Session 1 - Continuous Probability

Year 12 Advanced Mathematics

Objectives:
- Identify and verify the conditions required for a valid **Probability Density Function (PDF)**.
- Distinguish between **Probability Density Functions (PDFs)** and **Cumulative Distribution Functions (CDFs)**
- Apply PDFs and CDFs to calculate probabilities and solve problems involving quantiles (e.g., medians, percentiles).

Context

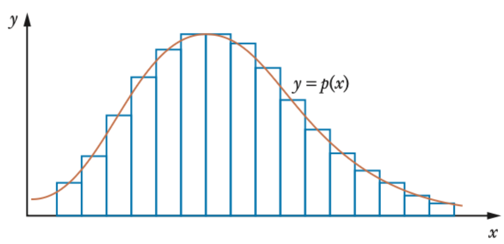
We need to have an understanding of some concepts from data,functions and integration.

Data:

- Discrete - represents figures that can be counted, and there are defined gaps/precision between values.
    - EG: number of students in a class.
- *Continuous* - Represents measurements that can take any value within a given range. Rather, values that fluctuate.
    - EG: Temperature of a classroom
        - We can consider temperature as a **continuous random variable.**

Functions/Integration:

- Continuous data can be represented through a continuous curve.
- The probability of a given range can be calculated by finding the area under the curve, through integration.
- We generally integrate a curve given by a function between a boundary, like a parabola between x = 0 and 2 etc.



Probability Density Functions (PDF)

PDF’s are the graphical representation of continuous data as a curve.

PDF represents the density at a point, meaning how likely a value is to occur in that range.

The function provided to you *f(x)* is the probability density function *p(x)*

They have two specific qualities that make them PDF’s.

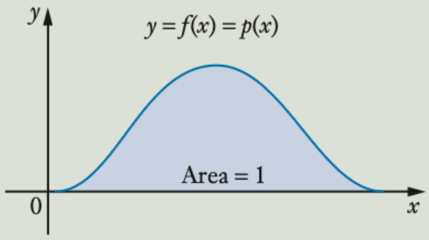
1. The area under the whole curve is equal to 1, as all probabilities add to 1
2. The curve is positive, meaning it exists only in quadrant 1.

Area under a probability density function

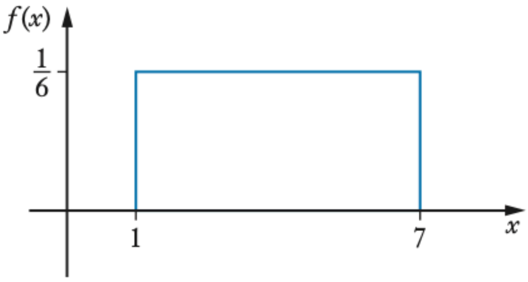
The area under a probability density function is 1.

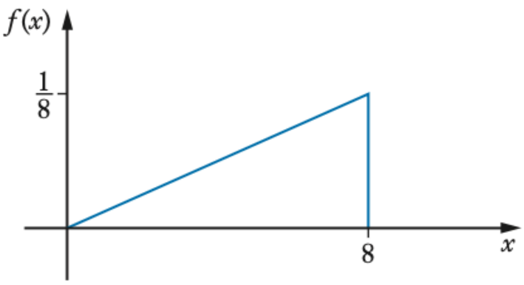
$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

where $f(x) \geq 0$ (since $0 \leq p(x) \leq 1$)



Proving a PDF EG:





Questions about PDF's EG:

a A function is given by $f(x) = \begin{cases} \dfrac{3x^2}{26} & \text{for } 1 \le x \le 3 \\ 0 & \text{for all other } x \end{cases}$.

Show that it is a continuous probability distribution.

b A function is given by $f(x) = ax^2$ defined for the domain $[0, 5]$.
Find the value of a for which this is a probability density function.

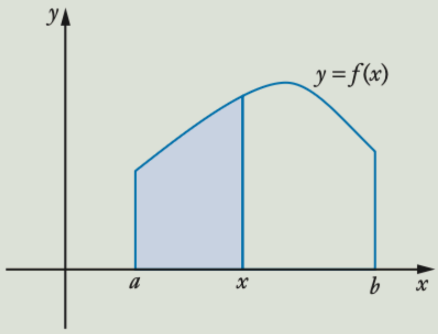
A probability density function has the equation $f(x) = \dfrac{x^4}{3355}$ over the domain $[2, b]$. Evaluate b.

