

Please check the examination details below before entering your candidate information

Candidate surname					Other names									
Pearson Edexcel					Centre Number					Candidate Number				
Level 3 GCE					<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>					<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>				
Goodness of Fit Tests														
<div> <div></div> <div></div> </div>														
Statistics Advanced Topic Test														
You must have: Statistical formulae and tables booklet Calculator										Total Marks <input type="text"/>				

**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 5 questions in this question paper. The total mark for this paper is 50.
- 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.

1. The caterpillars of the Great Mormon butterfly can develop into female adult butterflies of different colours.

According to a genetic theory, the butterflies of species will be black and blue, black and red or pale white in the ratio 5 : 3 : 2.

A random sample of caterpillars is collected, of which 132 develop into female Great Mormon butterflies, with the following results recorded:

Butterfly Colour	Black and Blue	Black and Red	Pale White
Frequency	82	19	31

Use a χ^2 test at the 5% significance level to see whether there is sufficient evidence to suggest the genetic theory is not supported by the data.

(Total 9 marks)

2. A small haulage company keeps a record of the number of customer complaints which are received on any particular week.

Number of customer complaints	0	1	2	3	4	5
Frequency	6	28	46	54	47	69

A new manager is hired by the company and wishes to review the performance of the company over the previous few years.

In order to analyse the frequency of customer complaints, it is suggested that the manager models the number of customer complaints using a Poisson distribution.

- (a) Estimate λ , the mean number of customer complaints received per week.

(1)

Elle is a statistician hired by the company who has started a Goodness of Fit test for these data.

Number of customer complaints	Observed frequency	Poisson Probability (4 d.p.)	Expected frequency (2 d.p.)
0	6	0.0384	9.60
1	28	0.1251	31.29
2	46	0.2040	51.00
3	54	0.2217	55.42
4	47	s	u
5 or more	69	t	v

- (b) Using the value of λ you found in part (a), find the values of s , t , u and v .

(2)

- (c) Carry out a χ^2 test to see whether there is evidence to suggest a Poisson distribution is not suitable for modelling the number of customer complaints received per week.

(7)

(Total 10 marks)

3. A surgical team has 5 surgery slots available for a new surgical procedure during any particular day. The risk of a patient developing a post-operative infection is suspected to be 2%.

The hospital administrator is monitoring the outcomes of this new procedure and records the number of patients who develop post-operative infections on any single day, for a random sample of 102 days.

Number of patients who develop post-operative infections	Number of days
0	81
1	16
2	4
3	1
4	0
5	0

Test to determine whether there is significant evidence to doubt that $B(5,0.02)$ is a suitable model for the number of patients operated on in a randomly chosen day who develop post-operative infections.

(Total 10 marks)

4. Eniola is the head of a large computing department that keeps records from which she can calculate the time, in hours, between network outages. She thinks that an exponential distribution with parameter $\lambda = 0.05$ will be a suitable model for the time between successive network outages. Eniola calculates the times between 51 consecutive network outages and produces the following table.

Time (in hours)	0 – 12	12 – 24	24 – 48	48 – 72	72 – 96
Observed frequency	24	13	8	4	1
Expected frequency	22.559	12.381	10.524	3.170	0.955

- (a) Show how Eniola calculated the expected frequency for the class 24 – 48 hours. (3)
- (b) Explain **two** adjustments which need to be made before a χ^2 Goodness of Fit test can be rigorously carried out. (2)

The adjustments required to complete the χ^2 Goodness of Fit test are made and the test statistic is correctly calculated as 0.405 to 3 significant figures.

- (c) Complete the test to determine whether there is significant evidence to doubt that an exponential model with parameter $\lambda = 0.05$ is a suitable model for the time intervals between network outages. (4)
- (d) Using your answer in part (c), suggest a possible distribution for the number of network outages per hour. (1)

(Total 10 marks)

5. A photocopier produces documents of nominal resolution 300 DPI. The resolutions of a random sample of 395 documents were measured and the results were recorded in a spreadsheet as follows:

	A	B	C	D	E	F
1	Min DPI		255	285	315	345
2	Max DPI	255	285	315	345	
3	Frequency	87	38	96	88	86

- (a) Given that there were no resolutions lower than 225 DPI or higher than 375 DPI, estimate the mean and standard deviation of the resolutions to 2 decimal places.

(1)

A χ^2 test is to be used to ascertain the suitability of the normal distribution to model the resolutions of documents. This test is recorded in the spreadsheet printout.

	A	B	C	D	E	F	G
1	Min DPI		255	285	315	345	
2	Max DPI	255	285	315	345		
3	Frequency	87	38	96	88	86	
4							Total
5	Probability	0.129365	0.203182	0.271415	0.227541		1
6	Expected	51.09907	80.25680	107.20894	89.87886		395
7	Contribution	25.22309	22.24905	1.17192	0.03928		
8							

- (b) Explain how spreadsheet software can be used to calculate the value in cell B6.

(2)

- (c) Calculate the values that should appear in cells F5, F6, F7 and G7.

(4)

- (d) Hence, complete the test to determine whether there is significant evidence to doubt the suitability of the normal distribution to model the resolution of documents.

You must clearly state your hypotheses.

(4)

(Total 11 marks)

TOTAL FOR PAPER: 50 MARKS