Cambridge AS & A Level · Mathematics 9709 · Notes

Continuous random variables A2

Read online at pastopic.com/cambridge/mathematics-9709/notes/continuous-random-variables

Last checked 29 September 2026

Read first: Discrete random variables, Integration, Integration

What you need to know

By the end of this topic you should be able to:

1. **Explain what a continuous random variable is**, and know the two properties every **probability density function** (pdf) has: it is never negative, and the total area under its graph is 1. Use them, for example to find an unknown constant or to decide whether a function could be a pdf. The pdf is always defined on a **single interval**, which may go on for ever, such as $x \geqslant 2$.
2. **Use a pdf to find probabilities** (areas under the graph between two values), and to work out the **mean** $E(X)$ and the **variance** $\mathrm{Var}(X)$.
3. **Find the median, quartiles or other percentiles** by finding the value that cuts off the right area under the graph.

The **cumulative distribution function** $F(x)$ is **not** in the syllabus: find every area directly from $f(x)$. The mode, and pdfs made of two or more different pieces, are not needed either.

Understand it

Discrete and continuous

A discrete random variable, such as the score on a dice, takes separate values, and each has its own probability. A **continuous** random variable, such as a time, a length or a mass, can take any value in an interval. There are so many possible values that the chance of hitting one exact value, like $2.5000\ldots$ hours, is 0. So for a continuous variable we only ever ask for the probability of an **interval**, and

$$P(X = a) = 0, \qquad P(a < X < b) = P(a \leqslant X \leqslant b).$$

Whether an inequality is strict or not makes no difference.

Probability is area

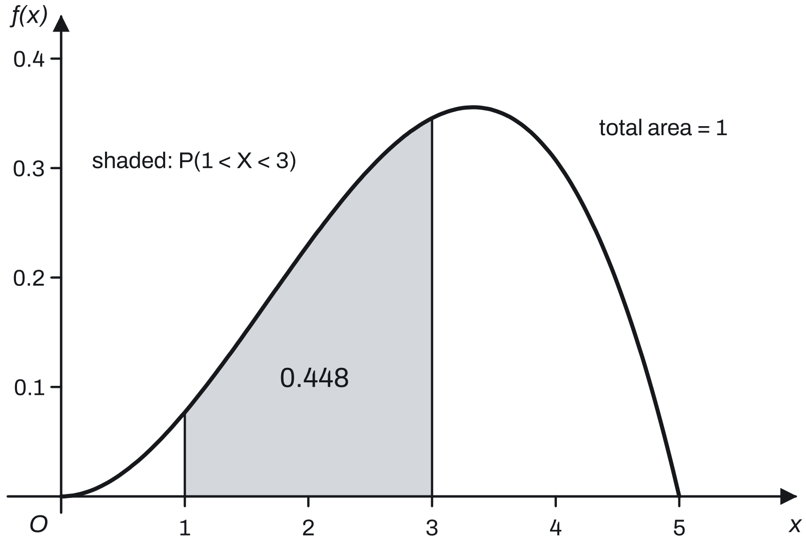
A continuous variable is described by its **probability density function**, $f(x)$. Its graph shows where the values are crowded (high parts) and where they are rare (low parts). The probability that X lands between a and b is the **area under the graph** between a and b:

$$P(a < X < b) = \int_a^b f(x)\,dx.$$

The height $f(x)$ itself is **not** a probability. It can be bigger than 1; only areas are probabilities.

As a running example, let X hours be the time a phone battery lasts after its low-power warning, with

$$f(x) = \begin{cases} \dfrac{12}{625}x^2(5-x) & 0 \leqslant x \leqslant 5, \\ 0 & \text{otherwise.} \end{cases}$$



Then

$$P(1 < X < 3) = \frac{12}{625}\int_1^3 (5x^2 - x^3)\,dx = \frac{12}{625}\left[\frac{5x^3}{3} - \frac{x^4}{4}\right]_1^3 = \frac{56}{125} = 0.448.$$

The two rules for a pdf

Because every value of X must be somewhere in the interval, and probabilities can't be negative:

- $f(x) \geqslant 0$ for every x;
- the total area is 1: $\displaystyle\int f(x)\,dx = 1$ over the whole interval where f is not zero.

The second rule finds an unknown constant. For $f(x) = k\,x^2(5-x)$ on $0 \leqslant x \leqslant 5$: $k\left[\frac{5x^3}{3} - \frac{x^4}{4}\right]_0^5 = k \times \frac{625}{12} = 1$, so $k = \frac{12}{625}$.

