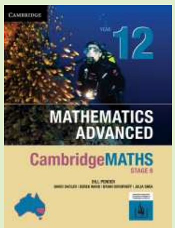
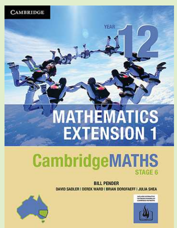
NORMANHURST BOYS HIGH SCHOOL

MATHEMATICS ADVANCED
MATHEMATICS EXTENSION 1/2
(YEAR 12 COURSE)

Topic summary and exercises:

With references to





(A) (X1) (X2) **Statistics/Continuous Probability Distributions**

Name: ....

Initial version by H. Lam, June 2020. Last updated August 10, 2021.
Various corrections by students & members of the Department of Mathematics at Normanhurst Boys High School.

Acknowledgements Pictograms in this document are a derivative of the work originally by Freepik at `http://www.flaticon.com`, used under 🅭 CC BY 2.0.

Symbols used

- Beware! Heed warning.
- Enrichment - not necessarily in the syllabus, or any formal assessment.
- Ⓐ Mathematics Advanced content.
- Ⓧ1 Mathematics Extension 1 content.
- Literacy: note new word/phrase.

$\mathbb{N}$ the set of natural numbers

$\mathbb{Z}$ the set of integers

$\mathbb{Q}$ the set of rational numbers

$\mathbb{R}$ the set of real numbers

$\forall$ for all

Syllabus outcomes addressed

MA12-8 solves problems using appropriate statistical processes

Syllabus subtopics

MA-S3 Random Variables

Gentle reminder

- For a thorough understanding of the topic, *every* question in this handout is to be completed!
- Additional questions from *CambridgeMATHS Year 12 Advanced* (Pender, Sadler, Ward, Dorofaeff, & Shea, 2019a) or *CambridgeMATHS Year 12 Extension* (Pender, Sadler, Ward, Dorofaeff, & Shea, 2019c) will be completed at the discretion of your teacher.
- Remember to copy the question into your exercise book!

Contents

Section 1

Probability Distributions

1.1 Discrete probability distributions

Definition 1

A **random variable** is a ....numerical........ ....quantity........ whose value depends on the outcome of a chance experiment.

Fill in the spaces

When one particular aspect or attribute of a given population S can be measured by a ....number........ (integer, fraction, or any real number), we call that attribute a ....random........ ....variable........ . Officially, a random variable X is a function from a population S to numbers.

Example 1

(Pender, Sadler, Ward, Dorofaeff, & Shea, 2019b, p.622) State whether each probability distribution is *numeric* or *categorical*. If it is numeric, state whether it is *discrete* or *continuous*.

1. The number showing when a die is thrown

2. The weight of a randomly chosen adult male in Australia

3. Whether it rains or not on a spring day in Sydney

4. The daily rainfall in Sydney on a September day

5. The colour of a ball drawn from a bag containing four red and three green balls

6. The colours of two balls drawn together from a bag containing four red and three green balls

7. The shoe size of a randomly chosen adult female in Australia

8. ATAR results for a particular year

Answer: **1.** Numeric/discrete **2.** Numeric/continuous **3.** Categorical **4.** Numeric/continuous **5.** Categorical **6.** Categorical **7.** Numeric/can be both discrete and continuous **8.** Numeric/discrete

1.2 Grouped data and relative frequencies

Learning Goal(s)

Knowledge
Discrete probabilities

Skills
Find quantiles and analyse results

Understanding
Histograms

By the end of this section am I able to:

32.1 Use relative frequencies and histograms obtained from data to estimate probabilities associated with a continuous random variable

Laws/Results

The relative frequency of an event occurring is an ....estimate.......... of the ....probabilities.......... .

Example 2

[2008 VCE Mathematical Methods Paper 1 Q7] Jane drives to work each morning and passes through three intersections with traffic lights. The number X of traffic lights that are red when Jane is driving to work is a random variable with probability distribution given by

x	0	1	2	3
$P(X = x)$	0.1	0.2	0.3	0.4

(a) What is the mode of X? **1**

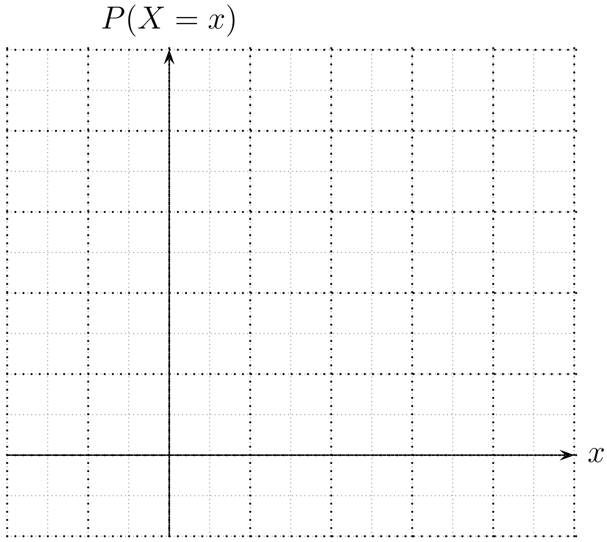
(b) Jane drives to work on two consecutive days. What is the probability that the number of traffic lights that are red is the same on both days? **2**

Answer: (a) 3 (b) 0.3

1.2.1 Relative frequency histogram and polygon

Refer to the table in Example 2, plot the *relative frequency histogram* and *relative frequency polygon*.

- No gaps between each bar.
- Polygon data point at the ...centre... of each histogram rectangle.
  - Commence and end at one score value below/above respectively, outside of any rectangles.



Laws/Results

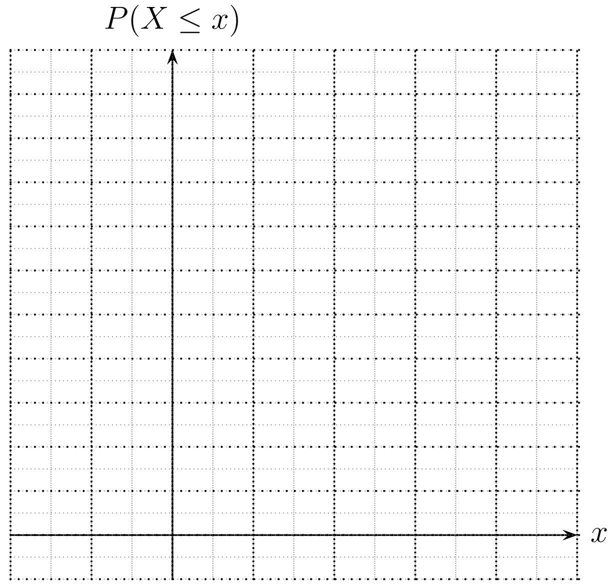
The area of rectangles in the relative frequency histogram ...sum... to ...1... when the width of the rectangles is ...1... .

1.2.2 Cumulative relative frequency histogram and polygon

Refer to the table in Example 2, plot the *cumulative relative frequency histogram* and *cumulative relative frequency polygon.*

x	0	1	2	3
$P(X = x)$	0.1	0.2	0.3	0.4
$P(X \leq x)$				

- No gaps between each bar.
- Polygon data point at the right of each histogram rectangle.
  - Commence and end at $\frac{1}{2}$-a-score's value below/above respectively, inside the lowest score/highest score's rectangle.



1.2.3 Quartiles, deciles and percentiles

Definition 2

Quartiles divide the scores into four equal sets.

Steps

Finding the value of the first quartile

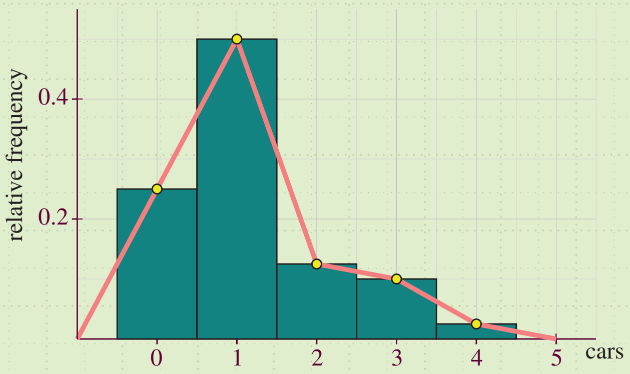
1. On a *cumulative relative frequency polygon*, find the location where cf_r = 0.25 (corresponding to Q_1)
2. Draw the horizontal line from the vertical axis where cf_r = 0.25 to the polygon
3. Draw the vertical line from the polygon to the horizontal axis to estimate the x value of the first quartile.

Important note

The *median* of the dataset occurs when $cf_r = \frac{1}{2}$. .

Example 3

[(A) **Ex 10A**/(x1) **16A Q9**] For town-planning purposes, the number of cars owned by each household in a suburb was recorded from census data. The results are displayed in the relative frequency histogram and polygon below. (This is a population, so the relative frequencies are probabilities.)



(a) What fraction of the households have no cars?

(b) What fraction of the households have fewer than 2 cars?

(c) What is the probability that a household chosen at random has three cars?

(d) Town planners will advise that additional on-street parking be provided if more than 40% of the households have 3 or more cars. Will they be advising that additional parking be provided? Explain your answer.

