

YEAR 2 STATISTICS • CHAPTER 2

Conditional Probability

Specification: Edexcel Paper 3 §3.1–3.3 — Probability • AQA section M1–M3

KEY FACTS — THIS CHAPTER

- $P(A | B)$ is the probability of A **given** that B has happened: $P(A | B) = \frac{P(A \cap B)}{P(B)}$.
- Knowing B shrinks the sample space to B : count, or add up probabilities, inside B only.
- **Multiplication rule:** $P(A \cap B) = P(A) \times P(B | A)$. On a tree diagram, the second-stage branches are conditional probabilities.
- $P(A | B)$ and $P(B | A)$ are usually different. To reverse a tree, divide: $P(\text{first} | \text{second}) = \frac{P(\text{both})}{P(\text{second})}$.
- A and B are **independent** exactly when $P(A | B) = P(A)$: knowing B doesn't change the chance of A . Then $P(A \cap B) = P(A) \times P(B)$.
- A probability model rests on assumptions, such as fixed probabilities and independence. Say whether they are realistic, and what happens if they fail.

Information changes chances. The probability that a student cycles to school depends on which year they are in; a positive medical test changes the chance that someone has the disease. **Conditional probability** is the probability of one event once you know that another has happened, and it is the idea behind every tree diagram you have drawn.

SET NOTATION

- $A \cap B$: A **and** B both happen.
- $A \cup B$: A **or** B or both happen.
- A' : A does **not** happen.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, from Year 1.

1. Conditional probability

$P(A \mid B)$ is read “the probability of A given B ”. Once you know that B has happened, every outcome outside B is ruled out: the sample space shrinks to B alone, and you ask what fraction of it lies in A .

WORKED EXAMPLE 2.1 • FROM A TWO-WAY TABLE

The table shows how 200 sixth-form students travel to school. A student is chosen at random. Find the probability that the student cycles, the probability that the student cycles given that they are in Year 12, and the probability that the student is in Year 13 given that they cycle.

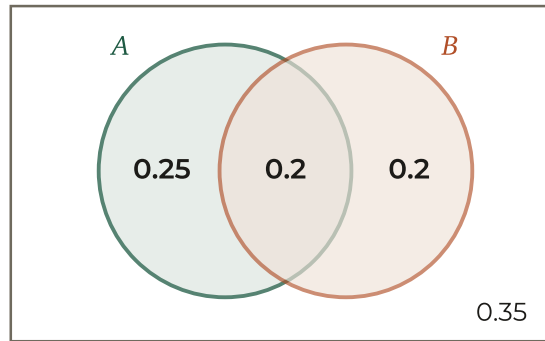
	Cycles	Does not cycle	Total
Year 12	27	63	90
Year 13	22	88	110
Total	49	151	200

1. $P(\text{cycles}) = \frac{49}{200} = 0.245$.
2. Given Year 12, only the 90 students in that row count, and 27 of them cycle: $P(\text{cycles} \mid \text{Y12}) = \frac{27}{90} = 0.3$.
3. Given that the student cycles, only the 49 cyclists count, and 22 of them are in Year 13: $P(\text{Y13} \mid \text{cycles}) = \frac{22}{49} = 0.449$ (3 s.f.).

Dividing counts is the same as dividing probabilities, since each count is a probability multiplied by 200. That gives the general rule.

CONDITIONAL PROBABILITY

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



Example 2.2: fill in the overlap first, then the rest of each circle, then the outside.

WORKED EXAMPLE 2.2 · FROM A VENN DIAGRAM

$P(A) = 0.45$, $P(B) = 0.4$ and $P(A \cup B) = 0.65$. Find $P(A \cap B)$, $P(A | B)$, $P(B | A)$ and $P(A' | B')$.

1. $P(A \cap B) = 0.45 + 0.4 - 0.65 = 0.2$. The Venn diagram follows: 0.25 in A only, 0.2 in B only and 0.35 outside both.
2. $P(A | B) = \frac{0.2}{0.4} = 0.5$ and $P(B | A) = \frac{0.2}{0.45} = \frac{4}{9}$.
3. $A' \cap B'$ is the region outside both circles, and $P(B') = 0.6$, so

$$P(A' | B') = \frac{0.35}{0.6} = \frac{7}{12}$$

COMMON MISTAKE

$P(A | B)$ and $P(B | A)$ are different questions with different answers. In Example 2.1, 30% of Year 12 students cycle, but only 55% of cyclists are in Year 12. The event after the bar is the one you already know about: it sets the denominator.

