Cambridge (CIE) A Level Maths: Probability & Statistics 2

Your notes

Continuous Random Variables

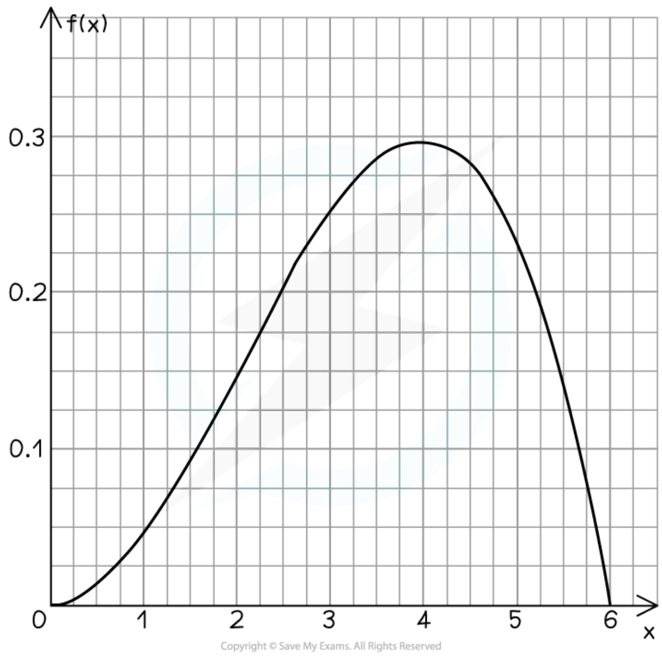
Contents

Your notes

Calculating Probabilities using PDF

What is a probability density function (p.d.f.)?

- For a **continuous random variable** , it is often possible to model probabilities using a function
  - This function is called a **probability density function (p.d.f.)**
  - For the continuous random variable,X , it would usually be denoted as a function of X (such as f(X) or g(X))
- The distribution (or **density**) of probabilities can be illustrated by the graph of f(X)
- The graph does not need to start and end on the x-axis



- For f(X) to **represent** a p.d.f. the following conditions must apply
  - $\mathrm{f}(x) \geq 0$ for **all** values of X
    - This is the equivalent to $P(X = x) \geq 0$ for a discrete random variable
  - The **area under the graph** must total 1

$$\int_{-\infty}^{\infty} \mathrm{f}(x)\, \mathrm{d}x = 1$$

- This is equivalent to $\Sigma\, \mathrm{P}(X = x) = 1$ for a discrete random variable

How do I find probabilities using a probability density function (p.d.f.)?

Your notes

- The probability that the continuous random variable X lies in the interval $a \leq X \leq b$, where X has the probability density function f(X) , is given by

$$P(a \leq X \leq b) = \int_a^b f(x)\ dx$$

- As with the **normal distribution** $P(a \leq X \leq b) = P(a < X < b)$
  - for any continuous random variable, $P(X = n) = 0$ for all values of n
  - One way to think of this is that $a = b$ in the integral above

How do I solve problems using the PDF?

- Some questions may ask for justification of the use of a given function for a probability density function
  - In such cases check that the function meets the two conditions $f(x) \geq 0$ for all values of X and the total area under the graph is 1
- If asked to find a probability
  - STEP 1
    Identify the **probability density function**, f(X), this may be given as a **graph**, an **equation** or as a **piecewise function**

    e.g. $f(x) = \begin{cases} 0.02x & 0 \leq x \leq 10 \\ 0 & \text{otherwise} \end{cases}$

- STEP 2
  Identify the **range** of X for a particular problem

  Remember that $P(a \leq X \leq b) = P(a < X < b)$

  Question: Can you explain why this is so?
  (Answer is at end of this section)

