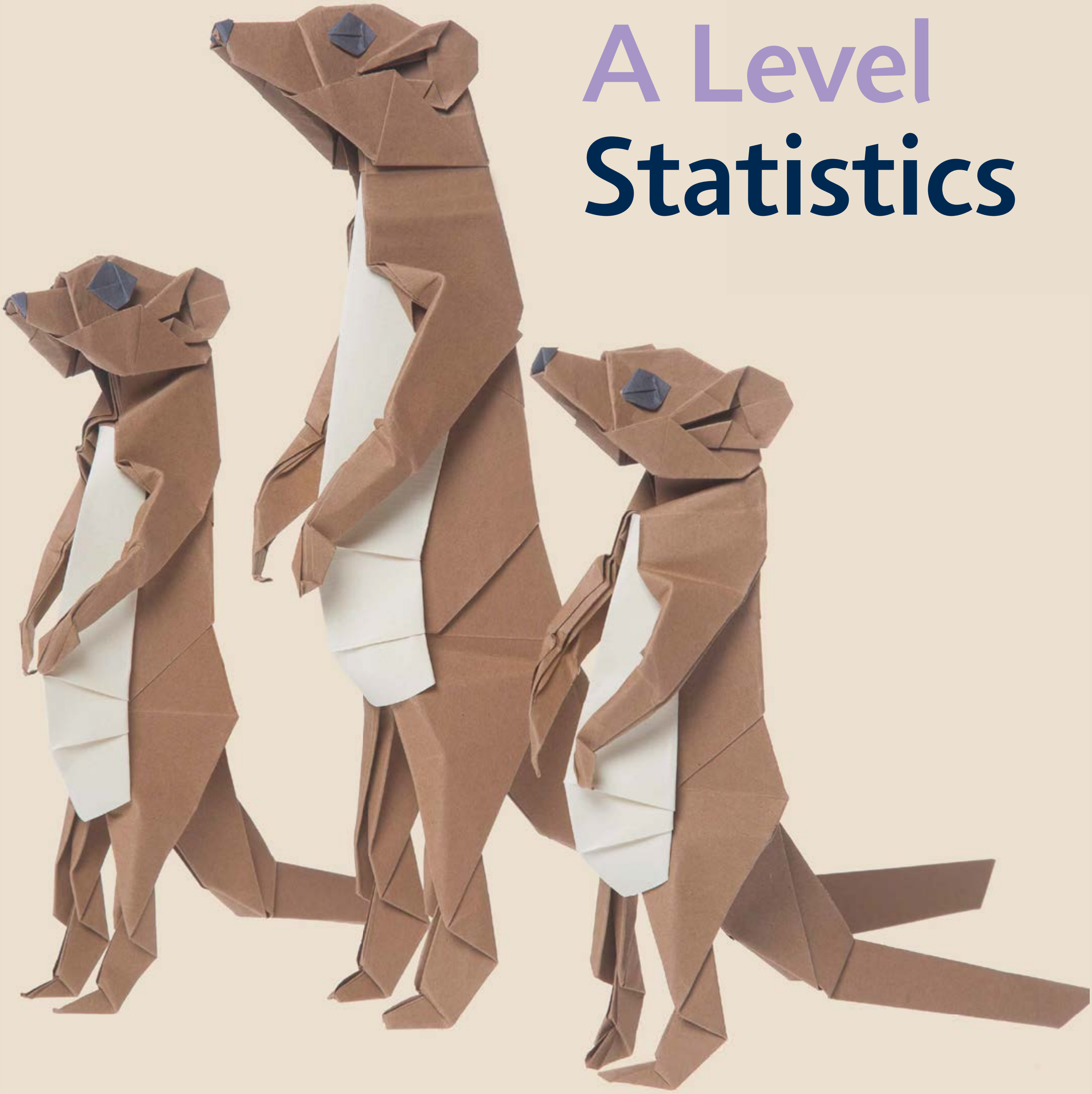
edexcel

A Level Statistics



Sample Assessment Materials

Pearson Edexcel Level 3 Advanced GCE in Statistics (9ST0)

First teaching from September 2017

First certification from 2019

Issue 1

Pearson

Edexcel, BTEC and LCCI qualifications

Edexcel, BTEC and LCCI qualifications are awarded by Pearson, the UK's largest awarding body offering academic and vocational qualifications that are globally recognised and benchmarked. For further information, please visit our qualification website at qualifications.pearson.com. Alternatively, you can get in touch with us using the details on our contact us page at qualifications.pearson.com/contactus

About Pearson

Pearson is the world's leading learning company, with 35,000 employees in more than 70 countries working to help people of all ages to make measurable progress in their lives through learning. We put the learner at the centre of everything we do, because wherever learning flourishes, so do people. Find out more about how we can help you and your learners at qualifications.pearson.com

References to third party material made in this sample assessment materials are made in good faith. Pearson does not endorse, approve or accept responsibility for the content of materials, which may be subject to change, or any opinions expressed therein. (Material may include textbooks, journals, magazines and other publications and websites.)