(e) Complete the following table for this probability distribution.

x	0	1	2	3	4
$P(X = x)$					

(f) Show that the sum of the probabilities is 1. How is this related to the area of the rectangles of the histogram?

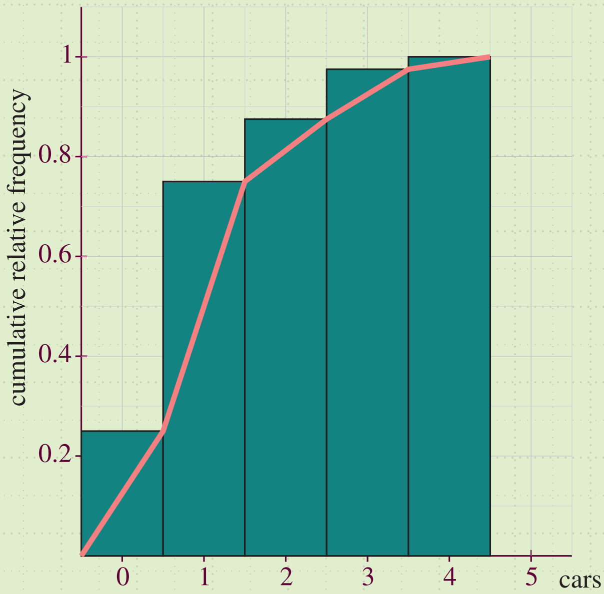
(g) Explain in your own words, and with reference to the graph above, why the area bounded by the relative frequency polygon and the horizontal axis will be the same as the area of the relative frequency histogram.

(h) Use your table to show that the mean number of cars per household is 1.15. What do you understand by this answer — how can a household have a fraction of a car?

(i) A street in the suburb is selected at random. If there are 100 households in the street, how many cars would you expect to belong to the households in the street in total? Are your assumptions for this estimate reasonable?

Example 3

[(A) Ex 10A/(X1) 16A Q9] (continued)



(j) Complete the following table for the cumulative relative frequencies of this probability distribution.

x	0	1	2	3	4
$P(X \leq x)$					

(k) A town planner constructs a cumulative relative frequency polygon and histogram from these data. His graph is shown above. Confirm that your data agree with this graph.

(l) By drawing horizontal lines at heights 0.25, 0.5 and 0.75, find the three quartiles Q_1, Q_2 and Q_3.

Further exercises

(A) Ex 10A

- Q10-11

(X1) Ex 16A

- Q10-11
- Q12

Additionally,

- Q1-5 (if a review of Discrete Probability Distributions is warranted)

1.3 Continuous probability distributions

1.3.1 Probability density function (PDF)

Learning Goal(s)

Knowledge
Area under a curve corresponding to probability

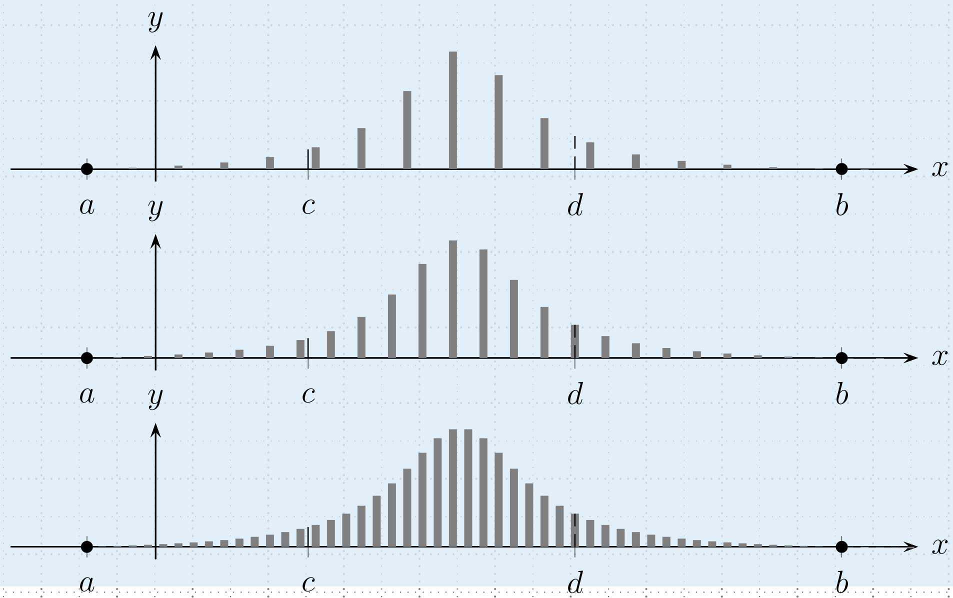
Skills
Find areas beneath curves via integration techniques

Understanding
Difference between discrete and continuous probability distribution

By the end of this section am I able to:
32.2 Understand and use the concepts of a probability density function of a continuous random variable

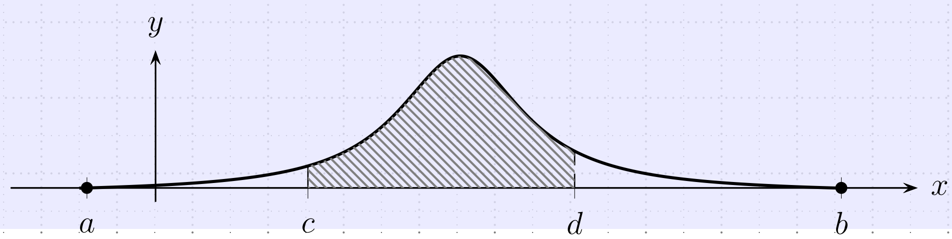
Fill in the spaces

As the number of values of discrete distribution increases, the ...graph... of the distribution may start to resemble a ...curve... .



Definition 3

Such a curve, is the ...probability... ...density... ...function... $f_X(x)$. It describes the ...probability... ...distribution... for a ...continuous... set of data.



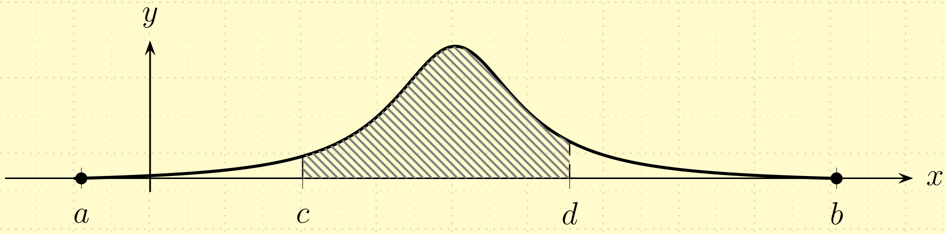
Important note

- The analogue in the discrete case: ...$P(X = x)$...
- ⚠ $f_X(x)$ is often abbreviated to $f(x)$.

Laws/Results

Properties For a continuous random variable X on the domain $x \in [a, b]$, such that

$$P(c \leq X \leq d) = \int_c^d f(x)\, dx$$



- $f(x) \geq$ 0 $\quad \forall x \in [a, b]$.

  Reason: ....

- $\int_a^b f(x)\, dx =$ 1 .

  Reason: ....

- $P(X = x) =$ 0 .

  Reason: ....

- Mode of the distribution: the x value where the local maximum occurs.

Example 4

(Haese, Haese, & Humphries, 2016, p.226) Consider the probability density function

$$f(x) = \frac{x^2}{9} \quad 0 \le x \le 3$$

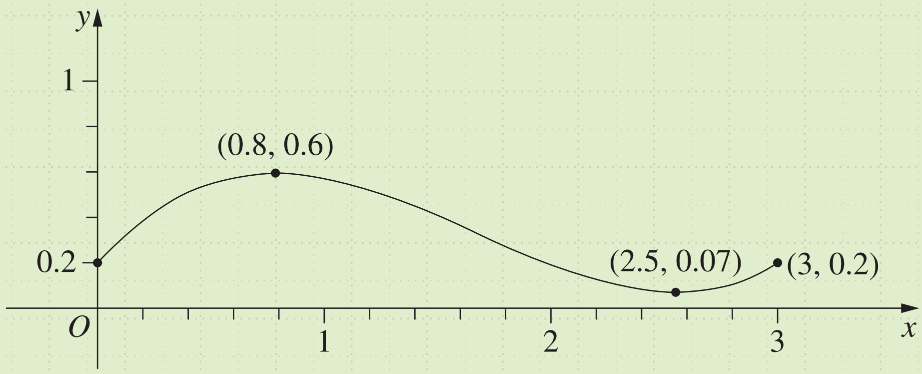
(a) Draw a sketch of $f(x)$ within its defined domain.

(b) Check that $f(x)$ is a valid probability density function.

(c) Find $P(1 \le X \le 2)$.

Example 5

[2020 Mathematics Advanced Sample HSC Q7] The diagram shows the graph of a continuous probability density function.



Which of the following is the mode?

(A) 0.07 (B) 0.6 (C) 0.8 (D) 3

1.3.2 Cumulative distribution function (CDF)

Definition 4

The *cumulative distribution function*, $F(x)$, is the ......signed...... ......area...... function.

$$P(X \le x) = F(x) = \int_a^x f(t)\,dt \quad a \le x \le b$$

Example 6

[2012 VCE Mathematical Methods Paper 1 Q8] The probability density function f of a random variable X is given by

$$f(x) = \begin{cases} \dfrac{x+1}{12} & 0 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of b such that $P(X \le b) = \dfrac{5}{8}$.

