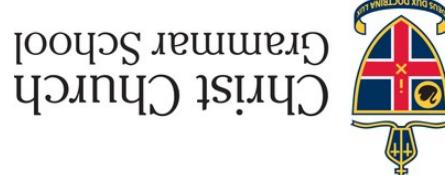


MATHEMATICS METHODS Year 12

Part One:
Take home investigation



Materials required/recommended for this section

Time allowed for this section: 2 weeks

Working time for this section:

This Question/Answer Booklet

To be provided by the supervisor

Standard items: pens (blue/black preferred), pencils (including coloured), sharpener, correction fluid/tape, eraser, ruler, highlighters

Special items: drawing instruments, templates, and up to three calculators approved for use in the WACE examinations

This investigation has 2 parts: a take home investigation that the student needs to research and complete at home, and an in-class validation that will be completed in test conditions during class.

The take home investigation will not be assessed but has to be completed prior to the validation. Solution for the take home investigation will be provided.

This paper is the take home investigation. It is handed out on Tuesday 27 June 2017.

The in-class validation will be assessed and it constitutes 7% of the course mark.

The in-class validation will be on Thursday 20 July 2017.

Instructions to candidates

1. Write your answers in this Question/Answer Booklet.
2. Answer all questions.
3. **Show all your working clearly.** Your working should be in sufficient detail to allow your answers to be checked readily and for marks to be awarded for reasoning. Incorrect answers given without supporting reasoning cannot be allocated any marks. For any question or part question worth more than two marks, valid working or justification is required to receive full marks. If you repeat an answer to any question, ensure that you cancel the answer you do not wish to have marked.
4. It is recommended that **you do not use pencil**, except in diagrams.

1. To achieve the objectives of this investigation and to prepare for the validation, students will have to work through all the notes and questions specified as Required Material.
2. Additional resources can be found in:
- O.T. Lee, WACE Revision Series, Mathematics Methods Year 12, Pages 203 – 210.
 - Creamlin Exam Questions, Mathematics Methods Units 3 & 4, Pages 117 – 123.

InSTRUCTIONS TO CANDIDATES

1. All the material contained in this booklet.
2. All the material found in:
- A.J. Sadler, Mathematics Methods Unit 4, Chapter 4 The Normal Distribution

REQUIRED MATERIAL

- At the end of this investigation, you should be able to:
- identify contexts, such as naturally occurring variation, that are suitable for modelling by normal random variables
 - identify features of the graph of the probability density function of the normal distribution
 - calculate probabilities and quantities associated with a given normal distribution using technology, and use these to solve practical problems

LEARNING OBJECTIVES

- The aim of this investigation is to examine events whose probabilities can be modelled by a Normal Distribution (bell shape), learn how to calculate the probabilities and quantities in a Normal Distribution and use them to solve practical problems.

AIM OF INVESTIGATION

THE NORMAL DISTRIBUTION

Investigation

In an agricultural sciences study, five hundred 10-acre plots of land were prepared and tended in identical fashion for a wheat crop. At harvest, the size of the crop yield of each plot was recorded in bushels.

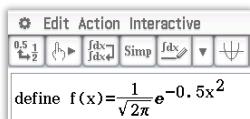
The crop yield of the plots was found to be normally distributed with a mean of 3.9 bushels and a standard deviation of 0.45 bushels.

Question 1

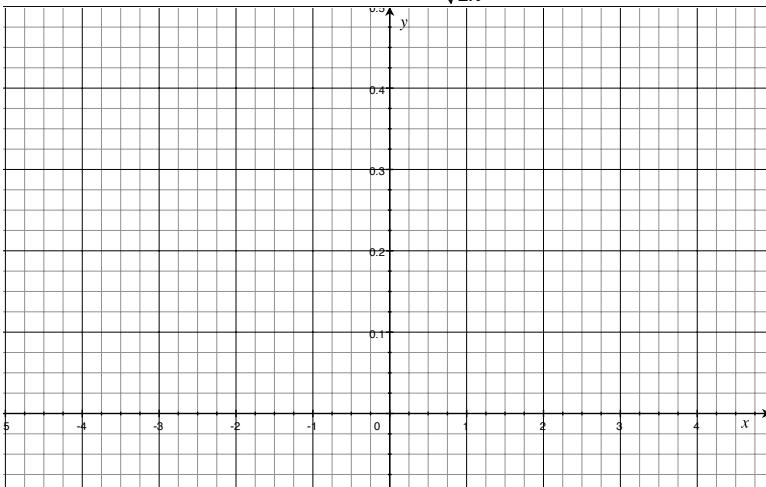
Consider the function $f(x) = \frac{1}{\sqrt{2\pi}} e^{-0.5x^2}$