To show a function **could be** a pdf, check both rules and say so. For example, $f(x) = \frac{1}{4}(3x-1)$ on $0 \leqslant x \leqslant 2$ has total area 1, but $f(0) = -\frac{1}{4} < 0$, so it **cannot** be a pdf.

When the graph is a straight line or part of a circle, you can find areas with geometry instead of integration: a triangle is $\frac{1}{2} \times \text{base} \times \text{height}$, a trapezium is $\frac{1}{2}(a+b)h$, a quarter circle of radius r is $\frac{1}{4}\pi r^2$.

An interval that goes on for ever

The interval can be infinite, for example $f(x) = \frac{6}{x^3}$ for $x \geqslant \sqrt{3}$. Integrate up to ∞: as $x \to \infty$, a term like $\frac{3}{x^2}$ tends to 0. So

$$\int_{\sqrt{3}}^{\infty} 6x^{-3}\,dx = \left[-\frac{3}{x^2}\right]_{\sqrt{3}}^{\infty} = 0 - \left(-\frac{3}{3}\right) = 1.$$

Mean and variance

For a discrete variable, $E(X) = \sum xp$. For a continuous variable the sum becomes an integral and p becomes $f(x)\,dx$:

$$E(X) = \int xf(x)\,dx, \qquad \mathrm{Var}(X) = \int x^2 f(x)\,dx - \{E(X)\}^2,$$

each over the interval where f is not zero. The mean is the **balance point** of the graph: if you cut the shape out of card, it would balance on a pin at $E(X)$.

For the battery: $E(X) = \frac{12}{625}\int_0^5 (5x^3 - x^4)\,dx = \frac{12}{625}\left[\frac{5x^4}{4} - \frac{x^5}{5}\right]_0^5 = 3$ hours, and $\int_0^5 x^2 f(x)\,dx = \frac{12}{625}\left[x^5 - \frac{x^6}{6}\right]_0^5 = 10$, so $\mathrm{Var}(X) = 10 - 3^2 = 1$.

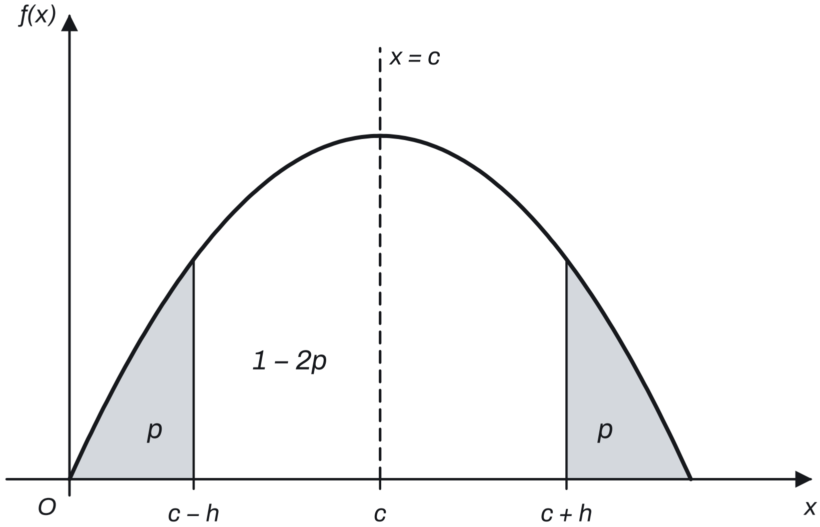
Median and percentiles

The **median** m splits the area into two halves: $\int_{\text{start}}^{m} f(x)\,dx = \frac{1}{2}$. The **lower quartile** cuts off $\frac{1}{4}$ of the area on the left, the **upper quartile** $\frac{3}{4}$, and the 90th percentile 0.9. You integrate from the start of the interval up to the unknown and solve the equation for it.

For the battery, $\frac{12}{625}\left(\frac{5m^3}{3} - \frac{m^4}{4}\right) = \frac{1}{2}$, which tidies up to $6m^4 - 40m^3 + 625 = 0$. Its root between 0 and 5 is $m = 3.07$ (3 s.f.). Always reject roots outside the interval. Here the mean (3) is a little below the median, because the long tail on the left pulls the balance point that way.

Symmetry

If the graph is **symmetrical** about $x = c$, then $E(X) = c$ and the median is c too, with no integration. Areas the same distance either side of c are equal, so you can often find a probability from one you already know.