Answer: $b = 3$

Example 7

[2010 VCE Mathematical Methods Paper 1 Q10] The continuous random variable X has a distribution with probability density function given by

$$f(x) = \begin{cases} ax(5-x) & 0 \le x \le 5 \\ 0 & x < 0 \text{ or } x > 5 \end{cases}$$

where a is a positive constant.

(a) Find the value of a. **3**

(b) Express $P(X < 3)$ as a definite integral. (Do **not** evaluate the definite integral). **1**

Answer: $\frac{6}{125}$

Example 8

[2008 VCE Mathematical Methods Paper 2 Q15] A probability density function f, is given by

$$f(x) = \begin{cases} \frac{1}{12}\left(8x - x^3\right) & 0 \le x \le 2 \\ 0 & \text{elsewhere} \end{cases}$$

The median, m, of this function satisfies the equation:

(A) $-m^4 + 16m^2 - 6 = 0$

(B) $-m^4 + 4m^2 - 6 = 0$

(C) $m^4 - 16m^2 = 0$

(D) $m^4 - 16m^2 + 24 = 0.5$

(E) $m^4 - 16m^2 + 24 = 0$

Example 9

[2008 VCE Mathematical Methods Paper 1 Q4] The function

$$f(x) = \begin{cases} k\sin(\pi x) & x \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

is a probability density function for the continuous random variable X.

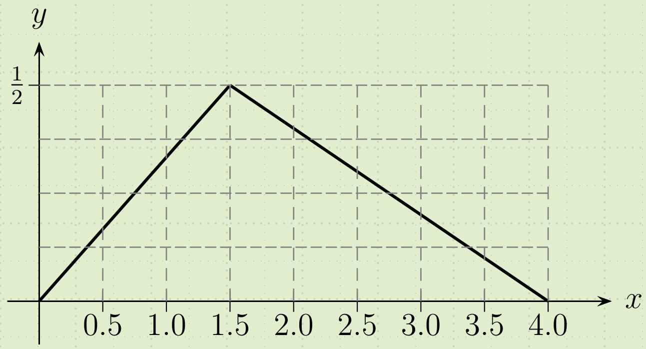
(a) Show that $k = \frac{\pi}{2}$. **2**

(b) Find $P\left(X \le \frac{1}{4} \middle| X \le \frac{1}{2}\right)$. **3**

Answer: $\frac{2-\sqrt{2}}{2}$

Example 10

[2019 WACE Mathematics Methods Section 1 Q3] Waiting times for patients at a hospital emergency department can be up to four hours. The associated probability density function is shown below.



(a) What is the probability a patient will wait less than one hour? **3**

(b) What is the probability a patient will wait between one hour and three hours? **4**

⚠ WACE marking schemes are quite different from HSC - and may attract more marks per question.

Answer: (a) $\frac{1}{6}$ (b) $\frac{11}{15}$

Example 11

[2020 Mathematics Advanced Sample HSC Q31] A bid made at an auction for a real estate property, in millions of dollars, can be modelled by the random variable X with the probability density function

$$f(x) = \begin{cases} k\left(16 - x^2\right) & 1 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that the value of k is $\dfrac{1}{27}$. **2**

(b) ⚠ Find the cumulative distribution function. **2**

(c) Find the probability that a bid of more than 3 million dollars will be made. **1**

Answer: (a) Show (b) $F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{81}\left(48x - x^3 - 47\right) & 1 \le x \le 4 \\ 1 & x > 4 \end{cases}$ (c) 0.1358

Important note

- What happens outside of the domain specified? Remember the CDF represents $P(X \le x)$!
- Don't forget the ....constant(s)........ of integration!

Example 12

A probability density function is defined piecewise by

$$f(x) = \begin{cases} k(4+x) & -4 \le x \le 0 \\ k(4-x) & 0 \le x \le 4 \end{cases}$$

(a) Find the value of the constant k. Hence write the equation of $f(x)$ and sketch it.

(b) What is the probability that $0 \le X \le 2$?

(c) Why is the median zero, and what is the mode?

(d) Find the CDF for $-4 \le x \le 0$, and the CDF for $0 \le x \le 4$. Then sketch the whole CDF.

1.3.3 Uniform distribution

Definition 5

A *uniform distribution*'s probability density function is a ....constant........ .

Example 13

Given each of the following for $0 \le x \le 10$:

(a) $f(x) = k$

(b) $f(x) = kx$

i. Find the value k that makes the function a probability density function:

ii. Find the corresponding CDF $F(x)$.

iii. Find the median and quartiles.

Answer: (a) i. $\frac{1}{10}$ ii. $\frac{1}{10}x$ iii. $Q_1 = \frac{5}{2}$, $Q_2 = 5$, $Q_3 = \frac{15}{2}$ (b) i. $\frac{1}{50}$ ii. $\frac{1}{100}x^2$ iii. $Q_1 = 5$, $Q_2 = 5\sqrt{2}$, $Q_3 = 5\sqrt{3}$

Example 14

[**2019 WACE Mathematics Methods Section 1 Q6**] The error X in digitising a communication signal has a uniform distribution with probability density function given by

$$f(x) = \begin{cases} 1 & -0.5 < x < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of $f(x)$. **2**

(b) What is the probability that the error is at least 0.35? **1**

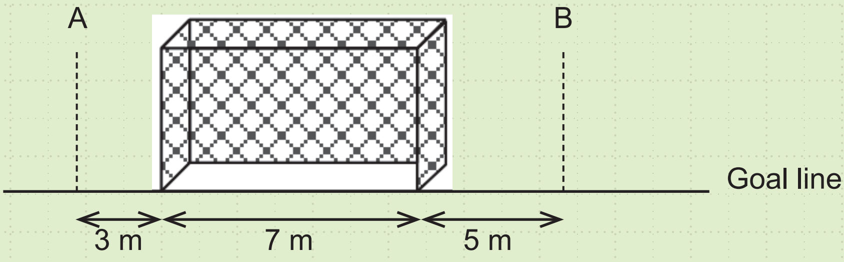
(c) If the error is negative, what is the probability that it is less than -0.35? **2**

(d) An engineer is more interested in the square of the error. What is the probability that the square of the error is less than 0.09? **2**

Answer: (a) Draw (b) 0.15 (c) 0.3 (d) 0.6

Example 15

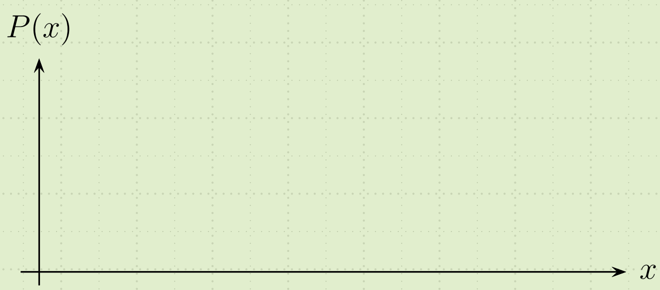
[2017 WACE Mathematics Methods Section 1 Q2] Michelle is a soccer goalkeeper and has built a machine to help her practise. The machine will shoot a soccer ball randomly along the ground at or near a goal that is seven metres wide. The machine is equally likely to shoot the ball so that the centre of the ball crosses the goal line anywhere between point A three metres left of the goal, and point B five metres right of the goal, as shown in the diagram below.



Michelle sets up a trial run without anyone in the goals. Assume the goal posts are of negligible width.

Let the random variable X be the distance the centre of the ball crosses the goal line to the right of point A.

(a) Complete the graphical representation of the probability density function for the random variable X. **2**



(b) What is the probability that the machine shoots a ball so that its centre misses the goal to the left? **1**

(c) What is the probability that the machine shoots a ball so that its centre is inside the goal? **1**

(d) If the machine shoots a ball so that its centre misses the goal, what is the probability that the ball's centre misses to the right? **2**

Answer: (a) $P(x) = \begin{cases} \frac{1}{15} & x \in [0, 15] \\ 0 & \text{otherwise} \end{cases}$. (b) $\frac{3}{15}$ (c) $\frac{7}{15}$ (d) $\frac{5}{8}$

1.3.4 Distributions with unbounded domains

Important note

A probability distribution function may have an unbounded domain and will result in an ....improper.......... ....integral.......... with property:

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

Example 16

(a) Find the area under the curve $y = \dfrac{1}{x^2}$ for $(1, a)$, $a > 1$.

(b) Take the limit as $a \to \infty$, and show that $y = \dfrac{1}{x^2}$, $x \geq 1$ is a probability distribution function.

(c) Find the corresponding cumulative distribution function.