This is equation of the graph of the standard normal distribution with mean = 0 and standard deviation = 1

This function may be conveniently entered into a CAS calculator for convenient recall and calculation. For a Casio ClassPad the syntax is:



- (a) Use your calculator to assist you to graph $f(x) = \frac{1}{\sqrt{2\pi}} e^{-0.5x^2}$ on the axes below



See next page

Question 8

Solution

(a) (i) $z \approx 0.375$

(ii)

$$X - \text{score} = -0.375(-0.45) + 3.9$$

$$= 3.73125$$

∴ Approx 3.73 bushels

(b) (i)

$$z = 1.25$$

$$\Rightarrow X - \text{score} = 1.25(0.45 + 3.9)$$

$$= 4.4625$$

∴ Approx 4.46 bushels

(ii)

$$\text{Min: } 3.73 \times 8.5 = 31.72$$

$$\text{Max: } 4.46 \times 8.5 = 37.93$$

∴ Expected return is from \$31.70 to \$37.90

Question 9

Solution

(a)

$$P(x < k) = 0.35 \Rightarrow k = 3.7266$$

∴ Minimum expected harvest is 3.73 bushels

(b)

$$P(x > k) = 0.9 \Rightarrow k = 4.4767$$

∴ Maximum expected harvest is 4.4767 bushels

(c) (i)

$$\text{Min: } 3.7266 \times 8.5 = 31.68$$

$$\text{Max: } 4.4767 \times 8.5 = 38.05$$

∴ Expected return is from \$31.70 to \$38.10

(ii)

$$\text{Min: } 500 \times 31.68 = 19025.97$$

$$\text{Max: } 500 \times 38.05 = 15838.09$$

∴ Expected return is from \$15 800 to \$19 000

Using the ClassPad, this can be done efficiently as $f(x)$ has already been defined

(e) Verify your answer in part (d) by evaluating $\int_{-\infty}^{\infty} e^{-0.5x^2} dx$ on your calculator.

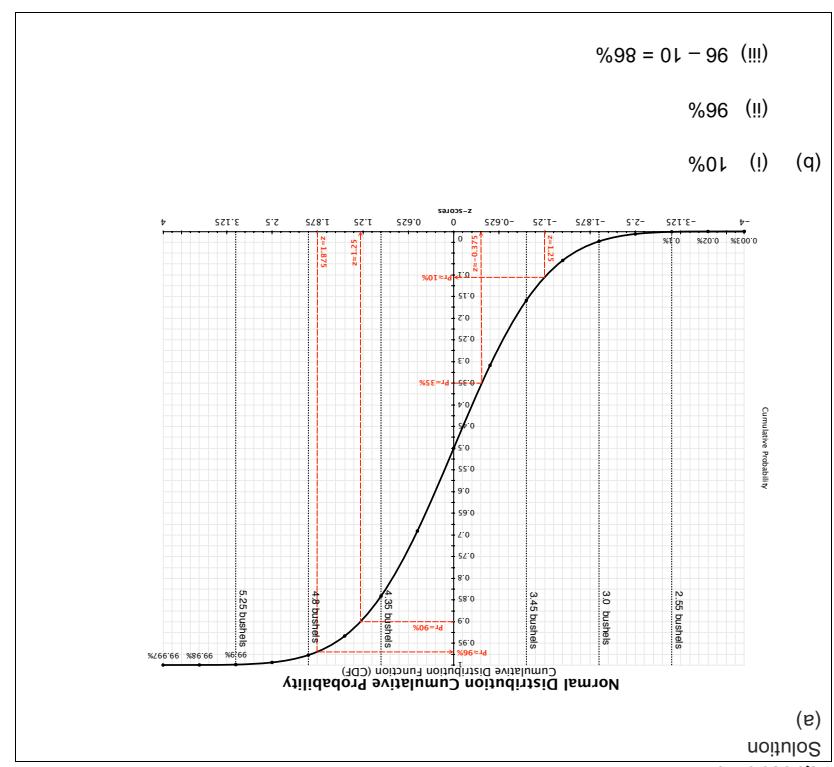
(d) What do you expect the area under the curve to be?