Calculating Probabilities in PDF's

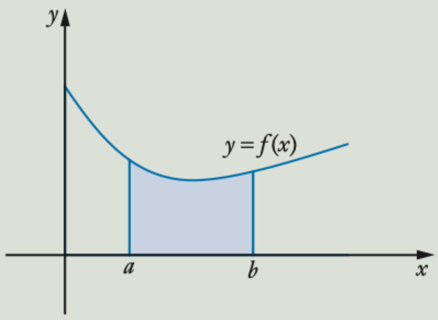
To calculate probabilities in a PDF, it is integrating the curve between a given boundary. That given boundary is the probability range in which we are calculating for.

Probabilities in probability density functions

$P(X \leq x) = \int_a^x f(x)\,dx$ where $y = f(x)$ is a PDF defined in the domain $[a, b]$.



$P(a \leq X \leq b) = \int_a^b f(x)\,dx$ where $y = f(x)$ is a PDF and a and b are in the defined domain.

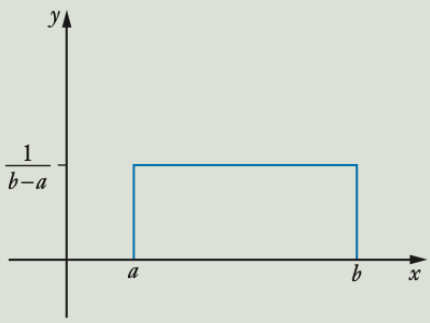


Uniform Distributions

Every outcome has the same probability, meaning the function is linear and creates a polygon in which we can find the area of, generally speaking.

Uniform continuous probability distributions

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{for all other } x \text{ values} \end{cases}$$



Equal intervals along the x-axis will have the same probability.

The continuous probability distribution is defined by $f(x) = ax^2$ in the domain [0, 5].

a Evaluate a.

b Find:

i $P(X \leq 3)$ **ii** $P(1 < X < 4)$ **iii** $P(X > 2)$

Cumulative Distribution Functions (CDF)

A CDF is similar to a PDF, where they now measure the total probability up to a particular value. It is like the running total of the area under the pdf.

Since it is values up to a specific value, we will always integrate between a top boundary of x as we can substitute that value in, and then the lower boundary of the given value.

Cumulative distribution function (CDF)

The cumulative distribution function is given by $F(x) = \int_a^x f(x)\,dx$ where $y = f(x)$ is a PDF defined in the domain $[a, b]$.

Questions will generally ask you to:

1. Find the cumulative distribution function, to do that we integrate between x and a
2. Use the CDF to find probabilities up to that point.

EG:

A continuous probability function is given by $f(x) = \dfrac{4x^3}{255}$ defined in the domain $[1, 4]$.

a Find the cumulative distribution function.

b Use the CDF to find:

i $P(X \le 3)$ **ii** $P(X < 1.6)$

Mode

The mode of a curve is the maximum point. Therefore, by deriving the function, we can solve for the max point and prove a stationary point.

When answering a mode question, we must:

1. Derive the function to find the gradient
2. Make the derivative equal to 0 [f'(x) = 0]
3. Solve for the mode.
4. Now we need to prove its max, so we find the second derivative.
5. If the second derivative is less than zero [f''(x) < 0]

EG:

A continuous probability distribution is defined on the interval $1 \le x \le 5$ and has equation $f(x) = \dfrac{3x(6-x)}{92}$. Find the mode.

Quantiles

Quantiles refer to a specific cumulative quantity at a single point.

Examples include: Median (0.5), Percentile (every 0.01), Decile (0.1), Quartile (0.25)

To answer a question asking us to find a quantile, we must:

1. Find the CDF by integrating between x and a.
2. Make the CDF equal to the quantile.
3.

A continuous probability distribution is defined as $f(x) = \dfrac{x^4}{11\,605}$ in the domain [4, 9].