All information in this document is correct at time of publication.

Original Origami Artwork designed by Beth Johnson and folded by Mark Bolitho
Origami photography: Pearson Education Ltd/Naki Kouyioumtzis

ISBN 978 1 4469 4613 8

All the material in this publication is copyright
© Pearson Education Limited 2017

Contents

Introduction

The Pearson Edexcel Level 3 Advanced GCE in Statistics is designed for use in schools and colleges. It is part of a suite of AS/A Level qualifications offered by Pearson.

These sample assessment materials have been developed to support this qualification and will be used as the benchmark to develop the assessment students will take.

The booklet '*Statistical formulae and tables*' will be provided for use with these assessments and can be downloaded from our website, qualifications.pearson.com.

General marking guidance

- All candidates must receive the same treatment. Examiners must mark the last candidate in exactly the same way as they mark the first.
- Mark schemes should be applied positively. Candidates must be rewarded for what they have shown they can do rather than be penalised for omissions.
- Examiners should mark according to the mark scheme – not according to their perception of where the grade boundaries may lie.
- All the marks on the mark scheme are designed to be awarded. Examiners should always award full marks if deserved, i.e. if the answer matches the mark scheme.
- Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme.
- Where some judgement is required, mark schemes will provide the principles by which marks will be awarded and exemplification/indicative content will not be exhaustive. However different examples of responses will be provided at standardisation.
- When examiners are in doubt regarding the application of the mark scheme to a candidate's response, a senior examiner must be consulted before a mark is given.
- Crossed-out work should be marked **unless** the candidate has replaced it with an alternative response.

Guidance on the use of codes within this mark scheme:

- **M** – Mark is for method
- **Mdep** – Mark is dependent on one or more M marks and is for method
- **A** – Mark is dependent on M or m marks and is for accuracy
- **B** – Mark is independent of M or m marks and is for method and accuracy
- **E** – Mark is for explanation
- **Ft** – Follow through from previous incorrect result
- **cao** – Correct answer only
- **cso** – Correct solution only
- **awfw** – Anything which falls within
- **awrt** – Anything which rounds to
- ***** – Answer given
- **SC** – Special case
- **o.e.** – Or equivalent
- **A2, 1** – 2 or 1 (or 0) accuracy marks
- **sf** – Significant figure(s)
- **dp** – Decimal place(s)

No method shown

Where the question specifically requires a particular method to be used, we must usually see evidence of use of this method.

Where the answer can be reasonably obtained without showing working and it is very unlikely that the correct answer can be obtained by using an incorrect method, we must award full marks. However, the obvious penalty to candidates showing no working is that incorrect answers, however close, earn no marks.

Where a question asks the candidate to state or write down a result, no method need be shown for full marks.

Where the permitted calculator has functions which reasonably allow the solution of the question directly, the correct answer without working earns full marks, unless it is given to less than the degree of accuracy accepted in the mark scheme, when it gains no marks.

Otherwise we require evidence of a correct method for any marks to be awarded.

Write your name here

Surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number | Candidate Number

Statistics

Advanced
Paper 1: Data and Probability

Sample Assessment Material for first teaching September 2017

Time: 2 hours

Paper Reference **9ST0/01**

You must have:
Statistical formulae and tables booklet, calculator

Total Marks

Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for algebraic manipulation, differentiation and integration, or 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.
- Inexact answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Statistical formulae and tables' is provided.
- There are 11 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

S58391A

©2017 Pearson Education Ltd.

1/1/1/

S 5 8 3 9 1 A 0 1 2 4

Pearson

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

1 On a weekday, calls received by a company's customer service helpdesk may be regarded as occurring at random with a constant average rate of 7.0 per hour.

Find the probability that the total number of calls received by the helpdesk during a period of two hours on a particular weekday is fewer than 10.