(ii) Label the crop yields in bushels at the bottom of the graph in line with the corresponding standard deviations of the x-axes.

(c) With reference to your graph in part (a),

The graph of $f(x)$ is known as the **probability density function** for the standard normal distribution. Use your calculator to find the limit of $f(x)$ as $x \rightarrow \pm\infty$ to determine how the function behaves relative to the x -axis.

| | |
|-----|--------|
| (a) | 0.1333 |
| (b) | 0.9869 |
| (c) | 0.0373 |



Question 2

(a) $f(x) = \frac{1}{\sqrt{2\pi}} e^{-0.5x^2}$ is a probability density function (PDF) with $\int_{-\infty}^{\infty} f(x) dx$

being the area under $f(x)$. Using this interpretation,

(i) explain the integral $\int_{-\infty}^k f(x) dx$ in terms of both area and probability.

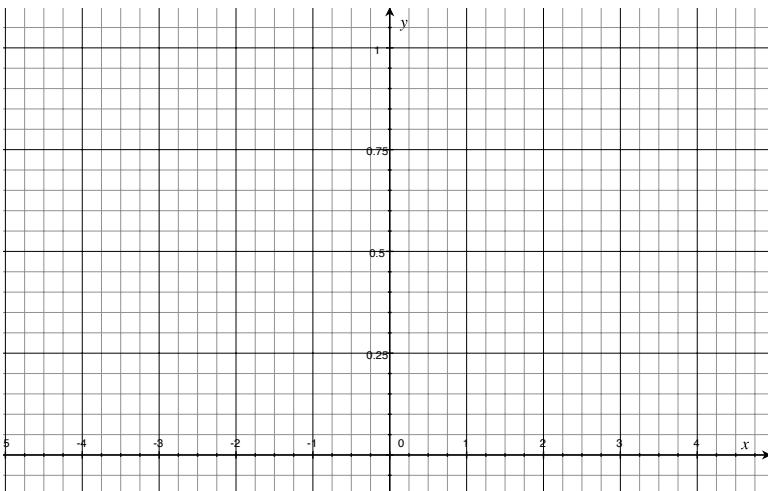
(ii) explain what $P(x) = \int_{-\infty}^x f(t) dt$ represents.

(b) Define $P(x) = \int_{-\infty}^x f(t) dt$ on your calculator and

(i) graph $P(x)$ for $\{-3, -2.5, -2, \dots, 3\}$ on the axes below.

Note: this function is too demanding for the graphing features of your calculator.

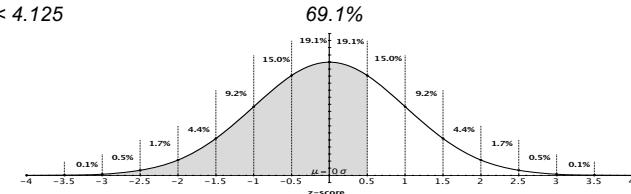
(ii) show crop yields in bushels along the bottom of the graph



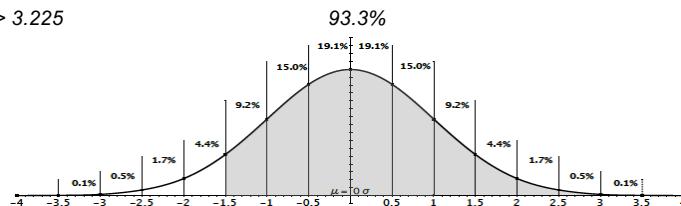
(c) This graph is a **cumulative distribution function** (CDF). Use limits to determine the behaviour of $P(x)$ as $x \rightarrow \pm\infty$ and explain this behaviour.