In the sketch, the tail to the right of $c+h$ has area p. By symmetry the tail to the left of $c-h$ is also p, the middle is $1-2p$, and the part from c to $c+h$ is $\frac{1}{2}-p$. Always sketch the graph and label the areas you know.

Moving a graph to the right by c adds c to the mean and leaves the variance alone. Squashing it to half its width (and doubling its height, to keep the area 1) keeps the centre and divides the variance by 4.

Key facts and formulas

"Given" means it is in the MF19 list of formulae in the exam. "Learn it" means you must remember it.

Formula	Status
$E(X)=\int xf(x)\,dx$	Given
$\mathrm{Var}(X)=\int x^2f(x)\,dx-\{E(X)\}^2$	Given
$P(a<X<b)=\int_a^b f(x)\,dx$	Learn it
$f(x)\geqslant 0$ for all x, and $\int f(x)\,dx=1$ over the whole interval	Learn it
$P(X=a)=0$, so $<$ and $\leqslant$ give the same probability	Learn it
Median m: $\int_{\text{start}}^{m} f(x)\,dx=\frac{1}{2}$; lower quartile $\frac{1}{4}$, upper quartile $\frac{3}{4}$	Learn it
Graph symmetrical about $x=c$: $E(X)=c$ and median $=c$	Learn it
$P(X>a)=1-P(X<a)$	Learn it
$\int x^n\,dx=\frac{x^{n+1}}{n+1}\ (n\neq -1)$, $\int \frac{1}{x}\,dx=\ln x$	Given
Integration by parts: $\int u\frac{dv}{dx}\,dx=uv-\int v\frac{du}{dx}\,dx$	Given

Formula	Status
$\int \sin x \, dx = -\cos x$, $\int \cos x \, dx = \sin x$	Given
$\int \cos(ax) \, dx = \dfrac{\sin(ax)}{a}$, $\int \sin(ax) \, dx = -\dfrac{\cos(ax)}{a}$	Learn it
Triangle $\frac{1}{2}bh$, trapezium $\frac{1}{2}(a+b)h$, circle πr^2	Learn it
Sector of a circle: area $\frac{1}{2}r^2\theta$ (θ in radians)	Given

For an infinite interval, terms such as $\frac{k}{x^n}$ (with $n > 0$) tend to 0 as $x \to \infty$.

How to do it

In the Paper 6 questions we have tagged for this topic (39 questions, 2021–2025; some papers share a question), the types came up this often. 37 of the 39 questions mix several types, so the counts add up to more than 39.

Question type	Questions
Finding a constant: total area is 1	25
The mean, E(X)	25
Using the graph: symmetry and shapes	24
Median, percentiles and unknown limits	19
Probabilities by integration	11
The variance, Var(X)	6
Models in context: meaning, limits and repeated values	3

1. Finding a constant: total area is 1 (25 questions)

1. **Set the total area to 1**: $\int f(x)\,dx = 1$ with the limits of the interval.
2. **Integrate**, keeping the constant outside: $k\int \ldots$. Expand brackets first unless the bracket integrates directly (for example $(x-4)^2 \to \frac{(x-4)^3}{3}$).
3. **Substitute both limits** and show the line with numbers in, then solve for the constant. This is usually "show that", so show each step.

Variations you will meet:

- **Straight-line or circle graphs.** Use the area of the triangle, trapezium, semicircle or quarter circle instead of integrating: for a semicircle of radius r, $\frac{1}{2}\pi r^2 = 1$ gives $r = \sqrt{2/\pi}$.
- **Writing $f(x)$ from a straight-line graph.** Find the gradient from two points, then $y = mx + c$. Check that the sign of the gradient and the intercept match the picture.

- **Two unknowns** (such as a and b, or the end of the interval as well as k). You need two equations: total area $= 1$, plus a given mean, median or point on the graph. The graph often gives the second one: if it reaches the x-axis at $x = 2$, then $f(2) = 0$. Solve them together.
- **An unknown end of the interval.** Integrate from the known end to p and set it equal to 1, then solve; reject values outside the sensible range.
- **"Show that f could be a pdf" or "verify"**: show the area is 1 **and** that $f(x) \geqslant 0$ on the interval, and finish with a sentence.
- **"Write down" a probability.** $P(X \geqslant \text{start})$ is 1; half of a symmetrical graph has area $\frac{1}{2}$. No integration is needed.