(3)

(Total for Question 1 is 3 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

2 David, a zoologist, investigated the snout-vent length for two species of Tropidus lizard.

He found that the snout-vent lengths for the species Spinulosus are normally distributed with mean 89.6 mm and standard deviation 15.4 mm.

He found that the snout-vent lengths for the species Insulanus are normally distributed with mean 75.2 mm and standard deviation 8.5 mm.

Sketch the distributions of snout-vent lengths for each species on a single diagram that David could use to illustrate his findings.

Source: Journal of Morphology, 2001

(3)

(Total for Question 2 is 3 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

3 The table below gives the prize values, £X, available in a national weekly lottery, and the approximate probabilities of winning each available weekly prize.

X	Probability
1 000 000	0.000 000 022
1 000	0.000 006 9
100	0.000 46
10	p
0	0.899 533 078

(a) Find the value of p.

(1)

(b) Maraig buys a ticket costing £2 in each of the 52 national weekly lottery draws during one year.

Find her expected loss, to the nearest penny, during the one year period.

(3)

(Total for Question 3 is 4 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

4 During mornings at a computer centre, the time, in minutes, between successive job submissions can be modelled by an exponential distribution with $\lambda = 4$

(a) For such mornings

(i) state the mean time between successive job submissions

(1)

(ii) find the probability that less than 1 minute will elapse between successive job submissions.

(2)

(b) Give one reason why this exponential model may not apply to the time, in minutes, between successive job submissions at this computer centre during afternoons.

(1)

(Total for Question 4 is 4 marks)

5 A vehicle insurance company classifies drivers as high risk, medium risk or low risk with regard to having an accident.

The company estimates that 20% of its insured drivers are high risk and 30% are low risk.

The probability that a high risk driver will have one or more accidents in a year is 0.06, and the corresponding probabilities for medium risk drivers and low risk drivers are 0.03 and 0.01 respectively.

(a) The company supplies insurance to a driver and within one year the driver has an accident.

Find the probability that the driver was classified as high risk.

(3)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(b) Bill is a new employee of the insurance company and is investigating driving accidents.

He states that the company should not offer insurance to drivers in its high risk category because they will have six times the number of accidents in a year than its low risk drivers.

Comment, giving reasons, on the validity of Bill's statement. You should include some numerical justification.

(3)

(Total for Question 5 is 6 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

6 The Parent Teacher Association (PTA) of a rural primary school wants to raise funds for the school by holding car boot sales. An enquiry to the National Association of PTAs reveals that the profits from such sales may be modelled by a normal distribution with mean £206 and standard deviation £28.

The PTA is aiming to raise £300 to pay for some ukuleles for the school and considers holding two car boot sales.

Two farmers from different villages each offer a field for the PTA to use for one car boot sale.

(a) Assuming that the profits from these two sales are independent, find the probability that the total profit will **not** be sufficient to pay for the ukuleles.

(5)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

Dougal, a local businessman, offers to match any money that the PTA raises.
(For example, if the PTA raises £200, then Dougal will also contribute £200.)

Following this offer, the PTA considers holding only one car boot sale.

(b) Find the probability that the profit from one car boot sale, together with Dougal's contribution, will **not** be sufficient to pay for the ukuleles.

(3)

(c) Comment on the advisability of holding **one** car boot sale, rather than two, in order to raise money for the ukuleles.

(2)

(d) Describe, briefly and in context, a situation that would lead you to doubt the validity of the independence assumption made in part (a).

(1)

(Total for Question 6 is 11 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

7 A new trainee, Dylan, in the town-planning department for a large town is investigating likely parking needs for a new housing development of 30 individual homes. He is required to plan adequate parking places for all the homes.

Dylan obtains up-to-date information, to use in his planning, from the DVLA (Driver & Vehicle Licensing Agency) regarding car registrations. The following table summarises this information that Dylan wants to use.

Number of cars registered per household	**0**	**1**	**2**	**$\geqslant$3**
Probability	0.18	0.47	0.25	0.10

Making any necessary assumptions, use an exact distribution and Dylan's information, to find the probability that, for the new development

(a) (i) more than 15 homes will have at most one car

(2)

(ii) at least 27 homes will have fewer than 3 cars.

(2)

(b) Explain why Dylan may not gain the accurate information that he needs regarding parking requirements for the new development using the DVLA data.
Include a suggestion to Dylan regarding his planning for parking needs.

(3)

(Total for Question 7 is 7 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

8 A trial was carried out to investigate the effect of an antibiotic treatment, amoxicillin, for an acute ear infection in children aged between 6 months and 2 years.