Question 5**Solution**

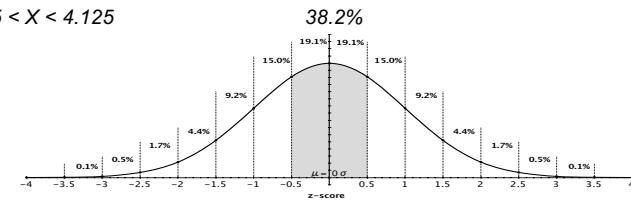
(a) i) $X < 4.125$



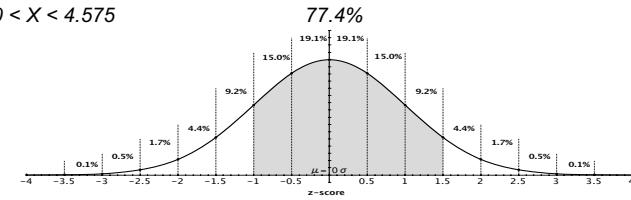
(b) $X > 3.225$



(c) $3.675 < X < 4.125$



(d) $3.450 < X < 4.575$



(e) (i) 15.9%

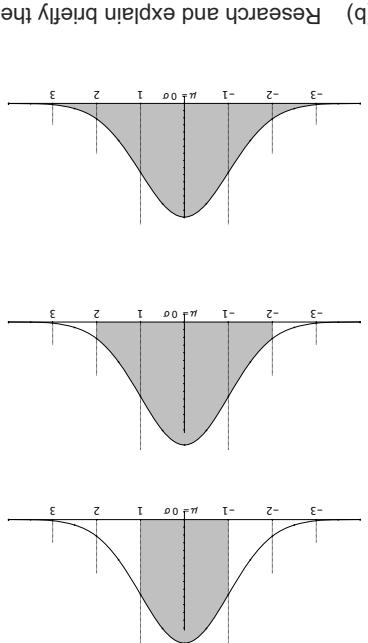
(ii) 6.7%

(iii) 77.4%

$$z = -1 \rightarrow X = 3.45 \text{ bushels}$$

$$z = 1.5 \rightarrow X = 4.575 \text{ bushels}$$

- (d) Given that $a < b$, interpret $P(b)$, $P(a)$ and $P(b) - P(a)$.
- Question 3**
- Cumulative distributions effectively represent cumulative probabilities and these are commonly overlaid onto the probability density function.
- (a) Use appropriate integrals to determine the proportions of the shaded regions for the PDF's shown below.



(b) Research and explain briefly the 68 – 95 – 99.7 rule.

Solution 3

(a) (i) $P(1) - P(-1) = \int_{-1}^1 \frac{1}{\sqrt{2\pi}} e^{-0.5x^2} dx = 0.6827$ $\approx 68.27\%$

(ii) $P(2) - P(-2) = \int_{-2}^2 \frac{1}{\sqrt{2\pi}} e^{-0.5x^2} dx = 0.9545$ $\approx 95.45\%$

(iii) $P(3) - P(-3) = \int_{-3}^3 \frac{1}{\sqrt{2\pi}} e^{-0.5x^2} dx = 0.9973$ $\approx 99.73\%$

(b) For a normally distributed population, 68% of the population will fall within ± 1 S.D. of the mean, 95% will fall within ± 2 S.D.s and 99.7% will fall within ± 3 S.D.s.

Solution 4

(a) $Z - \text{Score} = \frac{\sigma}{X - \mu}$

(i) $Z = \frac{0.45}{4.125 - 3.9} \iff Z = 0.5$

(ii) $Z = \frac{0.45}{3.225 - 3.9} \iff Z = -1.5$

(iii) $Z = \frac{0.45}{2.55 - 3.9} \iff Z = -3$

(iv) $Z = \frac{0.45}{5.025 - 3.9} \iff Z = 2.5$

Question 4

When considering the crop yields we can think about the yields in one of two ways:

- The raw score measurements, i.e. a particular quantity of bushels harvested. These are known as the **X-scores**.
- The number of standard deviations a particular quantity is above or below the mean. These are known as the **z-scores**.