- STEP 3
  Sketching the **graph** of y = f(X) if **simple** may help to find the probability
  - Look for **basic shapes** such as **triangles** or **rectangles**; finding **areas** of these is easy and avoids integration
  - Look for **symmetry** in the graph that may make the problem easier
  - **Integrate f(X)** and evaluate it between the two limits for the **required** probability
- Trickier problems may involve finding a limit of the integral given its value
  - i.e. one of the values in the range of *X*, given the probability
    e.g. Find the value of a given $P(0 \leq X \leq a) = 0.09$

- Answer to question in STEP 2:
  Since $\mathrm{P}(X=a)=\mathrm{P}(X=b)=0$, $\mathrm{P}(a \le X \le b)=\mathrm{P}(a < X < b)$

Worked Example

The continuous random variable, X, has probability density function

$$\mathrm{f}(x)=\begin{cases} 0.08x & 0 \le x \le 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $\mathrm{f}(x)$ can represent a probability density function

a) There are two conditions for a function to be a p.d.f

$$\int_0^5 0.08x \, dx = \left[\frac{0.08x^2}{2}\right]_0^5$$

$$= \left[0.04x^2\right]_0^5$$

$$= 0.04 \times 5^2 - 0 = 1$$

$$\therefore \int_{-\infty}^{\infty} \mathrm{f}(x) \, dx = 1$$

Also, $\mathrm{f}(x) \geq 0$ for all values of x so $\mathrm{f}(x)$ meets both conditions to represent a probability density function

(b) Find

(i) $\mathrm{P}(0 \le X \le 2)$

(ii) $\mathrm{P}(X = 3.2)$

(iii) $\mathrm{P}(X > 4)$

b) Step 1: Identify the p. d. f. – given as a piecewise function in the question

(i) Step 2: Identify the range: 0 – 2

$$P(0 \leqslant X \leqslant 2) = \int_0^2 0.08x \ dx$$

$$= \left[0.04x^2\right]_0^2$$

Step 3: Evaluate the integral

$P(0 \leqslant X \leqslant 2) = 0.16$

(ii) $P(X = 3.2) = 0$ $P(X = n) = 0$ for all values of n

(iii) Step 2: Range is 4 – 5

$$P(X > 4) = \int_4^5 0.08x \ dx$$ You may use $1 - P(0 \leqslant X \leqslant 4)$

$$= \left[0.04x^2\right]_4^5$$

Step 3: Evaluate

$= 1 - 0.64$

$P(X > 4) = 0.36$

Examiner Tips and Tricks

- If the graph is easy to draw, then a sketch of f(x) is helpful
  - This can highlight useful features such as the graph (and so probabilities) being **symmetrical**
  - Some p.d.f. graphs lead to common shapes such as **triangles** or **rectangles** whose areas are easy to find, avoiding the need for integration

Median and Mode of a CRV

What is meant by the median of a continuous random variable?

- The **median**, ***m***, of a continuous random variable, X, with **probability density function f(x)** is defined as the value of the **continuous random variable** X, such that

$$P(X < m) = P(X > m) = 0.5$$

- Since $P(X = m) = 0$ this can also be written as $P(X \leq m) = P(X \geq m) = 0.5$
- If the p.d.f. is **symmetrical** (i.e. the graph of y = f(x) is symmetrical) then the median will be halfway between the lower and upper limits of x
  - In such cases the graph of y=f(x) has axis of symmetry in the line x = m

Your notes

Your notes

How do I find the median of a continuous random variable?

- By solving one of the equations to find m

$$\int_{-\infty}^{m} \mathrm{f}(x)\,\mathrm{d}x = 0.5$$

and

$$\int_{m}^{\infty} \mathrm{f}(x)\,\mathrm{d}x = 0.5$$

- The equation that should be used will depend on the information in the question
- If the graph of $y = \mathrm{f}(x)$ is **symmetrical**, symmetry may be used to deduce the median

How do I find quartiles (or percentiles) of a continuous random variable?