Details of the design of the trial were stated as follows:

> *After we obtained consent, the children were randomly assigned to treatment with amoxicillin or with a placebo suspension with the same colour and taste.*
>
> *The suspensions were supplied to the participating doctors in a double blind fashion with two block randomisation; doctors, parents, and investigators remained blinded throughout the study.*

Source: British Medical Journal, 2000

(a) Explain, in the context of this trial, the purpose of

(i) the random assignment to treatment

(2)

(ii) a double blind trial.

(2)

(b) Explain how a randomised block design could be set up for this trial and why this would be preferable to a completely randomised design.

You should include a diagram to illustrate your suggested randomised block design.

(5)

(Total for Question 8 is 9 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

9 A survey of students taking a politics module at university were asked which, if any, of three influential political leaders, *N*, *A* and *F*, they admired.

Of these students:

60 admired *N*
55 admired *A*
21 admired *F*
45 admired *N* and *A*
12 admired *A* and *F*
14 admired *N* and *F*
8 admired **all three** leaders
and 1 admired **none** of the leaders.

For a randomly selected student
Event **N** is defined as 'student admires *N*'
Event **A** is defined as 'student admires *A*'
Event **F** is defined as 'student admires *F*'

(a) Draw a fully labelled Venn diagram to illustrate this information.

(4)

(b) State the number of these students who admired exactly one of the leaders.

(2)

(c) What proportion of these students who do not admire leader *F*, admire both of the other two leaders?

(3)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

(d) Find the probability that exactly 2 of a random sample of 3 of these students, will admire only N.

(4)

(Total for Question 9 is 13 marks)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

10 The first time that anyone ran one mile in under 4 minutes was in May 1954.
Figure 1 gives the world record time, t seconds in excess of 3 minutes 30 seconds, for running one mile as of 31st May in various subsequent years.

Year (x)	1955	1960	1965	1970	1975	1980	1985	1990	1995	2000
Time (t)	28.0	24.5	24.1	21.1	21.0	19.0	17.3	16.3	14.4	13.1

Figure 1

Figure 2 shows the scatter diagram of t plotted against year.