2. The mean, E(X) (25 questions)

1. Write $E(X) = \int x f(x)\,dx$ with the interval's limits. **Multiply by x first**, then integrate term by term.
2. Substitute the limits and simplify. Give an exact answer if asked.

Variations you will meet:

- **Symmetrical graph**: state $E(X) = c$ and give the reason, "by symmetry" (not "the middle of the end points").
- **A given mean finds a constant.** Set $\int x f(x)\,dx$ equal to the given value and solve, together with the total-area equation if there are two unknowns.
- **$\frac{k}{x^2}$ or $\frac{k}{x}$ terms.** $x \times \frac{k}{x^2} = \frac{k}{x}$, which integrates to $k \ln x$; expect $\ln$ in the answer.
- **Trigonometric pdfs**, such as $1 + \cos(\pi x)$ or $x(1 + \cos x)$. For probabilities, divide by the number multiplying x: $\int \cos(ax)\,dx = \frac{\sin(ax)}{a}$, so $\int (1 + \cos \pi x)\,dx = x + \frac{\sin \pi x}{\pi}$. For the mean, integrate $x \cos x$ by parts with $u = x$: $\int x \cos x\,dx = x \sin x + \cos x$. Similarly $\int x \cos(\pi x)\,dx = \frac{x \sin \pi x}{\pi} + \frac{\cos \pi x}{\pi^2}$.
- **A probability involving the mean**, such as $P(X < E(X))$: find $E(X)$, then integrate $f(x)$ from the start up to it.

3. Using the graph: symmetry and shapes (24 questions)

1. **Sketch the graph** (or use the one given) and mark the line of symmetry and the given areas.
2. **Use equal areas** either side of the line of symmetry and the total area of 1 (each half is $\frac{1}{2}$).
3. Build the area you want by adding or subtracting areas you know.

For example, if the graph is symmetrical about $x = 5$ and $P(X < 3) = 0.15$, then $P(X > 7) = 0.15$, so $P(3 < X < 7) = 1 - 2(0.15) = 0.7$, and $P(5 < X < 7) = 0.5 - 0.15 = 0.35$.

Variations you will meet:

- **Areas in terms of letters.** With $P(X < a) = p$, say, the same ideas give answers such as $1 - 2p$ or $\frac{1}{2} - p$. Label every piece on the sketch.
- **Explain without calculating**, for example why the median is below some value: say that the area to the left of that value is more than $\frac{1}{2}$ (the graph is higher there), so the median must be to its left.

- **Semicircle or quarter circle.** A probability can be a sector minus a triangle: find the angle with $\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}}$, then use $\frac{1}{2}r^2\theta$ and $\frac{1}{2}ab\sin C$.
- **Sketch a pdf with a given mean and variance** related to one you know: move the graph along for a new mean; make it narrower and taller for a smaller variance, keeping the area 1.
- **Symmetrical quadratic**, such as $k(x-1)(7-x)$: its axis of symmetry is halfway between the roots, $x = 4$, so $E(X) = 4$ and the median is 4.

4. Median, percentiles and unknown limits (19 questions)

1. **Write the equation**: $\int_{\text{start}}^{m} f(x)\,dx = \frac{1}{2}$ for the median (or $\frac{1}{4}$, $\frac{3}{4}$, 0.9 for other percentiles). You can use $\int_{m}^{\text{end}} f(x)\,dx = \frac{1}{2}$ instead; for a top percentile the right-hand tail is often quicker.
2. **Integrate and substitute** the limits, one of which is the unknown.
3. **Solve**: a quadratic in m (formula), a quadratic in m^2, a cube root, or $\ln$. Reject any root outside the interval and say why.

Variations you will meet:

- **"Show that m satisfies ..."**: rearrange carefully to the given equation, clearing fractions one step at a time.
- **Showing that a median (or other percentile) is between two given numbers.** Find the area from the start of the interval up to each number: one area should come out below $\frac{1}{2}$ and the other above, so the median is trapped between them. Another way is to put both numbers into the rearranged equation and show that it changes sign. Either way, write down both values and finish with a sentence.
- **Probability between the mean and the median**, such as $P(m < X < 3)$: you don't need m itself: it is $P(X < 3) - \frac{1}{2}$, or $\frac{1}{2} - P(X > 3)$.
- **Find c so that $P(-c < X < c) = \frac{1}{2}$** for a graph symmetrical about 0: use $\int_0^c f(x)\,dx = \frac{1}{4}$.
- **The median gives a constant.** If the median is known, $\int_{\text{start}}^{m} f(x)\,dx = \frac{1}{2}$ is an equation for the unknown constant.

