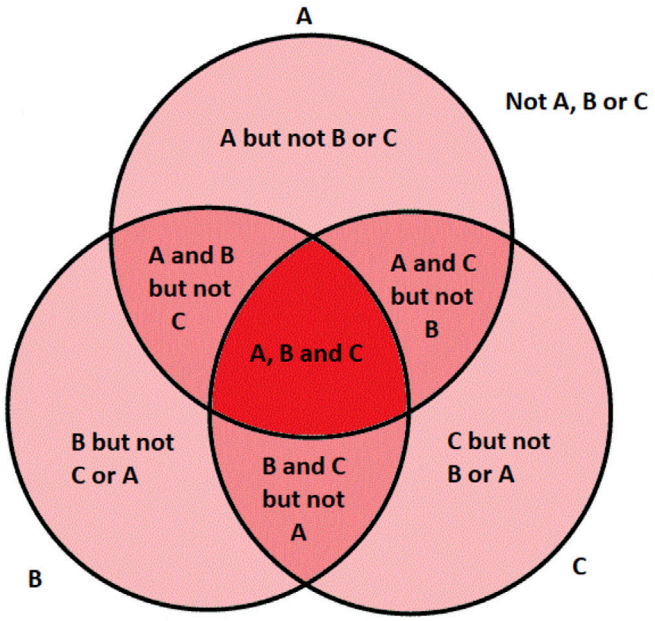
    $0.2 \div 0.8 = 0.25$, so our final answer for that branch would end up being 0.25.

TABLES

	CAUGHT COLD	DIDN'T CATCH COLD	TOTAL
GROUP A	16	24	40
GROUP B	15	45	60
TOTAL	31	69	100

- Just like in level 2, tables are often used in risk calculations. Remember, 'risk' is just another word for the probability of something happening for a particular group.
  - For example, the risk of catching a cold for group A is 16/40 = 0.4 and for group B is 15/60 = 0.25.
- **Relative risk** is a way of comparing risks for two groups.
  - RR = $\frac{\textit{Probability for group 1}}{\textit{Probability for group 2}}$
  - For the table above, the relative risk is 0.4/0.25 = 1.6
  - To interpret this, we say that people in group A were 1.6 times as likely to catch a cold compared to people in group B.

VENN DIAGRAMS



- The easiest way to keep track of all the information you're filling in on a three-way Venn diagram is to **colour code** the sections. We'd recommend bringing in highlighters or coloured pencils to your exam to do this.
- When you're placing the information from the question, try to **work from the centre outwards** if possible. If you're not given the value that goes in the very middle, look for any other information that tells you exactly where a number goes.
  - For example, in the diagram above, you could look for information about how many people were in 'C but not B or A', or 'Not A, B or C'.
- Once you're done filling in the diagram, **go back and add up all the numbers** to make sure they give you the total from the question!

OTHER CALCULATIONS

- There will probably be a question asking you whether two events are 'mutually exclusive' or 'independent'.
  - Remember, being 'exclusive' is like saying 'you can't sit with us' – those two events can't be seen together. Two events are **mutually exclusive** if the probability of both happening is 0.
  - Being **independent** is like saying 'I don't care what you do, I'll do what I want' – one event doesn't affect the other.
  - To check for independence, you need to use this formula:
    $P(A \cap B) = P(A) \times P(B)$
  - Calculate each side of the formula separately, and then decide whether they're equal or not. If they are, then A and B are independent. If not, they're non-independent.
  - Sometimes a question will ask you to **state any assumptions** you made while calculating your answer. The key here is almost always independence. Put the answer in context: instead of just saying "I assumed independence", say something like "I assumed that event A and event B were independent, so they did not affect each other".
- Finally, you'll want to be super familiar with this little formula for **conditional probability:** $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$
  - This is given to you, but it's up to you to recognise when to use it.
  - It's used any time you want to find the probability of an event, when you already know some information about another event.
  - Your main clue is whenever the question includes words like 'if' or 'given that' – for example "If B happens, what is the probability that A also happens?"

DISCUSSION QUESTIONS

- A classic discussion-based question is a **comparison between theoretical and experimental probabilities.** These questions are looking for you to explain why the experimental data doesn't come out exactly as expected. Your answer should include:
  - An explanation of the fact that experimental results are affected by random chance, so they don't usually exactly match the predicted results.
  - A numerical comparison of the two probabilities, explaining whether the differences are large or small.
  - A description of what could be done to compare the situation to the theoretical model better. This could involve increasing the sample size or running a simulation.
- The question will sometimes explicitly ask you to explain how **simulations** can be used to understand the effects of random variation.
  - Simulations are basically when you get a computer to randomly generate results to an experiment over and over again to see what kind of results you would get by random chance.
  - These let you compare your real result to the simulation results and see whether it was likely to happen by chance.
- As well as experimental and theoretical probabilities, these questions might also ask you about true probability. The key idea here is that **true probability** is never known, we can only estimate it by doing experiments and calculating theoretical probabilities.
- Finally, it's common for a question to start off by getting you to calculate a probability, and then for the next

part to ask you why this probability shouldn't necessarily be used to predict a future event. These are great questions to answer for some good marks even if you're not confident about calculation questions.

- As an example, let's imagine the question asked you to calculate what percentage of people from a sample passed a particular test. The next part might say "Explain why Sarah should be cautious about using these results to predict her own chances of passing the test next year". Things to take into account would be:
  - Was the sample size from the question small? The smaller the sample size, the more random chance will affect the results.
  - What other factors would you want to take into account before making a prediction? For example, the amount Sarah studied for the test would impact her results.
  - Are the results likely to change over time? In this example, maybe the test has gotten easier or harder to pass.