- In a similar way, to find the median
  - The lower quartile will be the value L such that $\mathrm{P}(X \le L) = 0.25$ or $\mathrm{P}(X \ge L) = 0.75$
  - The upper quartile will be the value U such that $\mathrm{P}(X \le U) = 0.75$ or $\mathrm{P}(X \ge U) = 0.25$
- Percentiles can be found in the same way
  - The 15^{th} percentile will be the value k such that $\mathrm{P}(X \le k) = 0.15$ or $\mathrm{P}(X \ge k) = 0.85$

What is meant by the mode of a continuous random variable?

- The **mode** of a continuous random variable, X, with **probability density function f(x)** is the value of x that produces the greatest value of f(x).

How do I find the mode of a PDF?

- This will depend on the type of function f(x); the easiest way to find the mode is by considering the shape of the graph of f(x)
- If the graph is a curve with a (local) **maximum point**, the mode can be found by **differentiating** and solving the equation f'(x) = 0
  - If there is more than one solution to f'(x) = 0, further work may be needed to deduce which answer is the mode
    - Look for valid values of from the **definition** of the p.d.f.
    - Use the **second derivative** (f''(x)) to deduce the nature of each **stationary point**
    - You may need to check the values of f(x) at the **endpoints** too

Your notes

Worked Example

The continuous random variable X has probability density function $f(x)$ defined as

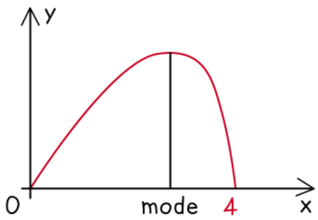
$$f(x) = \frac{1}{64}x(16 - x^2) \qquad 0 \leq x \leq 4$$

(a) Find the median of *X*, giving your answer to three significant figures

(b) Find the *exact* value of the mode of *X*

(a) Find the median of *X*, giving your answer to three significant figures

a) Sketch the graph



- Third root is $x = -4$
- Negative cubic: ↳↰
- Not necessarily symmetrical
- (Local) maximum

For the median, solve "$\int_0^m f(x)\ dx = 0.5$"

$$\frac{1}{64}\int_0^m (16x - x^3)\ dx = \frac{1}{2}$$

$$\left[8x^2 - \frac{1}{4}x^4\right]_0^m = 32$$

$$8m^2 - \frac{1}{4}m^4 = 32$$

$$m^4 - 32m^2 + 128 = 0$$

This is a 'hidden quadratic' equation in m^2
Using the formula

$m^2 = 16 \pm 8\sqrt{2}$

$\therefore m = 5.226\ 251...$ Out of range

or $m = 2.164\ 764...$

$\therefore$ Median $m = 2.16$ (3sf)

(b) Find the *exact* value of the mode of *X*

b) Differentiate, solving $f'(x) = 0$ to find the mode

$f'(x) = \frac{1}{64}(16 - 3x^2)$

$\therefore 16 - 3x^2 = 0$

$x^2 = \frac{16}{3}$

$x = \pm \frac{4\sqrt{3}}{3}$ Reject $x = -\frac{4\sqrt{3}}{3}$ since $0 \leqslant x \leqslant 4$

$\therefore$ Mode $= \frac{4}{3}\sqrt{3}$

Clearly from sketch of graph $x = \frac{4}{3}\sqrt{3}$ is a (local) maximum

Your notes

Examiner Tips and Tricks

- Avoid spending too long sketching the graph of $y = f(x)$, only do this if the graph is straightforward as finding the median and mode by other means can be just as quick

Your notes

E(X) & Var(X) (Continuous)

What are E(X) and Var(X)?

- E(X)is the **expected value**, or **mean**, of a **random variable X**
  - E(X) is the same as the population mean so can also be denoted by μ
- **Var (X)** is the **variance** of the continuous random variable X
  - **Standard deviation** is the **square root** of the **variance**

How do I find the mean and variance of a continuous random variable?