This equivalence will be reinforced each time you scale the crop yield (**X-scores**) against the standard deviations (**z-scores**) in Questions 4 & 5. Converting from the X-scores to the z-scores is known as **standardizing** the scores.

You have already performed these calculations informally. For example, in the case of our wheat crops we know

$$\mu = 3.9 \text{ bushels}$$

$$\sigma = 0.45 \text{ bushels}$$

So a yield of

$$4.8 \text{ bushels} = 3.9 + 2(0.45)$$

$$= \mu + 2\sigma$$

$$3.45 \text{ bushels} = 3.9 - 1(0.45)$$

$$= \mu - \sigma$$

Thus, $X - \text{score} = 4.8 \Rightarrow z - \text{score} = 2$ i.e two S.D.'s **above** the mean (μ).

Whilst, $X - \text{score} = 3.45 \Rightarrow z - \text{score} = -1$ i.e. one S.D. **below** the mean (μ).

(a) Determine and/or research the formula to convert from an X-score to a z-score

(b) Convert the following X-scores to z-scores using $\mu = 3.9$ & $\sigma = 0.45$.

(i) $X = 4.125 \text{ bushels}$

(ii) $X = 3.225 \text{ bushels}$

(iii) $X = 2.550 \text{ bushels}$

(iv) $X = 5.025 \text{ bushels}$

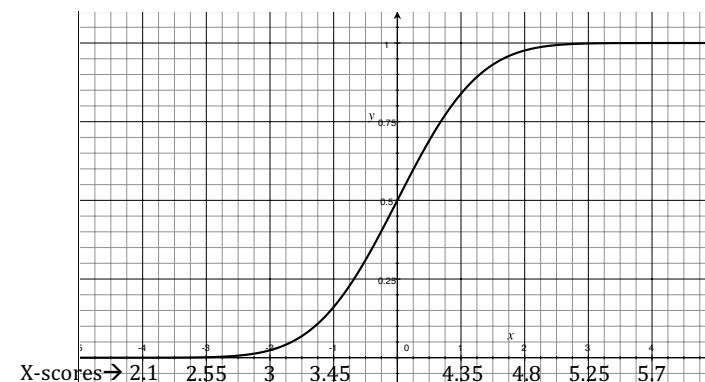
Question 2

Solution

(a) (i) This integral will calculate the area under the PDF from its lowest extreme up to the value k. Since the total area under the PDF = 1, this represents a proportion of the population and, hence, a percentage and a probability.

(ii) $P(x)$ formulates a probability function that gives the probability of a score being in the area below the upper boundary of x .

(b) (i)



(ii)

(c)

As $x \rightarrow -\infty$, $P(x) \rightarrow 0$ As $x \rightarrow -\infty$ the area bounded approaches zero
 As $x \rightarrow +\infty$, $P(x) \rightarrow 1$ As $x \rightarrow +\infty$ the area bounded approaches the total area under the PDF curve = 1

(d) $P(b)$ is the probability of a score being below 'b',
 $P(a)$, the probability of a score being below 'a' and
 $P(b) - P(a)$ the probability of a score falling within the area bounded at the lower end by 'a' and at the upper end by 'b'.

- (e) Two crop yields have their *z-scores* calculated as $z = -1$ and $z = 1.5$.

What is the probability of:

- (i) a harvest yield of *z-score* < -1 .

- (ii) a harvest yield of *z-score* > 1.5 .

- (iii) a harvest yield with $-1 < z\text{-score} < 1.5$.

Determine the actual yields (i.e. *X*-scores) in bushels for each *z-score*.

Inverse Normal CD

Question 9

Use your calculator to confirm and obtain more accurate answers (to 4 d.p.) for Question 8.

- (a) A farmer determines that the lowest 35% of harvests are not financially viable. What is the minimum harvest (in bushels) to be profitable?
- (b) The top 10% of harvests are considered to be too improbable for use in calculating a projected income range. What is the maximum harvest (in bushels) that can be considered?
- (c) Given that wheat fetches \$8.50 per bushel,
 - (i) what would be the expected return (as a range) of a random plot of land?
 - (ii) From the 500 plots of land what would be the expected return in total?