Further exercises

(A) Ex 10B

- Q1-14

(A) Ex 16B

- Q1-14

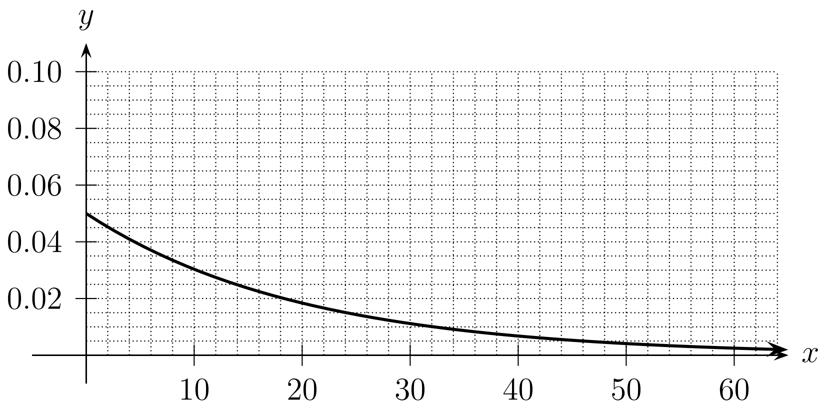
1.3.5 Supplementary exercises

1. [**2020 SACE Mathematical Methods Sample Paper Q10**] At a sporting event attended by thousands of people, individuals spend time in a queue in order to enter the stadium. The probability that a randomly chosen individual spends time, t, in the entry queue can be modelled by the probability density function

$$f(t) = 0.05e^{-0.05t}$$

where t is measured in minutes and $t \geq 0$.

The graph of $y = f(t)$ is shown below.



(a) i. Calculate the probability that a randomly chosen individual spends between 0 and 10 minutes in the entry queue. **2**

ii. On the graph above, draw a representation of your answer to part (a)i. **1**

(b) i. Calculate the probability that a randomly chosen individual spends less than 60 minutes in the entry queue. **1**

ii. Of 200 individuals who entered the stadium, how many does the model predict spent more than 60 minutes in the entry queue? **2**

(c) Describe *one* limitation of using a function of the form $f(t) = a \times e^{-at}$ to model the time that an individual spends in the entry queue. **1**

It is a management policy that once inside the stadium, all individuals who spend time in a queue to buy food must be served within 20 minutes. The probability that an individual spends time t in the food queue can be modelled by the probability density function

$$g(t) = B \times e^{-0.05t}$$

where B is a positive value, t is measured in minutes, and $0 \leq t \leq 20$.

(d) Find the *exact* value of B. **4**

(e) On the set of axes above, sketch the graph of $y = g(t)$. **2**

(f) Is the probability of an individual spending between 0 and 10 minutes in the food queue greater than or less than the probability of an individual spending between 0 and 10 minutes in the entry queue? **2**

Give a reason for your answer, without calculating the probability of spending between 0 and 10 minutes in the food queue

2. [**2016 VCE Mathematical Methods Paper 2 Q18**] The continuous random variable, X, has a probability density function given by

$$f(x) = \begin{cases} \frac{1}{4}\cos\left(\frac{x}{2}\right) & 3\pi \le x \le 5\pi \\ 0 & \text{elsewhere} \end{cases}$$

The value of a such that $P(X < a) = \frac{\sqrt{3}+2}{4}$ is

(A) $\frac{19\pi}{6}$ (B) $\frac{14\pi}{3}$ (C) $\frac{10\pi}{3}$ (D) $\frac{29\pi}{6}$ (E) $\frac{17\pi}{3}$

3. [**2015 VCE Mathematical Methods Paper 2 Q13**] The function f is a probability density function with rule

$$f(x) = \begin{cases} ae^x & 0 \le x \le 1 \\ ae & 1 < x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

The value of a is

(A) 1 (B) e (C) $\frac{1}{e}$ (D) $\frac{1}{2e}$ (E) $\frac{1}{2e-1}$

4. [**2008 VCE Mathematical Methods Paper 2 Q11**] The probability density function for the continuous random variable X is given by

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

The probability that $X < 1.5$ is equal to

(A) 0.125 (B) 0.375 (C) 0.5 (D) 0.625 (E) 0.75

Answers

1. (a) i. $1 - e^{-0.5} \approx 0.3935$ ii. Sketch area. (b) i. $1 - e^{-3} \approx 0.9502$ ii. $200e^{-3} \approx 10$ (c) There could be people who theoretically, are waiting indefinitely. (d) $B = \frac{e}{20(e-1)}$ (e) Sketch (f) See height of graphs. **2.** (B) **3.** (E) **4.** (D)

1.4 Expected value and variance

Important note

Not in the Mathematics Advanced syllabus, but provided for continuity of the ideas from Topic 14 - Discrete Probability Distributions.

- If formulae are given, students may be expected to make such calculations.

	Discrete	**Continuous**
Distribution (notation)	p_i	$f_X(x)$
Probability of event (notation)	$P(X = x)$	$P(c \le X \le d)$
Validity	$\sum p_i = 1$	$\int_{-\infty}^{\infty} f(x)\,dx = 1$
Expected value $E(X)$	$\sum x p_i$	$\int_{-\infty}^{\infty} x f(x)\,dx$
Variance $\text{Var}(X) = E\left(X^2\right) - \left(E(X)\right)^2$	$\sum x^2 p_i - \mu^2$	$\int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$

Example 17

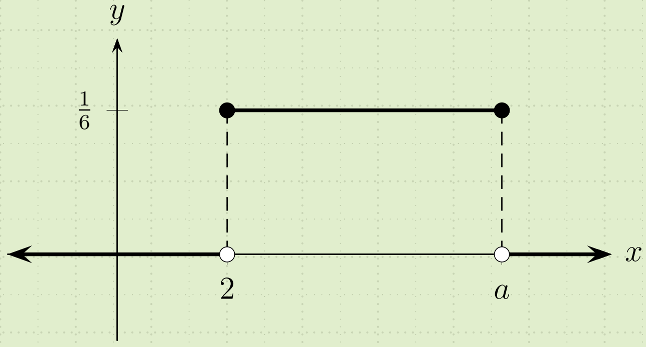
[**2013 VCE Mathematical Methods Paper 1 Q8**] (3 marks) A continuous random variable X, has a probability density function

$$f(x) = \begin{cases} \dfrac{\pi}{4}\cos\left(\dfrac{\pi x}{4}\right) & x \in [0, 2] \\ 0 & \text{otherwise} \end{cases}$$

Given that $\dfrac{d}{dx}\left(x \sin\left(\dfrac{\pi x}{4}\right)\right) = \dfrac{\pi x}{4}\cos\left(\dfrac{\pi x}{4}\right) + \sin\left(\dfrac{\pi x}{4}\right)$, find $E(X)$. **Answer:** $2 - \frac{4}{\pi}$

Example 18

[**2014 VCE Mathematical Methods Paper 2 Q9**] The graph of the probability density function of a continuous random variable, X, is shown below.



If $a > 2$, what is $E(X)$ equal to?

(A) 8 (B) 5 (C) 4 (D) 3 (E) 2

Example 19

[2017 WACE Mathematics Methods Section 2 Q11] A pizza shop estimates that the time X hours to deliver a pizza from when it is ordered is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \dfrac{4}{3} - \dfrac{2}{3}x & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the probability of a pizza being delivered within half an hour of being ordered? **2**

(b) Calculate the mean delivery time to the nearest minute **3**

(c) Calculate the standard deviation of the delivery time to the nearest minute. **4**

Answer: (a) $\frac{7}{12}$ (b) $\frac{4}{9}$ hrs, ≈ 27 min (c) $\sqrt{\frac{13}{162}} \approx 0.2833$, ≈ 17 min

Example 20

[2014 VCE Mathematical Methods Paper 2 Q16] The continuous random variable X, with probability density function $p(x)$, has mean 2 and variance 5. What is the value of

$$\int_{-\infty}^{\infty} x^2 p(x)$$

(A) 1 (B) 7 (C) 9 (D) 21 (E) 29

Further exercises

⚠ Not in the syllabus, but provided as an enrichment exercise.

Ⓐ **Ex 10C**

- Q1-7, Q8-14

(x1) **Ex 16C**

- Q1-7, Q8-14

Section 2

S2 The Normal Distribution

Learning Goal(s)

Knowledge
Normal Distribution (x) and Standard Normal (z) scores

Skills
Calculate, conversion of x to z scores

Understanding
Standard normal distribution make judgements

By the end of this section am I able to:

32.4 Identify the numerical and graphical properties of data that is normally distributed.

32.5 Calculate probabilities and quantiles associated with a given normal distribution using technology and otherwise, and use these to solve practical problems.

32.6 Understand and calculate the z-score (standardised score) corresponding to a particular value in a dataset

32.7 Use z-scores to compare scores from different datasets, for example comparing students' subject examination scores.

32.8 Use collected data to illustrate the empirical rules for normally distributed random variables.

32.9 Use z-scores to identify probabilities of events less or more extreme than a given event.

32.10 Use z-scores to make judgements related to outcomes of a given event or sets of data.

History

Carl Friedrich Gauss and the normal distribution

- The normal distribution was first used by Abraham De Moivre in 1734 and subsequently proven rigourously by Carl Friedrich Gauss in 1809, and expanded by other mathematicians such as Pierre Simon de Laplace around 1812.
- Sometimes known as a *Gaussian* curve, as attributed to C. F. Gauss' proof.

Gauss appears on one of the last German DM10 notes.

Notice the Standard Normal Distribution on the note!