- The **mean**, for a **continuous random variable X** is given by

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x\mathrm{f}(x)\ \mathrm{d}x$$

- This is equivalent to $\Sigma x\mathrm{P}(X = x)$ for discrete random variables
- If the graph of $y = \mathrm{f}(x)$ has **axis** of **symmetry,** x = a , then E(X) = a
- The **variance** is given by

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} x^2\mathrm{f}(x)\mathrm{d}x - [\mathrm{E}(X)]^2$$

- This is equivalent to $\Sigma x^2\mathrm{P}(X = x) - [E(X)]^2$ for discrete random variables
- Be careful about confusing $E(X^2)$ and $[E(X)]^2$
  - $E(X^2) = \int_{-\infty}^{\infty} x^2\mathrm{f}(x)\ \mathrm{d}x$ "mean of the squares"
  - $[E(X)]^2 = \left[\int_{-\infty}^{\infty} x\ \mathrm{f}(x)\ \mathrm{d}x\right]^2$ "square of the mean"
  - If you are happy with the difference between these and how to calculate them the variance formula becomes very straightforward

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$$

How do I calculate E(g(X))?

- $\mathrm{E}(g(X)) = \int_{-\infty}^{\infty} g(x)\mathrm{f}(x)\ \mathrm{d}x$

- In particular:
  - $\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2 \mathrm{f}(x)\,\mathrm{d}x$ as seen above

Your notes

Worked Example

A continuous random variable, X, is modelled by the probability distribution function $\mathrm{f}(x)$, such that

$$\mathrm{f}(x) = \begin{cases} 1.5x^2(1-0.5x) & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

(i) Find $E(X)$

a) $\text{"}E(x) = \int_{-\infty}^{\infty} x\mathrm{f}(x)\,dx\text{"}$

$$E(x) = \int_0^2 1.5x^3(1-0.5x)\,dx$$

$$E(x) = \int_0^2 (1.5x^3 - 0.75x^4)\,dx$$

$$E(x) = \Big[0.375x^4 - 0.15x^5\Big]_0^2$$

$$E(x) = 6 - 4.8$$

$$E(x) = 1.2$$

(ii) Find $\mathrm{Var}(X)$

b) $\text{"}\mathrm{Var}(x) = \int_{-\infty}^{\infty} x^2 \mathrm{f}(x)\,dx - \big[E(x)\big]^2 \text{ "or" } E(x^2) - \big[E(x)\big]^2\text{"}$

$$E(x^2) = \int_0^2 (1.5x^4 - 0.75x^5)\,dx$$

$$= \Big[0.3x^5 - 0.125x^6\Big]_0^2 = 1.6$$

$$\therefore \mathrm{Var}(x) = E(x^2) - \big[E(x)\big]^2$$

$$= 1.6 - (1.2)^2$$

$$\mathrm{Var}(x) = 0.16$$

Your notes

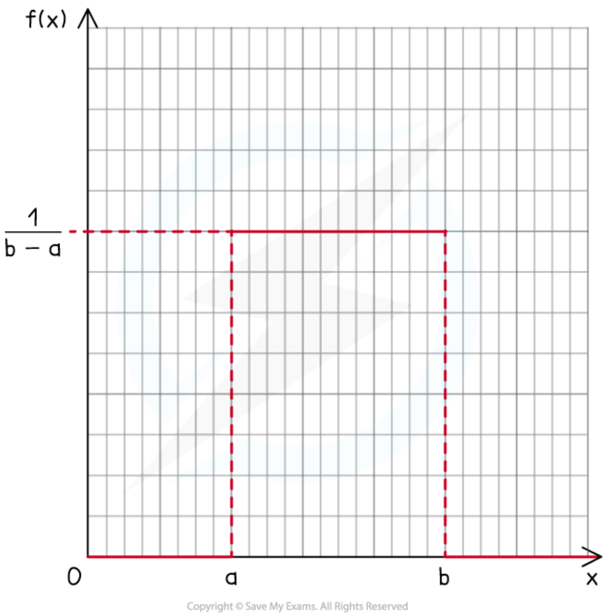
Examiner Tips and Tricks

- A **sketch** of the graph of $y = f(x)$ can highlight any **symmetrical** properties which can help reduce the work involved in finding the **mean** and **variance**
- Take care with awkward **values** and **negatives** – use the **memory** features on your calculator and **avoid rounding** until your **final** answer (if rounding at all!)

