Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number | Candidate Number

HA061477651

Pearson Edexcel Level 3 GCE

Wednesday 12 June 2024

Afternoon (Time: 2 hours)

Paper reference **9ST0/02**

Statistics

Advanced

PAPER 2: Statistical Inference

You must have:
Statistical formulae and tables booklet
Calculator

Total Marks

Candidates may use any calculator allowed by Pearson regulations. Calculators must not have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B).
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided
  – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Unless otherwise stated, inexact answers should be given to three significant figures.
- Unless otherwise stated, statistical tests should be carried out at the 5% significance level.

Information

- A booklet 'Statistical formulae and tables' is provided.
- There are 7 questions in this question paper. The total mark for this paper is 80.
- The marks for **each** question are shown in brackets
  – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.

Turn over ▶

P75705RA

F:1/1/1/1/1/1/1/1/1/1/

P 7 5 7 0 5 R A 0 1 2 4

Pearson

Answer ALL questions. Write your answers in the spaces provided.

1 Brenda is a yoga teacher. She is researching the benefits of regular yoga practice on reducing eye fatigue.

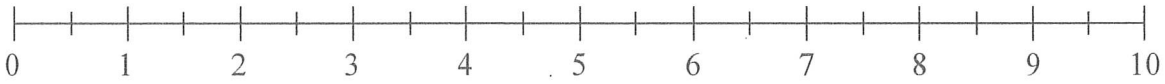
To carry out her research, Brenda seeks 12 volunteer students, who have never practised yoga before, to join her yoga classes.

She asks each student to complete a questionnaire **before** they start her classes.

In her questionnaire, the first question is as follows.

Indicate below your current level of eye fatigue by placing a cross on the line.

A score of 0 indicates no eye fatigue and a score of 10 indicates severe eye fatigue.



After attending her classes for three months, these 12 students are asked again to complete the same questionnaire.

Both scores for the first question, for each of these 12 students, are given in **Figure 1**

Student	A	B	C	D	E	F	G	H	I	J	K	L
Score before yoga	7.4	6.2	8.8	3.5	5.4	6.3	9.5	4.4	6.3	7.8	3.6	5.0
Score after yoga	4.3	5.0	4.7	3.5	6.7	4.5	5.0	6.2	3.5	2.4	4.8	4.5

Figure 1

Brenda believes that regular yoga practice can reduce the average eye fatigue score.

(a) Making any necessary assumptions, carry out a paired *t*-test to investigate whether there is evidence to support Brenda's belief.

(7)

Let X be the eye fatigue score before yoga and let Y be that for after yoga.

Let $D = X - Y$

$H_0: \mu_D = 0$

$H_1: \mu_D > 0$

Use a 1-tailed paired t-test at the 5% level using $\nu = 12 - 1 = 11$

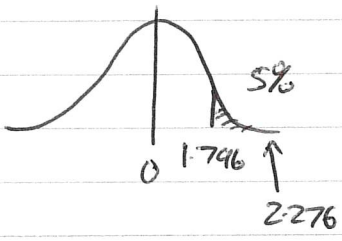
Differences: 3.1 1.2 4.1 0 -1.3 1.8 4.5 -1.8 2.8 5.4 -1.2 0.5

$\bar{d} = 1.59$

$s_d = 2.42$ (calculator)

Test statistic: $\frac{\bar{d} - \mu_D}{\frac{s_d}{\sqrt{n}}} = \frac{1.59 - 0}{\frac{2.42}{\sqrt{12}}} = 2.276$

Critical region



Result significant
Reject H_0

There is significant evidence to suggest, on average, regular yoga practice reduces eye fatigue score.

(b) Give **two** possible sources of bias in this experiment.

(2)

- Eye fatigue score is subjective and based to the student's opinion, which may change over the course of 3 months
- The students all attend her classes, this may not extend to other yoga classes
- Experiment is based to volunteers who are willing to do 3 months of yoga.

etc.

(Total for Question 1 is 9 marks)

2 In October 2021, just ahead of the COP26 UN Climate conference in Glasgow, 75% of adults, aged 18 years or over, in Great Britain agreed with the statement

"I am worried about the impact of climate change."

[Source: Three-quarters of adults in Great Britain worry about climate change Office for National Statistics (ons.gov.uk)]

Briony believed that, in Great Britain, a greater percentage of those aged under 18 were worried about the impact of climate change.