Find, correct to 2 decimal places:

a the 1st quartile

b the 38th percentile

c the 7th decile

HOMEWORK

10.01 - PDF

5 Given the continuous probability distribution $f(x) = \begin{cases} kx^3 & \text{for } 0 \le x \le 5 \\ 0 & \text{for all other } x \end{cases}$

find the value of k.

6 A probability density function is given by $f(x) = ae^x$ over a certain domain. Find the exact value of a if the domain is:

a $[1, 3]$ **b** $[1, 7]$ **c** $[0, 4]$

7 A function is given by $f(x) = \frac{x^2}{72}$.

Over what domain starting at $x = 0$ is this a probability density function?

8 A PDF is given by $f(x) = \frac{2x^5}{87\,381}$ over the interval $1 \le x \le b$. Find the value of b.

10.02 - Calculating Probabilities

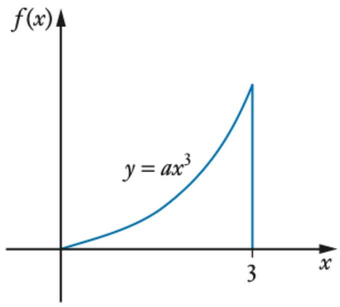
4 The continuous random variable X has the PDF shown.

a Evaluate a.

b Find:

i $P(1 \le X \le 3)$ **ii** $P(X < 2)$

iii $P(1 \le X \le 2)$ **iv** $P(X \le 1)$



5 A continuous probability function is given by $f(x) = ke^x$, defined on the domain $[1, 6]$.

a Find the exact value of k.

b Find each exact probability:

i $P(2 \le X \le 5)$ **ii** $P(X < 4)$ **iii** $P(X \ge 3)$

6 **a** Show that $y = \sin x$ is a probability density function in the domain $\left[0, \frac{\pi}{2}\right]$.

b Find each exact probability:

i $P\left(X \le \frac{\pi}{3}\right)$

ii $P\left(0 < X < \frac{\pi}{4}\right)$

iii $P\left(X > \frac{\pi}{6}\right)$

10.03 - Cumulative Distribution Function

1 Find the cumulative distribution function for each continuous probability distribution.

a $f(x) = \dfrac{x^2}{9}$ defined in the domain [0, 3]

b $f(x) = \dfrac{4x^3}{1296}$ defined in the domain [0, 6]

c $f(x) = \dfrac{e^x}{e^4 - 1}$ in the interval $0 \le x \le 4$

d $f(x) = \dfrac{4(x-2)^3}{625}$ in the domain [2, 7]

e $f(x) = \dfrac{3x(8-x)}{135}$ in the domain [2, 5]

2 **a** Find the cumulative distribution function for $f(x) = \begin{cases} \dfrac{5x^4}{7776} & \text{for } 1 \le x \le 6 \\ 0 & \text{for all other values} \end{cases}$.

b Find:

i $P(X \le 3)$ **ii** $P(X \le 2)$ **iii** $P(X < 5)$

iv $P(X > 4)$ **v** $P(2 \le X \le 4)$

3 A continuous probability distribution is given by $f(x) = \dfrac{4x^3}{2320}$ in the domain [3, 7].

a Find the cumulative distribution function.

b Find:

i $P(X \le 4)$ **ii** $P(X \le 6)$ **iii** $P(X \ge 5)$

iv $P(X > 4)$ **v** $P(4 \le X < 6)$

10.04 - Quantiles

1 Find the median of each continuous random variable correct to 2 decimal places.

a $f(x) = \dfrac{3x^2}{511}$ defined on the interval $1 \le x \le 8$

b $f(x) = \dfrac{4x^3}{2401}$ defined in the domain [0, 7]

c $f(x) = \dfrac{5x^4}{16\,807}$ in the interval $0 \le x \le 7$

d $f(x) = \dfrac{3(x-3)^2}{16}$ in the domain [1, 5]

2 For each continuous probability distribution, find:

i the 1st quartile **ii** the 2nd decile **iii** the 77th percentile

a $f(x) = \dfrac{3x^2}{973}$ defined in the domain [3, 10]

b $f(x) = \dfrac{x^3}{324}$ defined in the interval $0 \leq x \leq 6$

c $f(x) = \dfrac{5x^4}{3124}$ defined in the interval $1 \leq x \leq 5$