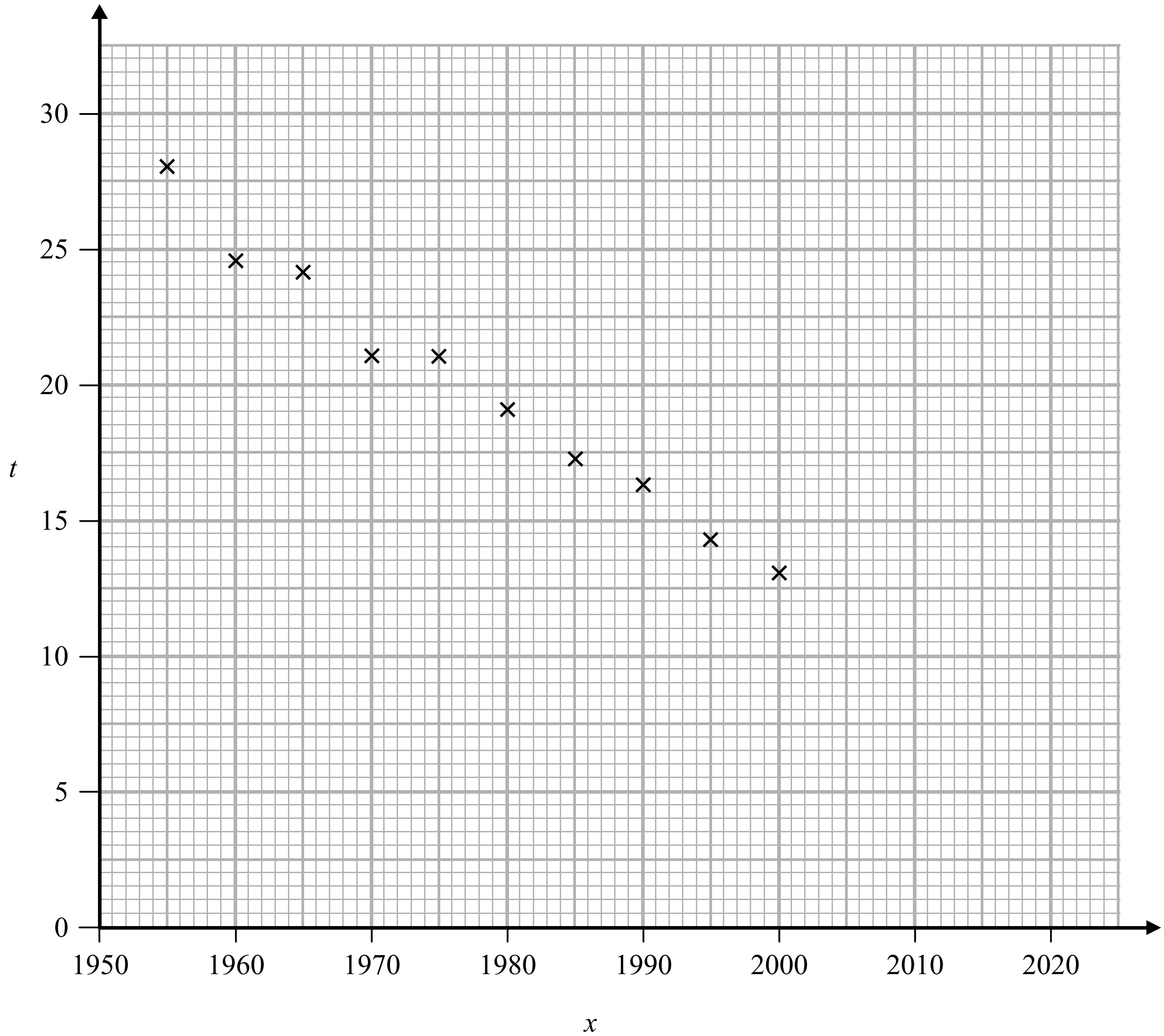


Figure 2

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(a) (i) Calculate an equation, using the data given in Figure 1, that can be used to predict the actual world record time for running one mile as of 31st May 2010.
Give your predicted time to the nearest one-tenth of a second.

(5)

(ii) Comment on the validity of your prediction in part (a)(i).

(1)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(b) The equation is also used to predict the world record time for running one mile as of 31st May 2020.

In fact, as of March 2016, the world record time for running one mile had not been broken since July 1999.

How would this information affect the prediction for 2020?

(2)

(c) The total value, £y, of government investment into athletics during the previous twelve months was also recorded on the same 10 dates as those in Figure 1.

It was observed that the value of y increased each time.

(i) State the value of Spearman's rank correlation coefficient for the 10 pairs of values of y and t.

(1)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

(ii) Using this information, a newspaper reported under the headline:

MORE INVESTMENT NEEDED SO OUR ATHLETES CAN BEAT THE WORLD!

Give **two** reasons to doubt the validity of this statement.

(2)

(Total for Question 10 is 11 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

11 Read the following two extracts on school workforce statistics for England in November 2015.

Extract 1

More than half a million children are being taught by unqualified teachers, Labour will reveal.

The number of staff without the key classroom grade of "qualified teacher status" leapt by 52% since ministers eased rules, the party's research found.

Labour examined Department for Education statistics showing there were 22,500 unqualified teachers in primary and secondary schools in England in 2015, with average class sizes of 25.4 pupils.

That means nearly 572,000 kids were taught by staff without the formal qualification.

Extract 2

The percentage of qualified teachers (those with Qualified Teacher Status) in all state funded schools is 95.1 per cent; a decrease from 95.5 per cent in 2014.

The total number of teachers without QTS increased between 2014 and 2015; from 20.3 thousand FTEs in 2014 to 22.5 thousand FTEs in 2015. This represents an increase in the percentage of FTE teachers without QTS; from 4.5 per cent to 4.9 per cent in 2015. As reported by schools the percentage of teachers without QTS that were working towards gaining QTS had also increased; from 15.0 per cent in 2010 to 19.2 per cent in 2014 to 20.0 per cent in 2015.

The percentage of FTE teachers that do not have QTS varies by school type. 3.1 per cent of teachers in all nursery/primary schools do not have QTS; compared with 5.9 per cent in all secondary schools.

(a) For these two extracts

(i) comment on the difference in report style.

(ii) state, with reasons, the likely target audience.

(7)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(b) Identify two drawbacks in using one or both of these extracts as a data source.