She set up an online survey with two questions.

Question A

Are you under 18?

Thòse who responded with a "no" to **Question A** could not continue with the survey.

Those who responded with a "yes" to **Question A** were then asked **Question B**.

Question B

Consider the statement
"I am worried about the impact of climate change."
Do you agree or disagree with this statement?

Of the 312 responses for **Question B**, 63 **disagreed** with the statement.

Using the results of Briony's survey carry out a hypothesis test, using **a distributional approximation**, to investigate whether the percentage of those aged under 18 who agree with the statement is greater than the percentage of adults in Great Britain who agree with the statement.

You may assume that responses were independent of one another.

(6)

Let X be the number of people under the age of 18 who were worried about climate change.

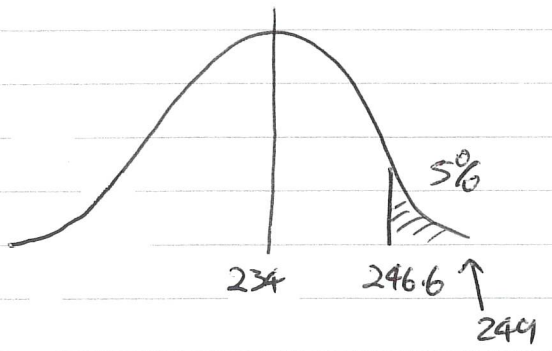
$H_0: \pi = 0.75$
$H_1: \pi > 0.75$

Use a 1-tailed proportion test at the 5% level using $B(312, 0.75) \approx N(234, 58.5)$

$np = 312 \times 0.75 = 234$
$np(1-p) = 312 \times 0.75 \times 0.25 = 58.5$

Test statistic: $312 - 63 = 249$

Question 2 continued



Result significant.
Reject H_0

There is significant evidence to suggest the percentage of people under 18 who are worried about the impact of climate change is higher than 75%.

(Total for Question 2 is 6 marks)

3 Adya, a retail researcher, wants to investigate whether there is a relationship between ages of members of the British public and their preferred supermarket.

Adya decides to post on social media asking members of the public living in her town to complete a survey. Her survey asks each person to categorise themselves as 'student', 'young adult', 'older adult' or 'retired' and to select their preferred supermarket A, B, C, D or E.

Her summarised data is given in **Figure 2**

		Supermarket					
		A	**B**	**C**	**D**	**E**	**Total**
Age category	**Student**	15	3	2	1	2	**23**
	Young adult	11	6	1	3	1	**22**
	Older adult	14	7	5	3	1	**30**
	Retired	6	15	2	1	1	**25**
	Total	**46**	**31**	**10**	**8**	**5**	**100**

Figure 2

Adya decides to perform a test for association on this data and calculates some of the expected frequencies, using the observed frequencies in **Figure 2**

(a) Complete the remaining expected frequencies, correct to one decimal place in **Figure 3**

(2)

Expected frequencies = $\frac{\text{row total} \times \text{column total}}{\text{grand total}}$ $\frac{23\times8}{100} = 1.84$ $\frac{22\times8}{100} = 1.76$

$\frac{30\times8}{100} = 2.4$

Note: Retired/E could be 1.3 if using formula

		Supermarket					
		A	**B**	**C**	**D**	**E**	**Total**
Age category	**Student**	10.6	7.1	2.3	1.8	1.2	**23**
	Young adult	10.1	6.8	2.2	1.8	1.1	**22**
	Older adult	13.8	9.3	3.0	2.4	1.5	**30**
	Retired	11.5	7.8	2.5	2	1.2	**25**
	Total	**46**	**31**	**10**	**8**	**5**	**100**

Figure 3

He last row and column calculated using totals

Question 3 continued

After calculating the expected frequencies, Adya combined the data, in **Figure 2** and **Figure 3**, for supermarkets C, D and E into a new category "Other" in order to make the test of association valid.

(b) Explain why Adya needed to do this.

(1)

All of the expected frequencies for these columns are under 5.
All three columns need to be combined to make them over 5.

Her combined observed frequencies are given in **Figure 4**

		Supermarket			
		A	**B**	**Other**	**Total**
Age category	**Student**	15	3	5	**23**
	Young adult	11	6	5	**22**
	Older adult	14	7	9	**30**
	Retired	6	15	4	**25**
	Total	**46**	**31**	**23**	**100**

Figure 4

(c) Making any necessary assumptions, carry out a test using the data in **Figure 3** and **Figure 4** to investigate whether there is any association between age and preferred supermarket.