Source: http://www.willamette.edu/~mjaneba/help/normalcurve.html

2.1 Definition

Theorem 1

When the number of samples taken becomes large (generally approaching the size of the population), and many other small independent factors contribute roughly equally, the distribution of the data will resemble a bell curve that is known as the *normal distribution*.

Important note

- The normal distribution is an accurate tool that describes the observations of natural phenomenon, population data and industrial processes.
- When a data set's distribution is not normal, it generally is due to other random phenomena contributing more than others.

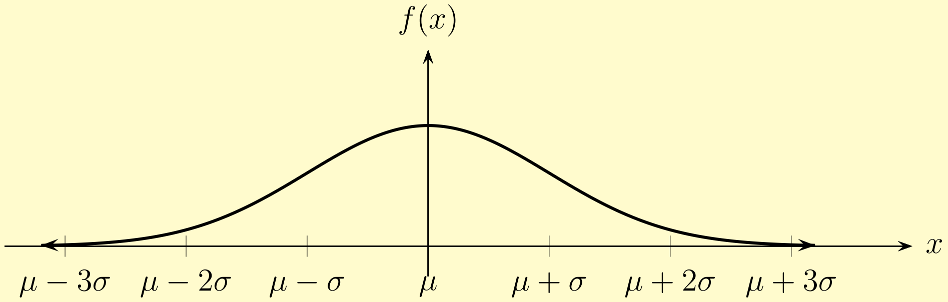
Laws/Results

If a random variable X is *normally distributed*, i.e. $X \sim N(\mu, \sigma^2)$ where μ is the ....mean.... and σ^2 is the ....variance....:

- Its ....probability.... ....density.... ....function.... is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

with domain $D = \{x : x \in \mathbb{R}\}$



- $f_X(x)$ ..$\geq$.. 0
- $\int_{-\infty}^{\infty} f_X(x)\,dx =$ ..1..
- x coordinates of the points of inflexion: $x = \mu \pm \sigma$

2.1.1 Standard normal distribution

Definition 6

The *standard normal distribution* has $\mu =$ ....0.... and $\sigma =$ ..1.. . If Z is a random variable represented by the standard normal distribution,

$$Z \sim \ldots N(0, 1^2) \ldots$$

Draw the distribution, marking out values up to 3 sd to the left and right of the mean:

Important note

The letter used for a random variable which follows the standard normal distribution, is usually Z.

2.2 CDF of a normally distributed random variable

Important note

(See also Definition 4 on page 13) ⚠ The *cumulative distribution function*, $F(x)$ for a normally distributed random variable cannot found easily via current integration techniques!

- Three techniques are required to calculate the area beneath a normal distribution curve, shown in increasing order of rigour.
- **Use symmetry** to find the required areas.
- Use the most appropriate technique or as stated in the question.

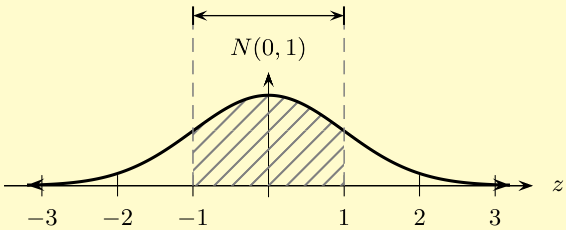
2.2.1 68-95-99.7 rule

- The standard normal distribution is used first.
- For other data sets, the same proportions apply.
- **Use symmetry** to find the required areas.

Laws/Results

Within 1 standard deviation of the mean

$$P(-1 < z < 1) \approx \underline{0.68}, \quad \underline{(2\text{ d.p.})}$$



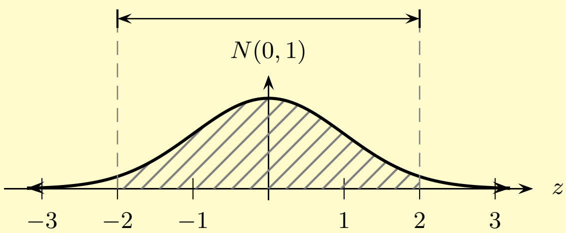
- 68% of data lies within 1 standard deviation from the mean.

$$P(\mu - \sigma < X < \mu + \sigma) = \int_{\mu-\sigma}^{\mu+\sigma} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\, dx \approx \underline{0.68}$$

Laws/Results

Within 2 standard deviations of the mean

$$P(-2 < z < 2) \approx \underline{0.95}, \quad \underline{(2\text{ d.p.})}$$



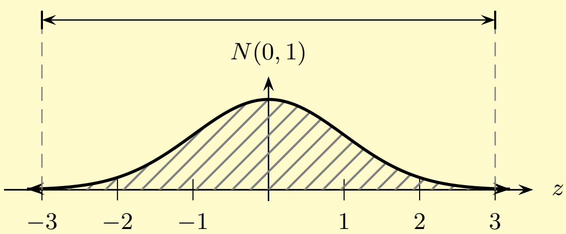
- 95% of data lies within 2 standard deviation from the mean.

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = \int_{\mu-2\sigma}^{\mu+2\sigma} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\, dx \approx \underline{0.95}$$

Laws/Results

Within 3 standard deviations of the mean

$$P(-3 < z < 3) \approx \underline{0.997}, \quad \underline{(3\text{ d.p.})}$$



- 99.7% of data lies within 3 standard deviation from the mean.

$$P(\mu - 3\sigma < X < \mu + 3\sigma) = \int_{\mu-3\sigma}^{\mu+3\sigma} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\, dx \approx \underline{0.997}$$

Example 21

(Haese et al., 2016, p.234) The chest measurements of 18 year old male footballers are normally distributed with mean 95 cm and standard deviation 8 cm.

(a) Find the percentage of footballers with chest measurements between 95 cm and 103 cm.

(b) From a group of two hundred 18 year old male footballers, how many would you expect to have a chest measurement between 87 cm and 111 cm?

(c) Find the value of k such that approximately 16% of chest measurements are below k cm.

Answer: (a) 34 (b) 164 (c) 87

Example 22

[**2006 VCE Mathematical Methods Paper 1 Q5**] Let X be a normally distributed random variable with a mean of 72 and a standard deviation of 8. Let Z be the standard normal random variable. Find:

(a) the probability that X is greater than 80 **1**

(b) the probability that $64 < X < 72$ **1**

(c) the probability that $X < 64$ given that $X < 72$ **2**

Answer: (a) 0.16 (b) 0.34 (c) $\frac{8}{25}$

Example 23

[**2017 WACE Mathematics Methods Section 2 Q19**] (2 marks) A global financial institution transfers a large aggregate data file every evening from offices around the world to its Hong Kong head office. Once the file is received it must be processed in the company's data warehouse. The time T required to process a file is normally distributed with a mean of 90 minutes and a standard deviation of 15 minutes.

An evening is selected at random. What is the probability that it takes more than two hours to process the file? **Answer:** 0.025

Laws/Results

Quartiles and percentiles

Lower quartile (25 th percentile) or Q_1 occurs at $x = a$ such that

$$\int_{-\infty}^{a} f(x)\,dx = 0.25$$

Median Q_2 equals the mean.

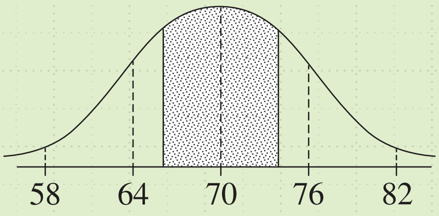
Upper quartile (75 th percentile) or Q_3 occurs at $x = a$ such that

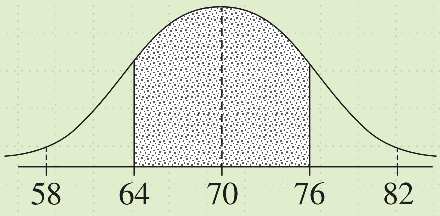
$$\int_{-\infty}^{a} f(x)\,dx = 0.75$$

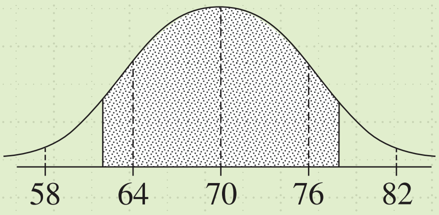
Example 24

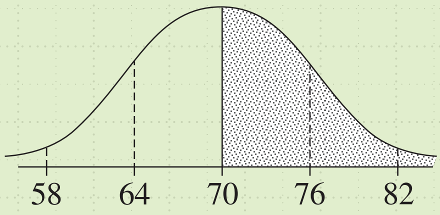
[2020 Mathematics Advanced Sample HSC Q9] The scores on an examination are normally distributed with a mean of 70 and a standard deviation of 6. Michael received a score on the examination between the lower quartile and the upper quartile of the scores.

Which shaded region most accurately represents where Michael's score lies?

A.


B.


C.


D.