5. Probabilities by integration (11 questions)

1. $P(a < X < b) = \int_a^b f(x)\,dx$. Use the part of the interval the question asks about, not beyond it.
2. **Substitute both limits** and subtract.
3. For "more than", you can integrate from the value to the end, or do $1 -$ (the rest). Pick whichever part is easier.

Variations you will meet:

- **"Without integrating, use part (a)"**: use symmetry and the answer you have, and show the working. For example, if the graph is symmetrical about $x = 4$ and part (a) gave $P(X > 4.8) = \frac{3}{20}$, then $P(3.2 < X < 4.8) = 1 - 2 \times \frac{3}{20} = \frac{7}{10}$.
- **"Show that the probability is ..."**: substitute the limits and show the intermediate line.

6. The variance, Var(X) (6 questions)

1. Find $E(X)$ (by symmetry if you can).
2. Work out $\int x^2 f(x)\,dx$ over the interval.
3. $\mathrm{Var}(X) = \int x^2 f(x)\,dx - \{E(X)\}^2$. **Don't forget to subtract** the square of the mean.

Variations you will meet:

- **Find the domain first.** "$f(x) = k(\ldots)$ where $f(x) \geqslant 0$" means the interval runs between the roots of the quadratic.
- **"Show that $\mathrm{Var}(X) = \ldots$"** in terms of a letter, for example for a flat pdf $f(x) = \frac{1}{a}$ on $0 \leqslant x \leqslant a$: $\frac{a^2}{3} - \frac{a^2}{4} = \frac{a^2}{12}$.

7. Models in context: meaning, limits and repeated values (3 questions)

These questions put the pdf in a real setting (times, distances) and ask you to interpret it, as well as the usual calculations.

1. **Read the model**: say in words what X measures, and what the ends of its interval mean in the setting.
2. **One value**: find a probability p for a single observation, as in type 5.
3. **Several independent values**: multiply. Both of two observations: p^2. At least one of three: $1 - (1-p)^3$. The largest of three is more than c exactly when **at least one** of them is.

Variations you will meet:

- **What the end points mean.** Since $f(x) = 0$ outside the interval, its two ends are the smallest and largest values the model allows. Say this in the setting: for a waiting time on $1 \leqslant x \leqslant 9$ minutes, nobody waits less than 1 minute or more than 9.
- **Criticising the model.** Give a reason tied to the setting, for example that the model rules out values beyond the ends of its interval, or makes the graph a smooth shape that real data may not follow.

Worked examples

The bold codes show where marks are earned, as in Cambridge mark schemes. The number is how many marks.

- **M1** method mark: for a correct method, even if the numbers slip.
- **A1** accuracy mark: for a correct answer; it needs the M mark before it.
- **B1** mark for a correct result on its own, with no method needed.

Example 1

The random variable X has probability density function given by

$$f(x) = \begin{cases} k(8 - x^3) & 0 \leqslant x \leqslant 2, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant.

(a) Show that $k = \frac{1}{12}$. **[3]**

(b) Find $P(X > 1)$. **[3]**

(c) Find $E(X)$. **[3]**

(d) Two independent values of X are taken. Find the probability that both are greater than 1. **[1]**

(a) $k \int_0^2 (8 - x^3)\, dx = 1$. **M1** for integrating $f(x)$ and setting it equal to 1.

$$k \left[8x - \frac{x^4}{4}\right]_0^2 = k(16 - 4) = 12k = 1.$$

A1 for the correct integral with the limits, **A1** for $k = \frac{1}{12}$ with the $12k = 1$ line shown.

(b)

$$P(X > 1) = \frac{1}{12}\left[8x - \frac{x^4}{4}\right]_1^2 = \frac{1}{12}\left(12 - \frac{31}{4}\right) = \frac{17}{48} = 0.354 \text{ (3 s.f.)}$$

M1 for integrating with limits 1 and 2, **A1** for the correct values substituted, **A1** for $\frac{17}{48}$.

(c)

$$E(X) = \frac{1}{12}\int_0^2 (8x - x^4)\, dx = \frac{1}{12}\left[4x^2 - \frac{x^5}{5}\right]_0^2 = \frac{1}{12}\left(16 - \frac{32}{5}\right) = \frac{4}{5}.$$

M1 for integrating $xf(x)$, **A1** for the correct integral and limits, **A1** for $\frac{4}{5}$ or 0.8.