(8)

Expected	A	B	Other
Student	10.6	7.1	5.3
Young Adult	10.1	6.8	5.1
Older Adult	13.8	9.3	6.9
Retired	11.5	7.8	5.7

↑ add C, D and E

(e.g. Student/A: $\frac{(15-10.6)^2}{10.6} = 1.826$)

$\frac{(O-E)^2}{E}$	A	B	Other
Student	1.826	2.368	0.017
Young Adult	0.080	0.094	0.002
Older Adult	0.003	0.569	0.639
Retired	2.630	6.646	0.507

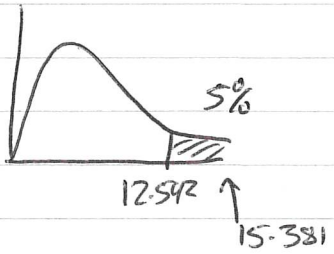
$$\chi^2 = \sum \frac{(O-E)^2}{E} = 15.381$$

H_0: There is no association between age of person and supermarket preference
H_1: There is an association between age of person and supermarket preference.

Use a χ^2 test for association at the 5% level using $\nu = (4-1)\times(3\times1) = 6$

Test stat: 15.381

critical region



Result significant
Reject H_0
There is significant evidence to suggest an association between the age of a person and their supermarket preference.

Question 3 continued

(d) Suggest **two** improvements that Adya could make to this investigation in order to reduce bias.

For each improvement, explain how it could lead to reduced bias.

(4)

- The age categories are self-selected and therefore biased to the opinion of the participant.
  Better to use their actual ages and use age groups to remove this bias

- The survey is posted on Social Media so it is biased to those who want to participate and are on social media.
  Better to use an electoral register and take a simple random sample.

- Only participants from her town are surveyed.
  Better to multiple towns to reduce the bias arising from the opinions of people from her town.

etc.

(Total for Question 3 is 15 marks)

4 In recent years, there have been changes in the patterns of working of employees in businesses throughout the UK. In March 2024, the Office for National Statistics published its latest findings from its ongoing Business Insights and Conditions Survey.

[Source: BICS Survey, www.ons.gov.uk]

Daniel, a researcher in a local university, read the ONS findings and believed that, if larger and smaller businesses in Northern Ireland were surveyed separately, results would have been different.

To test his belief Daniel surveyed businesses with fewer than 100 employees and businesses with more than 100 employees.

In his sample, he questioned, 77 larger businesses with more than 100 employees and 64 smaller businesses with fewer than 100 employees.

Of the 77 **larger businesses** surveyed, 13 stated that they used or intended to use increased homeworking as a permanent business model going forward.

Of the 64 **smaller businesses** surveyed, 9 stated that they used or intended to use increased homeworking as a permanent business model going forward.

(a) Conduct a hypothesis test to investigate whether the proportion of **larger businesses** in Northern Ireland using or intending to use increased homeworking was more than the proportion of **smaller businesses** in Northern Ireland using or intending to use this approach.

(8)

Let X be the number of businesses using or intending to use increased homeworking as a permanent business model for larger businesses and let Y be that for smaller businesses.

$p_x = \frac{13}{77} \approx 16.9\%$ $\quad p_y = \frac{9}{64} \approx 14.1\%$

$n_x = 77$ $\quad n_y = 64$

$H_0: \pi_x - \pi_y = 0$

$H_1: \pi_x - \pi_y > 0$

Use a one-tailed 2-sample proportion test at the 5% level using $N(0,1)$

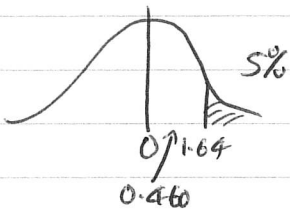
$$p = \frac{\frac{13}{77} \times 77 + \frac{9}{64} \times 64}{77 + 64} = \frac{22}{141}$$

$$\text{Standard error} = \sqrt{\frac{22}{141}\left(1 - \frac{22}{141}\right)\left(\frac{1}{77} + \frac{1}{64}\right)} = 0.06138$$

Question 4 continued

Test statistic : $\frac{\frac{13}{77} - \frac{9}{64}}{0.06138} = 0.460$

Critical region



Result not significant.
Do not reject H_0

There is insufficient evidence to suggest the proportion of larger businesses using a homeworking model is higher than smaller businesses.