2.2.2 Transforming from a normal to standard normal distribution

Laws/Results

Every random variable $X \sim N(\mu, \sigma^2)$ can be transformed into a random variable $Z \sim N(0,1)$ with the following transformation:

$$z = \frac{x-\mu}{\sigma}$$

Example 25

[2018 VCE Mathematical Methods NHT Paper 2 Q9] A continuous random variable X has a normal distribution with a mean of 40 and a standard deviation of 5. The continuous random variable Z has the standard normal distribution.
$P(-2 < Z < 1)$ is equal to:

(A) $P(40 < X < 55)$
(B) $P(35 < X < 50)$
(C) $P(30 < X < 50)$
(D) $P(10 < X < 30)$
(E) $P(X > 30) - P(X < 45)$

Example 26

[2009 VCE Mathematical Methods Paper 2 Q6] The continuous random variable X has a normal distribution with mean 14 and standard deviation 2. If the random variable Z has the standard normal distribution, then the probability that X is greater than 17 is equal to

(A) $P(Z > 3)$
(B) $P(Z < 2)$
(C) $P(Z < 1.5)$
(D) $P(Z < -1.5)$
(E) $P(Z > 2)$

Example 27

[2010 VCE Mathematical Methods Paper 1 Q5] Let X be a normally distributed random variable with mean 5 and variance 9 and let Z be the random variable with the standard normal distribution.

(a) Find $P(X > 5)$ **1**

(b) Find b such that $P(X > 7) = P(Z < b)$ **2**

Answer: (a) 0.5 (b) $b = -\frac{2}{3}$

Example 28

[2019 CSSA Standard 2 Trial Q23] The weight of packets of Felicity's Frozen Fish Fillets is found to be normally distributed with a mean of 512 grams and a standard deviation of 6 grams.

(a) Phuong has a packet of Felicity's Frozen Fish Fillets with a standardised weight, $z = 1.75$. **1**

What is the actual weight of the packet?

(b) A wholesaler sells boxes of 40 packets of Felicity's Frozen Fish Fillets. **2**

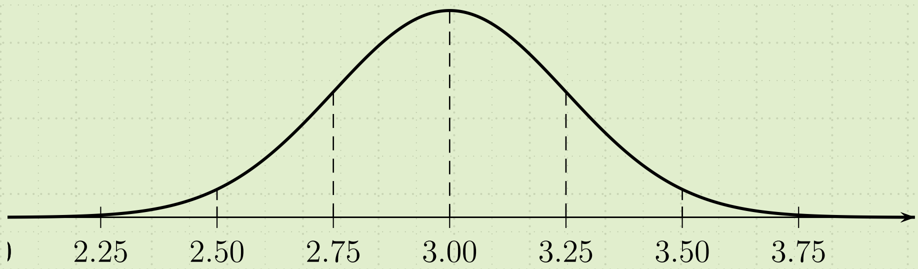
Given that each packet is labelled as containing 500 grams, how many packets in each box would you expect to be under the labelled weight? Support your answer with appropriate calculations.

Answer: (a) 522.5 (b) 1 pkt

Example 29

[**2019 Independent Standard 2 Trial Q16**] A Biologist found that the length of the wing span of a sample of butterflies captured from a specific location, was normally distributed.

The bell curve below was drawn by the Biologist after the wing spans of the butterflies were measured.



(b) What would have been the length (cm) of the wingspan of butterflies recorded with a z score of -1.3? **2**

(c) The Biologist wants to do more research on the butterflies whose wing spans were greater than 3.50 cm and less than 2.50 cm. **1**

What percentage of the sample would the Biologist use in her research?

(d) If there were 250 butterflies in the sample, explain why it would be unlikely for the Biologist to have found butterflies with a wingspan greater than 3.75 cm. **2**

Answer: (a) Not in course (b) 2.675 (c) 5% (d) 0.375 butterflies

Example 30

[2015 Mathematics General 2 HSC Q28] (2 marks) The results of two tests are normally distributed. The mean and standard deviation for each test are displayed in the table.

	Mathematics	English
$\overline{x}$	70	75
s	6.5	8

Kristoff scored 74 in Mathematics and 80 in English. He claims that he has performed better in English.

Is Kristoff correct? Justify your answer using appropriate calculations.

Example 31

[2012 General Mathematics HSC Q29] (2 marks) A machine produces nails. When the machine is set correctly, the lengths of the nails are normally distributed with a mean of 6.000 cm and a standard deviation of 0.040 cm.

To confirm the setting of the machine, three nails are randomly selected. In one sample the lengths are 5.950, 5.983 and 6.140.

The setting of the machine needs to be checked when the lengths of two or more nails in a sample lie more than 1 standard deviation from the mean.

Does the setting on the machine need to be checked? Justify your answer with suitable calculations.

2.2.3 Trapezoidal rule

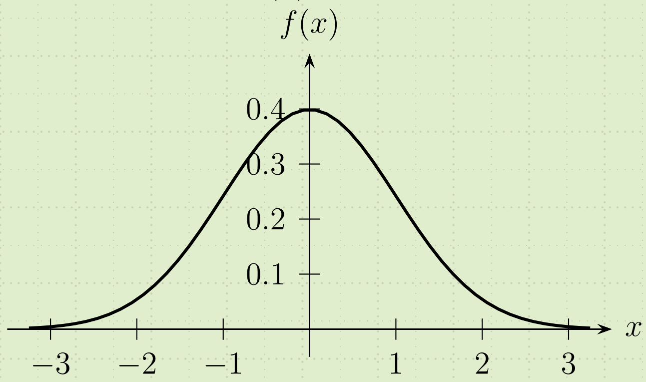
Example 32

[**2020 Mathematics Advanced Sample HSC Q29**] Let X denote a normal random variable with mean 0 and standard deviation 1. The random variable X has the probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

where $-\infty < x < \infty$.

The diagram shows the graph of $y = f(x)$.



(a) Complete the table of values for the function given. Give your answer correct to four significant figures. **1**

x	0	1	2	3
$f(x)$	0.3989	0.2420		0.004432

(b) Using the trapezoidal rule and the 4 function values in the table in part (a), show that **2**

$$P(-3 \leq X \leq 3) = \int_{-3}^{3} f(x)\, dx \approx 0.9953$$

(c) The IQ (Intelligence Quotient) scores for a large population are normally distributed with a mean of 100 and a standard deviation of 15. Using the result obtained in part (b), calculate the probability of randomly selecting a person with an IQ score above 145 from this large population. **2**

2.2.4 Statistical tables

Important note

Do NOT memorise ANY values from this!

- This will be supplied on the examination, if it is required. Otherwise, default to the 68-96-99.7 rule (Section 2.2.1 on page 30) or the Trapezoidal rule (Section 2.2.3 on the preceding page).
- Values are machine calculated.

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Table 2.1 – Standard normal distribution values - $Z \sim N(0,1)$, $P(Z < a) = \displaystyle\int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}\, dz$

Important note

Questions requiring statistical tables usually involve

- Standard deviations at non-integer values, and
- Calculation.

Example 33

[**2017 VCE Mathematical Methods NHT Paper 2 Q12**] The maximum temperature reached by the water heated in a kettle each time it is used is normally distributed with a mean of 95°C and a standard deviation of 2°C.

When the kettle is used, the proportion of times that the maximum temperature reached by the water is greater than 98°C is closest to

(A) 0.0671 (B) 0.0668 (C) 0.8669 (D) 0.9332 (E) 0.9342

Example 34

[2019 VCE Mathematical Methods NHT Paper 2 Q3] Concerts at the Mathsland Concert Hall begin L minutes after the scheduled starting time. L is a random variable that is normally distributed with a mean of 10 minutes and a standard deviation of four minutes.

(a) What proportion of concerts begin before the scheduled starting time, correct to four decimal places? **1**

(b) Find the probability that a concert begins more than 15 minutes after the scheduled starting time, correct to four decimal places. **1**

If a concert begins more than 15 minutes after the scheduled starting time, the cleaner is given an extra payment of \$200. If a concert begins up to 15 minutes after the scheduled starting time, the cleaner is given an extra payment of \$100. If a concert begins at or before the scheduled starting time, there is no extra payment for the cleaner.

Let C be the random variable that represents the extra payment for the cleaner, in dollars.

(c) i. Using your responses from part (a) and (b), complete the following table, correct to 3 decimal places. **1**

c	0	100	200
$P(C = c)$			

ii. Calculate the expected value of the extra payment for the cleaner, to the nearest dollar. **1**

iii. Calculate the standard deviation of C, correct to the nearest dollar. **1**

Answer: (a) 0.0062 (b) 0.1056 (c) i. 0.006, 0.888, 0.106 ii. \$110 iii. \$32

Further exercises

(A) **Ex 10D, 10E, 10F**

- As required

(X1) **Ex 16D, 16E, 16F**

- As required

2.2.5 Supplementary exercises

⚠ Questions in this subsection will almost certainly require Table 2.1 on page 40. It is unlikely to reflect the exact nature of HSC questions, but is provided here nonetheless as an additional source of practice.

1. [**2006 VCE Mathematical Methods Paper 2 Q21**] The times (in minutes) taken for students to complete a university test are normally distributed with a mean of 200 minutes and standard deviation 10 minutes.

The proportion of students who complete the test in less than 208 minutes is closest to

(A) 0.200 (B) 0.212 (C) 0.758 (D) 0.788 (E) 0.800

2. [**2007 VCE Mathematical Methods Paper 2 Q18**] The heights of the children in a queue for an amusement park ride are normally distributed with mean 130 cm and standard deviation 2.7 cm. 35% of the children are not allowed to go on the ride because they are too short.