(d) $\left(\frac{17}{48}\right)^2 = \frac{289}{2304} = 0.125$ (3 s.f.). **B1**

Example 2

The random variable X has probability density function given by

$$f(x) = \begin{cases} \dfrac{k}{x^4} & x \geqslant 2, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that $k = 24$. **[2]**

(b) Find $E(X)$ and $\mathrm{Var}(X)$. **[5]**

(c) Find the median of X. **[3]**

(a)

$$\int_2^{\infty} kx^{-4}\,dx = \left[-\frac{k}{3x^3}\right]_2^{\infty} = 0 + \frac{k}{24} = 1 \;\Rightarrow\; k = 24.$$

M1 for integrating and using the limits 2 and ∞, **A1** with the step $\frac{k}{24} = 1$ shown.

(b)

$$E(X) = \int_2^{\infty} 24x^{-3}\,dx = \left[-\frac{12}{x^2}\right]_2^{\infty} = \frac{12}{4} = 3.$$

M1 A1

$$\int_2^{\infty} x^2 \cdot 24x^{-4}\,dx = \int_2^{\infty} 24x^{-2}\,dx = \left[-\frac{24}{x}\right]_2^{\infty} = 12.$$

M1 for integrating $x^2 f(x)$. $\mathrm{Var}(X) = 12 - 3^2 = 3$. **M1** for subtracting the square of their mean, **A1**.

(c)

$$\int_2^{m} 24x^{-4}\,dx = \left[-\frac{8}{x^3}\right]_2^{m} = 1 - \frac{8}{m^3} = \frac{1}{2}.$$

M1 for integrating from 2 to m and setting it equal to $\frac{1}{2}$. **A1** for a correct equation. So $m^3 = 16$ and $m = \sqrt[3]{16} = 2.52$ (3 s.f.). **A1**

Example 3

The graph of the probability density function, f, of a random variable X is a straight line from $(2, c)$ to $(5, 4c)$, where c is a positive constant. Elsewhere $f(x) = 0$.

(a) Find the value of c. **[2]**

(b) Show that $f(x) = \frac{2}{15}(x-1)$ for $2 \leqslant x \leqslant 5$. **[2]**

(c) Find $E(X)$. **[3]**

(d) Find the median of X. **[3]**

(a) The region is a trapezium with parallel sides c and $4c$ and width 3: $\frac{1}{2}(c + 4c) \times 3 = \frac{15c}{2} = 1$, so $c = \frac{2}{15}$. **M1 A1**

(b) Gradient $= \dfrac{4c - c}{5 - 2} = c = \frac{2}{15}$. **M1** Through $(2, \frac{2}{15})$: $y - \frac{2}{15} = \frac{2}{15}(x - 2)$, so $f(x) = \frac{2}{15}(x-1)$. **A1**

(c)

$$E(X) = \frac{2}{15}\int_2^{5} (x^2 - x)\,dx = \frac{2}{15}\left[\frac{x^3}{3} - \frac{x^2}{2}\right]_2^{5} = \frac{2}{15}\left(\frac{175}{6} - \frac{2}{3}\right) = \frac{19}{5} = 3.8.$$

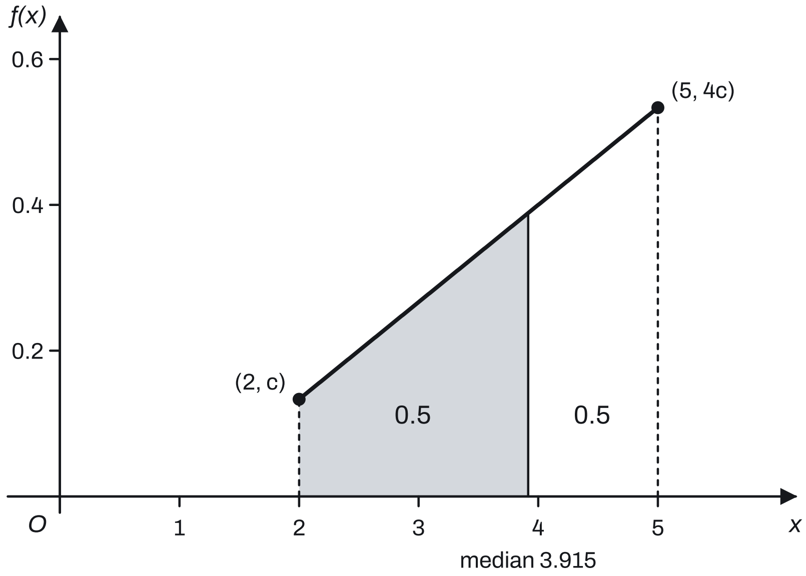
M1 for integrating $xf(x)$, **A1** for the correct integral with limits 2 and 5, **A1**.