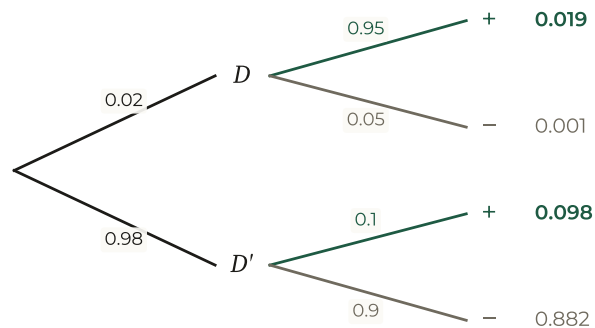
2. Tree diagrams

THE MULTIPLICATION RULE

$$P(A \cap B) = P(A) \times P(B | A)$$

On a tree diagram, each second-stage branch is a conditional probability: the probability of the second event **given** the branch you are already on.

Trees are drawn in the order things happen, but questions often ask the other way round: given what you can see (a test result), how likely is what you can't (the disease)? Find the two probabilities from the tree, then divide.



Example 2.3: two of the four paths end in a positive test.

WORKED EXAMPLE 2.3 • A POSITIVE TEST

A disease affects 2% of a population. A test for it is positive for 95% of people who have the disease, and also, wrongly, for 10% of people who do not. A person chosen at random tests positive. Find the probability that they have the disease.

1. Let D be “has the disease” and $+$ “tests positive”. The two paths that end in a positive test give

$$P(D \cap +) = 0.02 \times 0.95 = 0.019$$

$$P(D' \cap +) = 0.98 \times 0.1 = 0.098$$

2. So $P(+) = 0.019 + 0.098 = 0.117$, and

$$P(D | +) = \frac{0.019}{0.117} = 0.162 \text{ (3 s.f.)}$$

IN CONTEXT

Only about 16% of people who test positive have the disease. The disease is rare, so the few true positives are outnumbered by false positives from the much larger healthy group.

COMMON MISTAKE

The test is “95% accurate” for people with the disease, but that is $P(+ \mid D)$, not $P(D \mid +)$. Confusing the two is a classic error, in courtrooms and clinics as well as exams.

3. Independence

Two events are independent when knowing that one has happened tells you nothing about the other. Conditional probability says exactly what that means.

INDEPENDENCE

A and B are independent exactly when

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

and then also $P(B \mid A) = P(B)$ and $P(A \cap B) = P(A) \times P(B)$. Any one of these shows independence. If A and B are independent, so are A and B' .

WORKED EXAMPLE 2.4 · TESTING FOR INDEPENDENCE

Using the table in Example 2.1, determine whether cycling to school is independent of being in Year 12.

$P(\text{cycles}) = 0.245$, but $P(\text{cycles} \mid \text{Y12}) = 0.3$. These are different, so the events are **not** independent.

IN CONTEXT

Knowing that a student is in Year 12 raises the chance that they cycle, from 24.5% to 30%.

WORKED EXAMPLE 2.5 · USING INDEPENDENCE

A and B are independent events with $P(A) = 0.3$ and $P(A \cup B) = 0.65$. Find $P(B)$ and $P(B \mid A')$.

1. Independence gives $P(A \cap B) = 0.3P(B)$, so

$$0.65 = 0.3 + P(B) - 0.3P(B)$$

$$0.35 = 0.7P(B)$$

and $P(B) = 0.5$.

2. B is independent of A , so also of A' : $P(B \mid A') = P(B) = 0.5$.
-
-

4. Modelling with probability

Every probability model rests on assumptions: that outcomes are equally likely, that probabilities stay fixed, that events are independent. A good answer names the assumption, says whether it is realistic in context, and says what would change if it failed.

WORKED EXAMPLE 2.6 · CRITICISING A MODEL

An analyst models a tennis player's first serves: each lands in with probability 0.62, independently of every other serve. Give two criticisms of the model and the likely effect of each.