(b) Briefly summarise your conclusion to the hypothesis test in a manner suitable for publishing in a newspaper article.

(2)

In Northern Ireland, the percentage of larger businesses using or intending to use increased homeworking is very similar to that of smaller businesses

This may indicate that the size of business may not affect decisions relating to homeworking.

Question 4 continued

(c) State **one** assumption you have made with regards to the sample when carrying out the hypothesis test.

(1)

- Assumed the samples for larger and smaller businesses were obtained randomly

(d) Comment on the likely validity of your conclusion in (a) for businesses across the UK.

(1)

- The businesses selected were all based in Northern Ireland and may not extend to the whole of the UK, so the conclusion may not be valid.

(Total for Question 4 is 12 marks)

5 Miku is interested in testing whether participants skydiving for the first time will experience a higher level of stress than those who had previously been skydiving.

Cortisol is a hormone released by humans when experiencing stressful events and situations and can be measured by either taking a sweat or saliva sample.

Miku chooses to measure the amount of cortisol in sweat samples of 13 first-time skydivers prior to their first jump.

The amounts of cortisol measured, in μg/dL, are shown below.

4.20	4.12	3.76	3.85	3.64	4.35	3.73	4.52	3.82	4.13	4.02	4.10	3.80

[Data source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4373275/]

From a larger study conducted on experienced skydivers it was found that their average cortisol measure prior to jumping was 3.82 μg/dL.

(a) Conduct a Wilcoxon signed-rank hypothesis test to investigate Miku's claim.

State the necessary assumption that must be made about the distribution of cortisol measurements for this test to be valid.

(8)

Let X be the cortisol measure of a first-time skydiver

H_0: $\eta = 3.82$
H_1: $\eta > 3.82$

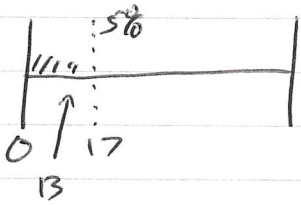
Assumption: The Cortisol measures of first-time skydivers are symmetrically distributed.

Use a 1-tailed Wilcoxon Signed-Rank test at the 5% level using $n = 12$

Differences:	0.38	0.3	-0.06	0.03	-0.18	0.53	-0.09	0.7	0	0.31	0.2	0.28	-0.02
Rank	10	8	3	2	5	11	4	12	X	9	6	7	1
	+	+	−	+	−	+	−	+		+	+	+	−

$T_+ = 65$ $T_- = 13$ so $T = 13$

Critical region:



Result significant
Reject H_0

There is significant evidence to suggest, on average, Cortisol levels of first-time skydivers is higher than 3.82 μg/dL.

Question 5 continued

It was later discovered that the average cortisol levels for experienced skydivers had been measured using saliva samples.

(b) Based on this information, explain why the conclusion in (a) may not be valid.

(1)

Cortisol levels may be different in sweat than in saliva.

The conclusion that Cortisol levels in first time skydives is, on average, greater than 3.82 μg/dL is valid, but not that they may be higher than experienced skydives.

After further research Miku, found that there are several scientific studies that had suggested a strong positive association between the values of cortisol levels in sweat and saliva samples.

Miku found, in a secondary source, cortisol measurements, obtained from both sweat and saliva, for each of 9 adults. These are shown in **Figure 5**

	A	B	C	D	E	F	G	H	I
Sweat reading (μg/dL)	3.74	3.47	2.58	4.12	3.97	4.33	4.67	4.68	4.51
Saliva reading (μg/dL)	6.24	6.31	6.41	6.66	6.72	6.81	7.02	7.12	7.13

Figure 5

Miku believes that the sample data are **not** from a bivariate normal distribution.

(c) Using the data in **Figure 5**, investigate whether there is a positive association between the cortisol levels in sweat and saliva readings.

(6)

H_0: There is no association between Cortisol measurements in sweat and saliva

H_1: There is a positive association between Cortisol measurements in sweat and saliva.

Use a 1-tailed Spearman's Rank test at the 5% level using $n = 9$.