The minimum acceptable height correct to the nearest centimetre is

(A) 126 (B) 127 (C) 128 (D) 129 (E) 130

3. [**2009 VCE Mathematical Methods Paper 2 Q3**] The Bouncy Ball Company (BBC) makes tennis balls whose diameters are normally distributed with mean 67 mm and standard deviation 1 mm. The tennis balls are packed and sold in cylindrical tins that each hold four balls. A tennis ball fits into such a tin if the diameter of the ball is less than 68.5 mm.

(a) What is the probability, correct to four decimal places, that a randomly selected tennis ball produced by BBC fits into a tin? **2**

(b) BBC management would like each ball produced to have diameter between 65.6 and 68.4 mm. **2**

What is the probability, correct to four decimal places, that the diameter of a randomly selected tennis ball made by BBC is in this range?

(c) What is the probability, correct to four decimal places, that the diameter of a tennis ball which fits into a tin is between 65.6 and 68.4 mm? **1**

(d) BBC management wants engineers to change the manufacturing process so that 99% of all balls produced have diameter between 65.6 and 68.4 mm. The mean is to stay at 67 mm but the standard deviation is to be changed. **3**

What should the new standard deviation be (correct to two decimal places)?

4. **[2009 VCE Mathematical Methods Paper 2 Q13]** The continuous random variable X has a normal distribution with mean 20 and standard deviation 6. The continuous random variable Z has the standard normal distribution.

The probability that Z is between -2 and 1 is equal to

(A) $P(18 < X < 21)$

(B) $P(14 < X < 32)$

(C) $P(14 < X < 26)$

(D) $P(8 < X < 32)$

(E) $P(X > 14) + P(X < 26)$

5. **[2010 VCE Mathematical Methods Paper 1 Q8]** The random variable X is normally distributed with mean 100 and standard deviation 4. **2**
If $P(X < 106) = q$, find $P(94 < X < 100)$ in terms of q.

6. **[2012 VCE Mathematical Methods Paper 2 Q11]** The weights of bags of flour are normally distributed with mean 252 g and standard deviation 12 g. The manufacturer says that 40% of bags weigh more than x g.

The maximum possible value of x is closest to

(A) 249.0 (B) 251.5 (C) 253.5 (D) 254.5 (E) 255.0

7. **[2014 VCE Mathematical Methods Paper 2 Q4]** Patricia is a gardener and she owns a garden nursery. She grows and sells basil plants and coriander plants.
The heights, in centimetres, of the basil plants that Patricia is selling are distributed normally with a mean of 14 cm and a standard deviation of 4 cm. There are 2000 basil plants in the nursery.

(a) Patricia classifies the tallest 10 per cent of her basil plants as **super**. **1**

What is the minimum height of a super basil plant, correct to the nearest millimetre?

(b) Patricia decides that some of her basil plants are not growing quickly enough, so she plans to move them to a special greenhouse. She will move the basil plants that are less than 9 cm in height. **2**

How many basil plants will Patricia move to the greenhouse, correct to the nearest whole number?

(c) (X2) The heights of the coriander plants, x centimetres, follow the probability density function $h(x)$, where **(3?)**

$$h(x) = \begin{cases} \dfrac{\pi}{100}\sin\left(\dfrac{\pi x}{50}\right) & 0 < x < 50 \\ 0 & \text{otherwise} \end{cases}$$

State the mean height of the coriander plants (Note: See Section 1.4 on page 24)

8. **[2014 VCE Mathematical Methods Paper 1 Q6]** Let the random variable X be normally distributed with mean 2.5 and standard deviation 0.3 Let Z be the standard normal random variable, such that $Z \sim N(0, 1)$.

(a) Find b such that $P(X > 3.1) = P(Z < b)$. **1**

(b) Using the fact that, correct to two decimal places, $P(Z < -1) = 0.16$, find $P(X < 2.8|X > 2.5)$. Write the answer correct to two decimal places. **2**

9. **[2016 VCE Mathematical Methods Paper 2 Q3]** **2**

(c) The time for which a laptop will work without recharging (the battery life) is normally distributed, with a mean of three hours and 10 minutes and standard deviation of six minutes. Suppose that the laptops remain out of the recharging trolley for three hours. **2**

For any one laptop, find the probability that it will stop working by the end of these three hours. Give your answer correct to four decimal places.

(e) It is known that when laptops have been used regularly in a school for six months, their battery life is still normally distributed but the mean battery life drops to three hours. It is also known that only 12% of such laptops work for more than three hours and 10 minutes. **2**

Find the standard deviation for the normal distribution that applies to the battery life of laptops that have been used regularly in a school for six months, correct to four decimal places.

10. **[2018 VCE Mathematical Methods NHT Paper 2 Q2]**

(c) A hole is drilled into each motor. The depth of the hole is normally distributed with a mean of 20 mm and a standard deviation of 0.3 mm. **1**

What is the probability that, for a randomly selected motor, the depth of the hole is greater than 20.6 mm? Give your answer correct to four decimal places.

(d) The depth of the hole drilled into a motor must be within 0.5 mm of the mean, otherwise the motor is defective. **2**

What is the probability that a motor is defective, correct to four decimal places?

(f) A knob is attached to each controller. The height of a knob is normally distributed with a mean of 30 mm. If the knob on a controller has a height greater than 30.4 mm or less than 29.6 mm, then the controller is defective. **2**

Rebecca wants to ensure that less than 2% of all controllers manufactured are defective.

What is the maximum standard deviation of the height of a knob, in millimetres, that can be attached to a controller so that less than 2% of controllers are defective? Give your answer correct to two decimal places.

11. [**2019 VCE Mathematical Methods Paper 2 Q14**] The weights of packets of lollies are normally distributed with a mean of 200 g. If 97% of these packets of lollies have a weight of more than 190 g, then the standard deviation of the distribution, correct to one decimal place, is

(A) 3.3 g (B) 5.3 g (C) 6.1 g (D) 9.4 g (E) 12.1 g

12. [**2019 VCE Mathematical Methods Paper 2 Q4**] The Lorenz birdwing is the largest butterfly in Town A. The probability density function that describes its life span, X, in weeks, is given by

$$f(x) = \begin{cases} \dfrac{4}{625}\left(5x^3 - x^4\right) & 0 \le x \le 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the mean life span of the Lorenz birdwing butterfly. **2**

(b) In a sample of 80 Lorenz birdwing butterflies, how many butterflies are expected to live longer than two weeks, correct to the nearest integer? **2**

(c) What is the probability that a Lorenz birdwing butterfly lives for at least four weeks, given that it lives for at least two weeks, correct to four decimal places? **2**

The wingspans of Lorenz birdwing butterflies in Town A are normally distributed with a mean of 14.1 cm and a standard deviation of 2.1 cm.

(d) Find the probability that a randomly selected Lorenz birdwing butterfly in Town A has a wingspan between 16 cm and 18 cm, correct to four decimal places. **1**

(e) A Lorenz birdwing butterfly is considered to be **very small** if its wingspan is in the smallest 5% of all the Lorenz birdwing butterflies in Town A. **1**

Find the greatest possible wingspan, in centimetres, for a very small Lorenz birdwing butterfly in Town A, correct to one decimal place.

NESA Reference Sheet – calculus based courses

NSW Education Standards Authority

2020 HIGHER SCHOOL CERTIFICATE EXAMINATION

Mathematics Advanced
Mathematics Extension 1
Mathematics Extension 2

REFERENCE SHEET

Measurement

Length

$$l = \frac{\theta}{360} \times 2\pi r$$

Area

$$A = \frac{\theta}{360} \times \pi r^2$$

$$A = \frac{h}{2}(a + b)$$

Surface area

$$A = 2\pi r^2 + 2\pi rh$$

$$A = 4\pi r^2$$

Volume

$$V = \frac{1}{3}Ah$$

$$V = \frac{4}{3}\pi r^3$$

Functions

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

For $ax^3 + bx^2 + cx + d = 0$:

$$\alpha + \beta + \gamma = -\frac{b}{a}$$

$$\alpha\beta + \alpha\gamma + \beta\gamma = \frac{c}{a}$$

and $\alpha\beta\gamma = -\frac{d}{a}$

Relations

$$(x - h)^2 + (y - k)^2 = r^2$$

Financial Mathematics

$$A = P(1 + r)^n$$

Sequences and series

$$T_n = a + (n - 1)d$$

$$S_n = \frac{n}{2}[2a + (n - 1)d] = \frac{n}{2}(a + l)$$

$$T_n = ar^{n-1}$$

$$S_n = \frac{a(1 - r^n)}{1 - r} = \frac{a(r^n - 1)}{r - 1}, \ r \neq 1$$

$$S = \frac{a}{1 - r}, \ |r| < 1$$

Logarithmic and Exponential Functions

$$\log_a a^x = x = a^{\log_a x}$$

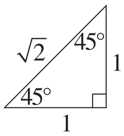
$$\log_a x = \frac{\log_b x}{\log_b a}$$

$$a^x = e^{x \ln a}$$

Trigonometric Functions

$\sin A = \dfrac{\text{opp}}{\text{hyp}}, \quad \cos A = \dfrac{\text{adj}}{\text{hyp}}, \quad \tan A = \dfrac{\text{opp}}{\text{adj}}$