(d)

$$\frac{2}{15}\int_2^m (x-1)\,dx = \frac{1}{15}\Big[(x-1)^2\Big]_2^m = \frac{(m-1)^2-1}{15} = \frac{1}{2}.$$

M1 for integrating from 2 to m and setting it equal to $\frac{1}{2}$. So $(m-1)^2 = 8.5$ and $m = 1+\sqrt{8.5}$ (the negative root gives $m < 2$, outside the interval). **A1** for a correct equation, **A1** for $m = 3.92$ (3 s.f.).

The median is to the right of the middle of the interval (3.5), because the graph is higher on the right. The diagram below shows the answers: the line with $c = \frac{2}{15}$ and the median splitting the area into two halves.



Example 4

The random variable X has probability density function given by

$$f(x) = \begin{cases} k(x-2)^2 & 0 \leqslant x \leqslant 4, \\ 0 & \text{otherwise.} \end{cases}$$

(a) State the value of $E(X)$, giving a reason. **[1]**

(b) Find the value of k. **[2]**

(c) Find $\mathrm{Var}(X)$. **[3]**

(d) Find the value of d such that $P(X > d) = 0.2$. **[3]**

(e) Write down $P(4-d < X < d)$. **[1]**

(a) The graph is symmetrical about $x = 2$, so $E(X) = 2$. **B1**

(b) $k\left[\frac{(x-2)^3}{3}\right]_0^4 = k\left(\frac{8}{3}+\frac{8}{3}\right) = \frac{16k}{3} = 1$, so $k = \frac{3}{16}$. **M1 A1**

(c) Expand: $x^2(x-2)^2 = x^4 - 4x^3 + 4x^2$.

$$\int_0^4 x^2 f(x)\, dx = \frac{3}{16}\left[\frac{x^5}{5} - x^4 + \frac{4x^3}{3}\right]_0^4 = \frac{3}{16} \times \frac{512}{15} = \frac{32}{5}.$$

M1 for integrating $x^2 f(x)$. $\mathrm{Var}(X) = \frac{32}{5} - 2^2 = \frac{12}{5} = 2.4$. **M1** for subtracting 2^2, **A1**.

(d) d is above 2. $\dfrac{3}{16}\left[\dfrac{(x-2)^3}{3}\right]_d^4 = \dfrac{8-(d-2)^3}{16} = 0.2$. **M1** for the right-hand tail set equal to 0.2.

$(d-2)^3 = 4.8$, **A1**, so $d = 2 + \sqrt[3]{4.8} = 3.69$ (3 s.f.). **A1**

(e) By symmetry $P(X < 4-d) = 0.2$ too, so $P(4-d < X < d) = 1 - 0.2 - 0.2 = 0.6$. **B1**

Common mistakes

- **Integrating $f(x)$ when the mean needs $xf(x)$**, or multiplying by x^2 for the mean. Examiners see this every year. Write $E(X) = \int xf(x)\,dx$ before you start, and $\int x^2 f(x)\,dx$ only for the variance.
- **Forgetting to subtract $\{E(X)\}^2$** at the end of a variance.
- **Losing the constant or a minus sign.** Keep k (or $\frac{3}{16}$) in front throughout, and keep the brackets: $-k(x-2)(x-6)$ and $c - d(x^2 - x)$ both need care when expanding.
- **Wrong limits.** Integrate only over the interval where f is not zero (not from 0 if it starts at 1), and don't swap the limits round.
- **Giving the mean of a symmetrical pdf as "the middle of the end points".** The reason is **symmetry of the graph**; say so.
- **Integrating when a sketch would do.** Many symmetry and straight-line questions need only areas of shapes. Draw the graph, label the known areas, and add or subtract them.
- **Writing $f(x)$ from a graph with the wrong sign of gradient or intercept.** Check your line against the picture: does it slope the right way, and cross where the diagram shows?
- **Keeping a root outside the interval.** A median or percentile must lie inside the interval. Reject the others with a reason.
- **Assuming an area is symmetrical when it isn't.** Only intervals the same distance either side of the line of symmetry have equal areas. $P(\text{mean} < X < \text{median})$ or $P(X > 2.5)$ on a graph symmetrical about $x = 3$ needs working, not a guess from the picture.
- **Stopping one step early** in context questions, such as giving the probability for one ball when the question asks about the largest of three.
- **"Show that" with missing steps.** The answer is given, so the working must be complete, including the line with the limits substituted.
- **Saying "the area is 1" but not "so f can be a pdf"** when asked to verify; and $f(x) \geqslant 0$ must hold too.