Continuous Uniform Distribution

What is meant by the continuous uniform distribution?

- This is a special case of a probability density function for a continuous random variable
  - The **normal distribution** is another special case covered in S1
- The **uniform**, or **rectangular**, **distribution** is a p.d.f. that is **constant** and **non-zero** over a range of values but zero everywhere else



- Since the area under the graph has to total 1, the height of the uniform distribution would be

$$\frac{1}{b-a}$$

- Therefore the probability density function is given by

$$\mathrm{f}(x)=\begin{cases}\dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise}\end{cases}$$

How do I find probabilities for a continuous uniform distribution?

- **Sketch** the graph of y= f(x)

- Probabilities are the **area under** the **graph**, all such **areas** will now be **rectangles**
  - Finding the area of a rectangle is likely to be easier than integration!
- The **symmetrical** properties of rectangles may also be used to find probabilities

How do I find the mean, median, mode and variance of a continuous uniform distribution?

- The **mean**, or expected value, is given by

$$\mathrm{E}(X) = \frac{1}{2}(a+b)$$

- This is the (vertical) **axis** of **symmetry** of the **rectangle**
- Should the above be forgotten, $E(X) = \int_a^b x\mathrm{f}(x)\,\mathrm{d}x$ can still be applied
  - You be may asked to use this to prove the result
- The **median** can also be found by **symmetry** and will be **equal** to the **mean**
- There is **no mode** as f(x) is equal - and so at its greatest - for all values of x
- The **variance** is given by

$$\mathrm{Var}(X) = \frac{1}{12}(b-a)^2$$

- Should the above be forgotten, $\mathrm{Var}(X) = \int_{-\infty}^{\infty} x^2\mathrm{f}(x)\,\mathrm{d}x - [E(X)]^2$ or $Var(X) = E(X^2) - [E(X)]^2$ can still be applied
  - You may be asked to use this to prove the result
  - The **standard deviation** is the **square root** of the **variance**

Worked Example

A continuous random variable, X, is modelled by the uniform distribution such that $\mathrm{f}(x) = 0.4$ for $a \le x \le 4$ and $\mathrm{f}(x) = 0$ otherwise.
a is a constant.

(a) Show that the value of a is 1.5 .

Your notes

a) "$f(x) = \frac{1}{b-a}$" for all values of x

$\frac{1}{4-a} = 0.4$

$1 = 0.4(4 - a)$

$0.4a = 1.6 - 1$

$a = \frac{0.6}{0.4}$

$\therefore a = 1.5$

(b) Find

(i) $P(2.5 \le X \le 3)$

(ii) $E(X)$

b)(i) Shade your graph from earlier

$P(2.5 \leqslant X \leqslant 3) = 0.5 \times 0.4 = 0.2$

$P(2.5 \leqslant X \leqslant 3) = 0.2$

(ii) "$E(X) = \frac{1}{2}(a + b)$"

$E(X) = \frac{1}{2}(1.5 + 4) = 2.75$

$E(X) = 2.75$

(c) Find the standard deviation of *X*, giving your answer in the form $a\sqrt{3}$, where a is a rational number.

c) "$\text{Var}(X) = \frac{1}{12}(b - a)^2$"

$\text{Var}(X) = \frac{1}{12}(4 - 1.5)^2$

$\text{Var}(X) = \frac{6.25}{12} = \frac{25}{48}$

$\text{St. Dev.} = \sqrt{\frac{25}{48}} = \frac{5\sqrt{3}}{12}$

Standard deviation is $\frac{5}{12}\sqrt{3}$

Examiner Tips and Tricks

- A **sketch** of the graph of a uniform distribution is quick and will highlight the **symmetry** in a uniform distribution
- Use **areas** of **rectangles** to find **probabilities** rather than integrating

Your notes