	A	B	C	D	E	F	G	H	I
Rank Sweat	3	2	1	5	4	6	8	9	7
Rank Saliva	1	2	3	4	5	6	7	8	9

DO NOT WRITE IN THIS AREA

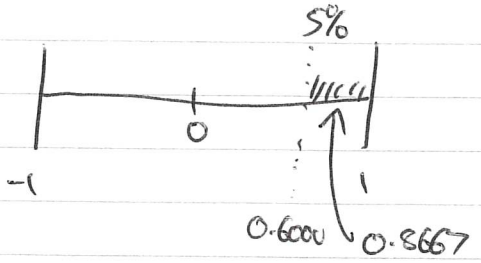
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 5 continued

Test statistic : $r_s = 0.8667$

Critical region :



Result significant
Reject H_0

There is significant evidence to suggest a positive association between Cortisol levels in sweat and saliva

Based on the result of the hypothesis test in (c), Miku decides that the test in (a) is still valid.

(d) Explain why Miku is **not** correct.

(3)

- As Cortisol levels in saliva increases, Cortisol levels in sweat increases
- But there is still no comparison between the levels of Cortisol in sweat and saliva.
- Only an association has been indicated, not a difference in levels.
- This means the conclusion in (a) may still not be valid

(Total for Question 5 is 18 marks)

6 Regularly engaging in aerobic exercise is associated with improved asthma control and quality of life for people with mild asthma.

Adults with mild asthma are recommended to take part in exercise to improve their symptoms.

In 2020 a six-week study was conducted to see if high intensity interval training (HIIT) could improve mild asthma **more** than the usual recommended exercise.

[Data source: Low volume high intensity interval training leads to improved asthma control in adults: Journal of Asthma: Vol 58, No 9 (tandfonline.com)]

Forty adults with mild asthma were randomly assigned to two groups.

The first group participated in a HIIT programme for 20 minutes, 3 times a week.

The second group continued with their usual recommended exercise.

After six weeks the VO2 max levels, the maximum amount of oxygen the body can use while exercising, of these adults, was recorded. This data is shown in **Figure 6**

	VO2 max levels (ml/kg/min)		
	Mean	**Standard deviation**	**Number of adults**
HIIT programme	41.3	4.57	20
Usual exercise	39.0	4.41	20

Figure 6

A higher VO2 max level indicates improved asthma control.

A *t*-test for comparing the mean VO2 max levels was carried out and found to have a *p*-value of 0.0568

(a) Using the information in **Figure 6** calculate the value of Cohen's *d* for the effect of HIIT training on VO2 max levels.

You may assume that VO2 max levels in the two groups are normally distributed and have a common variance.

(2)

$$s = \sqrt{\frac{(20-1)\times 4.57^2 + (20-1)\times 4.41^2}{20+20-2}} = 4.49$$

$$d = \frac{41.3 - 39.0}{4.49} = 0.512$$

Question 6 continued

(b) Use the results of the *t*-test **and** the value of Cohen's *d* calculated in (a) to draw conclusions about the difference in average VO2 max levels for adults with mild asthma participating in the HIIT programme and those who are participating in their usual recommended exercise.

(3)

- The p-value of 0.0568 indicates insufficient evidence at the 5% level that, on average, VO_2 max levels are higher for the HIIT programme than usual recommended exercise
- The d-value of 0.512 indicates a medium effect.
- The results indicate insufficient evidence to suggest the HIIT programme improves mild asthma, however, if it did the effect would be a medium effect.
  This indicates that maybe the investigation should be repeated with a larger sample.

(Total for Question 6 is 5 marks)

7 In January 2023, Michael was working at a financial advisory firm in the North West of England.

He was interested in investigating whether people living in Manchester and Liverpool have different levels of **savings**.

He decided to use his company records for his investigation. An extract of these records is shown in **Figure 7**

	A	B	C	D	E	F	G
1	**Customer_ref**	**Surname**	**First**	**Branch**	**Type_of_account**	**Asset_value (£)**	**Savings_account (£)**
2	1022564	Kelly	Nigel	Preston	Flex	312 458	18 210
3	1022565	Littler	Mark	Manchester	Fixed	256 874	34 116
4	1022566	Leng	Moira	Liverpool	High risk	303 451	26 415
5	1022567	Tang	Peter	Chester	Fixed	512 744	22 457
6	1022568	Prest	Heather	Manchester	Low risk	211 110	10 223

Figure 7

(a) Explain how Michael could use a spreadsheet function to help **select** appropriate data for this investigation.