1. **The probability is unlikely to be fixed.** It will fall as the player tires, and vary with the score and the conditions. A single value of 0.62 hides this, so the model will be least accurate late in long matches.
 2. **Serves are unlikely to be independent.** A player who has just served well may gain confidence, and one who has missed may play safe. Successes and failures would then come in longer runs than the model predicts.
-

Practice

Twelve questions in three tiers. Try each one properly before opening the solution. The solutions show the full method, not just answers. Give answers as exact fractions or decimals, or to 3 significant figures where they aren't exact.

Warm-up

Question 1

[3 marks]

80 adults were asked whether they prefer tea or coffee.

	Tea	Coffee	Total
Under 40	12	28	40
40 or over	26	14	40
Total	38	42	80

One adult is chosen at random. Find (a) $P(\text{tea})$, (b) $P(\text{tea} \mid \text{under 40})$, (c) $P(\text{under 40} \mid \text{coffee})$.

SOLUTION

(a) $\frac{38}{80} = 0.475$. (b) $\frac{12}{40} = 0.3$. (c) $\frac{28}{42} = \frac{2}{3}$.

Question 2

[2 marks]

$P(A) = 0.6$, $P(B) = 0.5$ and $P(A \cap B) = 0.35$. Find $P(A \mid B)$ and $P(B \mid A)$.

SOLUTION

$$P(A \mid B) = \frac{0.35}{0.5} = 0.7 \text{ and } P(B \mid A) = \frac{0.35}{0.6} = \frac{7}{12}.$$

Question 3

[3 marks]

A bag holds 4 red and 6 blue counters. Two are taken without replacement. Find the probability that the second is red given that the first is red, and the probability that both are red.

SOLUTION

After a red is taken, 3 of the 9 left are red: $P(R_2 | R_1) = \frac{3}{9} = \frac{1}{3}$.

$$P(R_1 \cap R_2) = \frac{4}{10} \times \frac{1}{3} = \frac{2}{15}.$$

Question 4

[2 marks]

$P(A) = 0.4$, $P(B) = 0.25$ and $P(A \cap B) = 0.1$. Are A and B independent?

SOLUTION

$$P(A | B) = \frac{0.1}{0.25} = 0.4 = P(A), \text{ so yes: } A \text{ and } B \text{ are independent.}$$

Standard**Question 5**

[6 marks]

Of 60 students, 30 study Maths (M), 25 Physics (P) and 20 Chemistry (C). 12 study both Maths and Physics, 8 Maths and Chemistry, 7 Physics and Chemistry, and 4 study all three.

- (a) Draw a Venn diagram showing the number in each region. [3]
- (b) A student is chosen at random. Find $P(M | P)$ and $P(C | M \cap P)$. [2]
- (c) Find the probability that the student studies none of the three. [1]

SOLUTION

(a) Start from the centre, 4, then the pairs without it:

- Maths and Physics only: $12 - 4 = 8$; Maths and Chemistry only: $8 - 4 = 4$; Physics and Chemistry only: $7 - 4 = 3$.
- Maths only: $30 - 8 - 4 - 4 = 14$; Physics only: $25 - 8 - 3 - 4 = 10$; Chemistry only: $20 - 4 - 3 - 4 = 9$.
- Outside all three: $60 - 52 = 8$.

(b) $P(M | P) = \frac{12}{25}$, and $P(C | M \cap P) = \frac{4}{12} = \frac{1}{3}$.

(c) $\frac{8}{60} = \frac{2}{15}$.

Question 6

[5 marks]

Machine X makes 60% of a factory's components and machine Y the rest. 3% of X 's components are faulty, and 5% of Y 's. A component chosen at random is faulty. Find the probability that it was made by machine X .

SOLUTION

$$P(X \cap F) = 0.6 \times 0.03 = 0.018$$

$$P(Y \cap F) = 0.4 \times 0.05 = 0.02$$

$$\text{So } P(F) = 0.038 \text{ and } P(X | F) = \frac{0.018}{0.038} = \frac{9}{19} = 0.474.$$

Question 7

[4 marks]

A and B are independent, with $P(A) = 0.4$ and $P(B) = 0.3$. Find $P(A \cup B)$, $P(A | B')$ and $P(A' \cap B')$.