How to get the marks

- **Write the integral with its limits** before integrating: the method mark is for integrating the right thing (f, xf or $x^2 f$) and the next mark is for the correct limits.

- **Show the substituted limits.** In "show that" answers, one line with the numbers in is expected before the given answer.
- **"Write down" or "state"** means no working is needed: use symmetry or the total area of 1.
- **"Without calculation" or "without performing an integration"** means use the graph, symmetry or an earlier answer; an integral may earn only part of the marks.
- **"Hence"** means use the previous part, for example the value of k or the probability just found.
- **"Exact"** means fractions, surds, π or $\ln$, no decimals. Otherwise give 3 significant figures.
- **Checking that a median lies between two numbers** needs both areas (or both values of the rearranged expression) worked out, and a sentence saying why the median is between them.
- **Reasons in words**: symmetry for a mean, "$f(x) \geqslant 0$ and total area 1" for a pdf, the context for a model's weakness.
- **Method marks follow through**: a wrong k used correctly still earns later method marks, so keep going.

Quick check

1. The random variable X has pdf $f(x) = k\sqrt{x}$ for $0 \leqslant x \leqslant 4$, and 0 otherwise. Find k and $E(X)$.
2. The random variable X has pdf $f(x) = \frac{1}{6}(x+1)$ for $1 \leqslant x \leqslant 3$, and 0 otherwise. Find the median of X, giving your answer correct to 3 significant figures.
3. The pdf of X is symmetrical about $x = 5$, and X takes values from 2 to 8 only. Given that $P(X < 4) = 0.3$, find $P(4 < X < 6)$ and $P(5 < X < 6)$.
4. The random variable X has pdf $f(x) = 4x^3$ for $0 \leqslant x \leqslant 1$, and 0 otherwise. Find $\mathrm{Var}(X)$.
5. Explain why $f(x) = \frac{1}{4}(3x-1)$ for $0 \leqslant x \leqslant 2$ (and 0 otherwise) is not a probability density function, even though the area under it is 1.

Answers

1. $k\int_0^4 x^{1/2}\,dx = k \times \frac{2}{3} \times 8 = \frac{16k}{3} = 1$, so $k = \frac{3}{16}$. $E(X) = \frac{3}{16}\int_0^4 x^{3/2}\,dx = \frac{3}{16} \times \frac{2}{5} \times 32 = \frac{12}{5} = 2.4$.
2. $\frac{1}{6}\int_1^m (x+1)\,dx = \frac{1}{12}\left[(x+1)^2\right]_1^m = \frac{(m+1)^2-4}{12} = \frac{1}{2}$, so $(m+1)^2 = 10$ and $m = \sqrt{10} - 1 = 2.16$ (3 s.f.). The other root is negative, outside the interval.
3. By symmetry $P(X > 6) = 0.3$, so $P(4 < X < 6) = 1 - 0.3 - 0.3 = 0.4$, and $P(5 < X < 6) = 0.5 - 0.3 = 0.2$.
4. $E(X) = \int_0^1 4x^4\,dx = \frac{4}{5}$ and $\int_0^1 4x^5\,dx = \frac{2}{3}$, so $\mathrm{Var}(X) = \frac{2}{3} - \frac{16}{25} = \frac{2}{75} = 0.0267$ (3 s.f.).
5. $f(0) = -\frac{1}{4}$, and $f(x) < 0$ for $0 \leqslant x < \frac{1}{3}$. A pdf can never be negative, so this is not a pdf.

Past questions to try

- **9709/62/M/J/24 Q7**: a constant, verifying the median lies between two values, and an exact mean by parts.
- **9709/62/O/N/24 Q6**: a model in context, the meaning of its end points, constants from the mean, and the median.
- **9709/62/M/J/23 Q7**: straight-line pdfs read from graphs, the mean and a median.
- **9709/62/F/M/23 Q3**: symmetry of the graph, a constant and the variance.