(2)

Filter Column D to display only Manchester or Liverpool, and use the data in Column G.

Question 7 continued

Having selected the relevant data Michael calculated the following summary statistics for levels of savings.

City	COUNT	SUM (£)	STDEV (£)
Liverpool	50	314250	425
Manchester	42	268800	511

(b) Making any necessary assumptions, carry out a hypothesis test to investigate whether people living in Manchester and Liverpool have different **average** levels of savings.

(8)

Let X be the levels of savings for people in Liverpool and Y be that for Manchester.

$\bar{x} = \frac{314250}{50} = 6285$ $\quad$ $\bar{y} = \frac{268800}{42} = 6400$

$s_x^2 = 425^2$ $\quad$ $s_y^2 = 511^2$

$n_x = 50$ $\quad$ $n_y = 42$

$H_0: \mu_Y - \mu_X = 0$

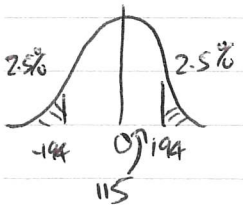
$H_1: \mu_Y - \mu_X \neq 0$

Use a two-tailed 2-sample z-test at the 5% level using:

Method 1

$\bar{Y} - \bar{X} \sim N\left(0, \frac{425^2}{50} + \frac{511^2}{42}\right)$

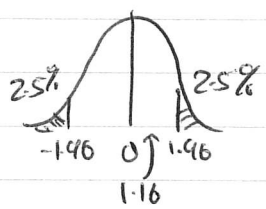
test stat: $6400 - 6285 = 115$



Method 2

$Z \sim N(0,1)$

$\frac{(6400 - 6285) - 0}{\sqrt{\frac{425^2}{50} + \frac{511^2}{42}}} = 1.16$



Method 3

$P(\bar{Y} - \bar{X} \geq 115) = P(Z \geq 1.16)$

$= 0.123$

p-value $= 2 \times 0.123$

$= 0.246$

> 0.05

Result not significant

Do not reject H_0.

There is insufficient evidence to suggest a difference in average levels of savings between people in Manchester and Liverpool.

Question 7 continued

(c) Explain

(i) why the only error that could have occurred in (b) is a Type II error,

(1)

The result of the test was that we did not reject H_0, and this can only be a Type II error.

(ii) the meaning of this error in the context of (b)

(1)

Concluding that there is no difference in the average levels of savings between people of Manchester and Liverpool, when in fact there was.

Question 7 continued

During November 2022, a **large** survey was carried out across the United Kingdom relating to savings.

It was found that, on average, the difference between savings per household for those in Manchester and in Liverpool was actually £270

[Source: Savings statistics: Average savings in the UK (finder.com)]

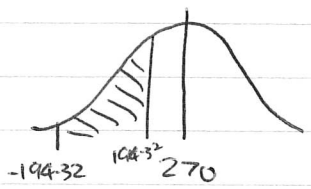
(d) Using the information from this large survey, calculate the probability of the possible Type II error described in (c)

(3)

in (b) Do not reject H_0 if $-194.32 \leq \bar{Y} - \bar{X} \leq 194.32$

(using $\bar{Y} - \bar{X} \sim N\left(0, \frac{425^2}{50} + \frac{511^2}{42}\right)$)

If $\bar{Y} - \bar{X} \sim N\left(270, \frac{425^2}{50} + \frac{511^2}{42}\right)$

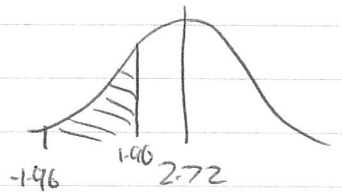


$P(-194.32 \leq \bar{Y} - \bar{X} \leq 194.32) = 0.223$

Alternatively: $z\text{-shift} = \frac{270 - 0}{\sqrt{\frac{425^2}{50} + \frac{511^2}{42}}} = 2.72$,

using $Z \sim N(0,1)$, do not reject H_0 if $-1.96 \leq Z \leq 1.96$

If $Z_T \sim N(2.72, 1)$



$P(-1.96 \leq Z_T \leq 1.96) = 0.223$

(Total for Question 7 is 15 marks)

TOTAL FOR PAPER IS 80 MARKS

BLANK PAGE

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

BLANK PAGE

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

BLANK PAGE