SOLUTION

$$P(A \cap B) = 0.4 \times 0.3 = 0.12, \text{ so } P(A \cup B) = 0.4 + 0.3 - 0.12 = 0.58.$$

$$A \text{ is independent of } B' \text{ too, so } P(A | B') = P(A) = 0.4.$$

$$P(A' \cap B') = 0.6 \times 0.7 = 0.42 \text{ (or } 1 - 0.58).$$

Question 8

[4 marks]

$P(A) = 0.3$, $P(B) = 0.4$ and $P(A | B) = 0.25$. Find $P(A \cap B)$, $P(B | A)$ and $P(A \cup B)$.

SOLUTION

$$P(A \cap B) = P(B) \times P(A | B) = 0.4 \times 0.25 = 0.1.$$

$$P(B | A) = \frac{0.1}{0.3} = \frac{1}{3}, \text{ and } P(A \cup B) = 0.3 + 0.4 - 0.1 = 0.6.$$

Question 9

[4 marks]

A weather model assumes that it rains on any day with probability 0.3, independently of every other day. It predicts that the chance of a week without rain is $0.7^7 = 0.0824$. Criticise two assumptions of the model and say how each might affect the prediction.

SOLUTION

Independence: weather comes in spells, so a dry day makes another dry day more likely. Dry weeks are then more common than the model predicts, so 0.0824 is probably too low.

A fixed probability: the chance of rain varies with the season, so one value of 0.3 all year will overestimate dry weeks in wet months and underestimate them in dry ones.

Exam style

Question 10

[7 marks]

For events A and B , $P(A) = 0.35$ and $P(B) = 0.4$, and the probability that neither happens is 0.35.

- (a) Find $P(A \cap B)$ and draw a Venn diagram. [3]
- (b) Find $P(A \mid B)$ and $P(B \mid A')$. [3]
- (c) Determine whether A and B are independent. [1]

SOLUTION

(a) $P(A \cup B) = 1 - 0.35 = 0.65$, so $P(A \cap B) = 0.35 + 0.4 - 0.65 = 0.1$. The regions are 0.25 (A only), 0.1 (both), 0.3 (B only) and 0.35 (neither).

(b) $P(A \mid B) = \frac{0.1}{0.4} = 0.25$. $P(A') = 0.65$ and $P(B \cap A') = 0.3$, so $P(B \mid A') = \frac{0.3}{0.65} = \frac{6}{13} = 0.462$.

(c) $P(A \mid B) = 0.25$ but $P(A) = 0.35$, so they are not independent.

Question 11

[8 marks]

A student's bus is late on 15% of days. When the bus is late, the student is late for school with probability 0.6; when it isn't, with probability 0.1.

- (a) Find the probability that the student is late on a given day. [3]
- (b) Given that the student is late, find the probability that the bus was late. [3]
- (c) Assuming days are independent, find the probability that the student is late on exactly one of two days. [2]

SOLUTION

(a) $P(L) = 0.15 \times 0.6 + 0.85 \times 0.1 = 0.09 + 0.085 = 0.175$.

(b) $P(B \mid L) = \frac{0.09}{0.175} = \frac{18}{35} = 0.514$.

(c) Late on the first day only, or the second only: $2 \times 0.175 \times 0.825 = 0.289$ (3 s.f.).

Question 12

[8 marks]

Events A , B and C have $P(A) = 0.2$, $P(B) = 0.5$ and $P(C) = 0.3$. A and C are mutually exclusive, A and B are independent, and $P(B \cap C) = 0.1$.

- (a) Find $P(A \cap B)$. [1]
- (b) Draw a Venn diagram showing the probability in each region. [4]
- (c) Find $P(B \mid C)$ and $P(A \cup B \cup C)$. [3]

SOLUTION

(a) Independent, so $P(A \cap B) = 0.2 \times 0.5 = 0.1$.

(b) A and C don't overlap, so no region is in all three. The regions are A only 0.1, $A \cap B$ 0.1, B only 0.3, $B \cap C$ 0.1, C only 0.2, and outside all three $1 - 0.8 = 0.2$.

(c) $P(B \mid C) = \frac{0.1}{0.3} = \frac{1}{3}$, and $P(A \cup B \cup C) = 0.1 + 0.1 + 0.3 + 0.1 + 0.2 = 0.8$.