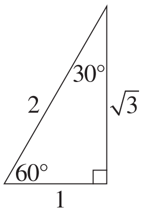
$A = \dfrac{1}{2}ab\sin C$



$\dfrac{a}{\sin A} = \dfrac{b}{\sin B} = \dfrac{c}{\sin C}$

$c^2 = a^2 + b^2 - 2ab\cos C$

$\cos C = \dfrac{a^2 + b^2 - c^2}{2ab}$



$l = r\theta$

$A = \dfrac{1}{2}r^2\theta$

Trigonometric identities

$\sec A = \dfrac{1}{\cos A}, \ \cos A \neq 0$

$\operatorname{cosec} A = \dfrac{1}{\sin A}, \ \sin A \neq 0$

$\cot A = \dfrac{\cos A}{\sin A}, \ \sin A \neq 0$

$\cos^2 x + \sin^2 x = 1$

Compound angles

$\sin(A + B) = \sin A\cos B + \cos A\sin B$

$\cos(A + B) = \cos A\cos B - \sin A\sin B$

$\tan(A + B) = \dfrac{\tan A + \tan B}{1 - \tan A\tan B}$

If $t = \tan\dfrac{A}{2}$ then $\sin A = \dfrac{2t}{1 + t^2}$

$\cos A = \dfrac{1 - t^2}{1 + t^2}$

$\tan A = \dfrac{2t}{1 - t^2}$

$\cos A\cos B = \dfrac{1}{2}\left[\cos(A - B) + \cos(A + B)\right]$

$\sin A\sin B = \dfrac{1}{2}\left[\cos(A - B) - \cos(A + B)\right]$

$\sin A\cos B = \dfrac{1}{2}\left[\sin(A + B) + \sin(A - B)\right]$

$\cos A\sin B = \dfrac{1}{2}\left[\sin(A + B) - \sin(A - B)\right]$

$\sin^2 nx = \dfrac{1}{2}(1 - \cos 2nx)$

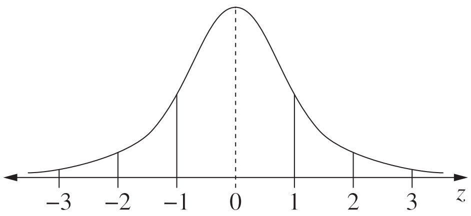
$\cos^2 nx = \dfrac{1}{2}(1 + \cos 2nx)$

Statistical Analysis

$z = \dfrac{x - \mu}{\sigma}$

An outlier is a score
less than $Q_1 - 1.5 \times IQR$
or
more than $Q_3 + 1.5 \times IQR$

Normal distribution



- approximately 68% of scores have z-scores between −1 and 1
- approximately 95% of scores have z-scores between −2 and 2
- approximately 99.7% of scores have z-scores between −3 and 3

$E(X) = \mu$

$\operatorname{Var}(X) = E\left[(X - \mu)^2\right] = E\left(X^2\right) - \mu^2$

Probability

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

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A|B) = \dfrac{P(A \cap B)}{P(B)}, \ P(B) \neq 0$

Continuous random variables

$P(X \le x) = \displaystyle\int_a^x f(x)\,dx$

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

Binomial distribution

$P(X = r) = {}^nC_r p^r (1 - p)^{n - r}$

$X \sim \operatorname{Bin}(n, p)$

$\Rightarrow \ P(X = x)$

$= \dbinom{n}{x} p^x (1 - p)^{n - x}, \ x = 0, 1, \ldots, n$

$E(X) = np$

$\operatorname{Var}(X) = np(1 - p)$

Differential Calculus

Function	Derivative
$y = f(x)^n$	$\frac{dy}{dx} = nf'(x)\left[f(x)\right]^{n-1}$
$y = uv$	$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$
$y = g(u)$ where $u = f(x)$	$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$
$y = \frac{u}{v}$	$\frac{dy}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$
$y = \sin f(x)$	$\frac{dy}{dx} = f'(x)\cos f(x)$
$y = \cos f(x)$	$\frac{dy}{dx} = -f'(x)\sin f(x)$
$y = \tan f(x)$	$\frac{dy}{dx} = f'(x)\sec^2 f(x)$
$y = e^{f(x)}$	$\frac{dy}{dx} = f'(x)e^{f(x)}$
$y = \ln f(x)$	$\frac{dy}{dx} = \frac{f'(x)}{f(x)}$
$y = a^{f(x)}$	$\frac{dy}{dx} = (\ln a)f'(x)a^{f(x)}$
$y = \log_a f(x)$	$\frac{dy}{dx} = \frac{f'(x)}{(\ln a)f(x)}$
$y = \sin^{-1} f(x)$	$\frac{dy}{dx} = \frac{f'(x)}{\sqrt{1-\left[f(x)\right]^2}}$
$y = \cos^{-1} f(x)$	$\frac{dy}{dx} = -\frac{f'(x)}{\sqrt{1-\left[f(x)\right]^2}}$
$y = \tan^{-1} f(x)$	$\frac{dy}{dx} = \frac{f'(x)}{1+\left[f(x)\right]^2}$

Integral Calculus

$$\int f'(x)\left[f(x)\right]^n dx = \frac{1}{n+1}\left[f(x)\right]^{n+1} + c$$

where $n \neq -1$

$$\int f'(x)\sin f(x)\,dx = -\cos f(x) + c$$

$$\int f'(x)\cos f(x)\,dx = \sin f(x) + c$$

$$\int f'(x)\sec^2 f(x)\,dx = \tan f(x) + c$$

$$\int f'(x)e^{f(x)}dx = e^{f(x)} + c$$

$$\int \frac{f'(x)}{f(x)}dx = \ln\left|f(x)\right| + c$$

$$\int f'(x)a^{f(x)}dx = \frac{a^{f(x)}}{\ln a} + c$$

$$\int \frac{f'(x)}{\sqrt{a^2 - \left[f(x)\right]^2}}dx = \sin^{-1}\frac{f(x)}{a} + c$$

$$\int \frac{f'(x)}{a^2 + \left[f(x)\right]^2}dx = \frac{1}{a}\tan^{-1}\frac{f(x)}{a} + c$$

$$\int u\frac{dv}{dx}dx = uv - \int v\frac{du}{dx}dx$$

$$\int_a^b f(x)\,dx$$

$$\approx \frac{b-a}{2n}\left\{f(a) + f(b) + 2\left[f(x_1) + \cdots + f(x_{n-1})\right]\right\}$$

where $a = x_0$ and $b = x_n$

Combinatorics

$^{n}P_r = \frac{n!}{(n-r)!}$

$\binom{n}{r} = {}^{n}C_r = \frac{n!}{r!(n-r)!}$

$(x+a)^n = x^n + \binom{n}{1}x^{n-1}a + \cdots + \binom{n}{r}x^{n-r}a^r + \cdots + a^n$

Vectors

$\left|\underset{\sim}{u}\right| = \left|x\underset{\sim}{i} + y\underset{\sim}{j}\right| = \sqrt{x^2 + y^2}$

$\underset{\sim}{u} \cdot \underset{\sim}{v} = \left|\underset{\sim}{u}\right|\left|\underset{\sim}{v}\right|\cos\theta = x_1x_2 + y_1y_2$,

where $\underset{\sim}{u} = x_1\underset{\sim}{i} + y_1\underset{\sim}{j}$

and $\underset{\sim}{v} = x_2\underset{\sim}{i} + y_2\underset{\sim}{j}$

$\underset{\sim}{r} = \underset{\sim}{a} + \lambda\underset{\sim}{b}$

Complex Numbers

$$\begin{aligned} z = a + ib &= r(\cos\theta + i\sin\theta) \\ &= re^{i\theta} \end{aligned}$$

$$\begin{aligned} \left[r(\cos\theta + i\sin\theta)\right]^n &= r^n(\cos n\theta + i\sin n\theta) \\ &= r^n e^{in\theta} \end{aligned}$$

Mechanics

$\frac{d^2x}{dt^2} = \frac{dv}{dt} = v\frac{dv}{dx} = \frac{d}{dx}\left(\frac{1}{2}v^2\right)$

$x = a\cos(nt + \alpha) + c$

$x = a\sin(nt + \alpha) + c$

$\ddot{x} = -n^2(x - c)$

References

Haese, M., Haese, S., & Humphries, M. (2016). *Mathematics for Australia 12 Mathematical Methods* (2nd ed.). Haese Mathematics.

Pender, W., Sadler, D., Ward, D., Dorofaeff, B., & Shea, J. (2019a). *CambridgeMATHS Stage 6 Mathematics Advanced Year 12* (1st ed.). Cambridge Education.

Pender, W., Sadler, D., Ward, D., Dorofaeff, B., & Shea, J. (2019b). *CambridgeMATHS Stage 6 Mathematics Extension 1 Year 11* (1st ed.). Cambridge Education.

Pender, W., Sadler, D., Ward, D., Dorofaeff, B., & Shea, J. (2019c). *CambridgeMATHS Stage 6 Mathematics Extension 1 Year 12* (1st ed.). Cambridge Education.