Question 7**CAS Calculator: z-score to probability (x-axis → y-axis)**

$P(x) = \int_{-\infty}^x f(t) dt$ where $f(x) = \frac{1}{\sqrt{2\pi}} e^{-0.5x^2}$ is computationally intensive. So CAS calculators have specialized functions to handle these calculations more efficiently and automatically standardize scores, using the cumulative distribution function to calculate left & right tailed and centred cumulative probabilities.

Notation: To indicate that a variable score X is part of a normal distribution we use the

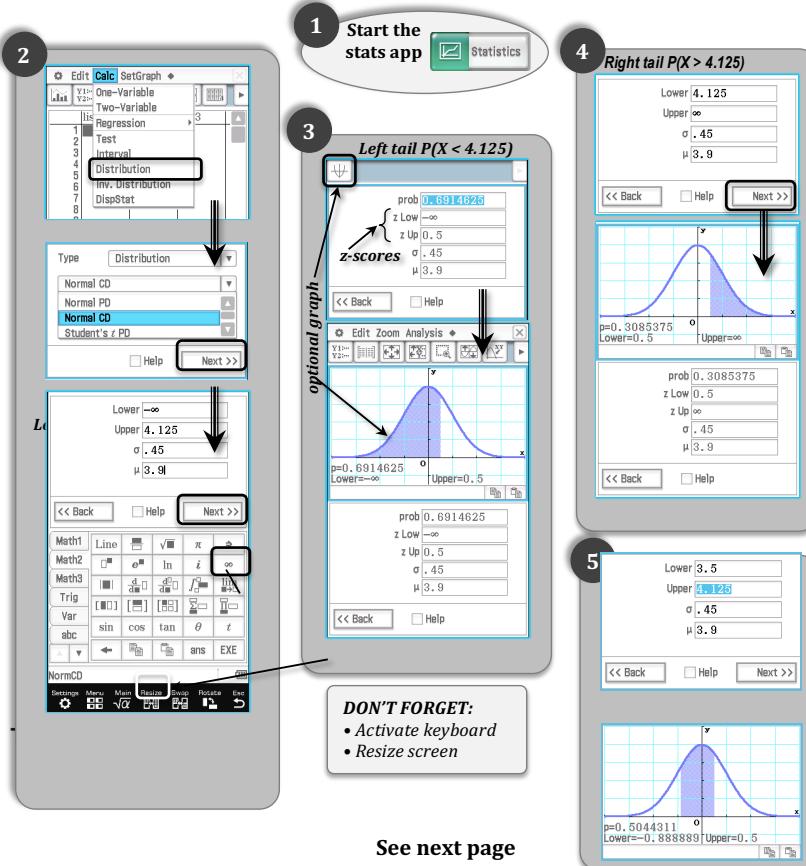
notation: $X \sim (\mu, \sigma^2)$.

Thus for our data set with $\mu = 3.9$ and $\sigma = 0.45$ we write $X \sim (3.9, 0.45^2)$

For example: To determine the proportion of scores below $X = 4.125$ we write

$$P(X < 4.125) \quad \text{where} \quad X \sim (3.9, 0.45^2)$$

Thus on a ClassPad calculator: $P(X < 4.125)$ - a "left-tailed" cumulative probability - is found thus:



See next page

Use your calculator to determine the following (to 4 d.p.)

- (a) $P(X < 3.4)$
- (b) $P(X > 2.9)$
- (c) $P(2.4 < X < 3.1)$

Question 8**Graphical: Probability to z-score (y-axis → x-axis)**

Refer back to the graph of the cumulative distribution function (CDF) on page 11 to answer the following questions. Draw construction lines on the CDF for each question to show your solution.

- (a) A farmer determines that lowest 35% of harvests are not financially viable.
 - (i) What *z-score* does this correspond to?
 - (ii) Hence, what is the minimum harvest (in bushels) to be profitable?
- (b) The top 10% of harvests are considered to be too improbable for use in calculating a projected income range.
 - (i) What is the maximum harvest (in bushels) that can be considered?
 - (ii) Given that wheat fetches \$8.50 per bushel, what would be the expected return (as a range) on a random plot of land?
(Exclude the lowest 35% and the top 10% of the harvests.)

See next page