(2)

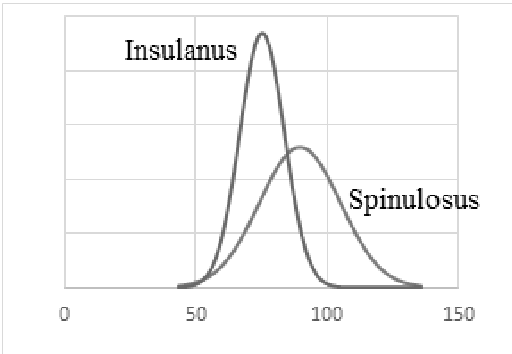
(Total for Question 11 is 9 marks)

TOTAL FOR PAPER IS 80 MARKS

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Paper 1: Data and Probability Mark Scheme

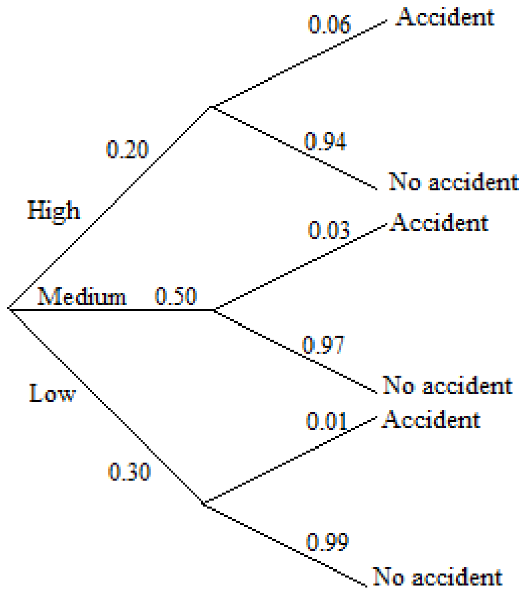
Question	Scheme	Marks	AO	Notes
1	$X \sim \text{Po}(14)$	M1	1.2	X (number of emails recd during two hours) ~ Po
		M1	2.1a	M1 for $\lambda = 14$ stated or implied
	$\text{P}(Y < 10) = \text{P}(Y \leqslant 9) = 0.109$	A1	1.2	awrt 0.109
	Total	**3**		

Question	Scheme	Marks	AO	Notes
2		B1	1.1	Two 'bell-shaped' curves drawn.
		M1	1.1	Clear difference in means seen.
		A1	1.1	Overall shape (heights and spread) correct and labelled.
	Total	**3**		

Question	Scheme	Marks	AO	Notes
3(a)	0.1	B1	1.2	
3(b)	E(X) = (0×0.899533078) + 10×0.10 + 100×0.00046 +1000×0.0000069	M1	1.2	Attempt at $\sum x \times P(X = x)$
	+1000000×0.000000022 = 1.0749			1.0749 or 1.075
	Exp loss = (2 – 1.0749) × 52	M1	2.1a	[2 – 'their' E(X)] × 52
	= £48.11	A1	1.2	cao
	Total	**4**		

Issue 1 – September 2017 © Pearson Education Limited 2017

Question	Scheme	Marks	AO	Notes
4(a)(i)	Mean = $\frac{1}{\lambda} = \frac{1}{4}$ minute (or 15 seconds)	B1	1.2	Correct units required.
4(a)(ii)	P(time < 1) $= 1 - e^{-4}$ $= 0.982$	M1 A1	1.2 1.2	
4(b)	λ (time between jobs) may be different in afternoons. **or** There may be follow-up calls/ follow-on problems from morning jobs so events may not be independent. **or** Rate of calls may not be constant in the afternoon.	E1	3.1a	One relevant comments made in context (other relevant comments allowed).
	Total	**4**		

Question	Scheme	Marks	AO	Notes
5(a)	P(accident)	M1	1.2	Attempt to obtain P(accident).
	$= 0.012 + (0.5 \times 0.03) + (0.3 \times 0.01)$ $= 0.03$	M1dep	1.2	
	$\text{P(High} \mid \text{Accident)} = \dfrac{0.012}{0.03} = 0.4$	A1	1.2	Correct fraction. $\dfrac{\text{P(High} \cap \text{Accident)}}{\text{P(Accident)}}$
	Alternative – Use of tree diagram gains full marks  			
5(b)	From part (a) For any one accident, it is 4 times more likely to be a High risk driver rather than a low risk driver involved.	E1	2.1b	Comment relating to probability found in part (a).
	Probability may be 6 times higher but this does not mean 6 × the number of accidents. **or** $\text{P(Low} \mid \text{Accident)} = \dfrac{0.003}{0.03} = 0.1$	E1	2.1b	Other numerical back up.
	Bill's statement is not true.	E1dep	2.1b	Dep any one valid reason given.
	Total	**6**		

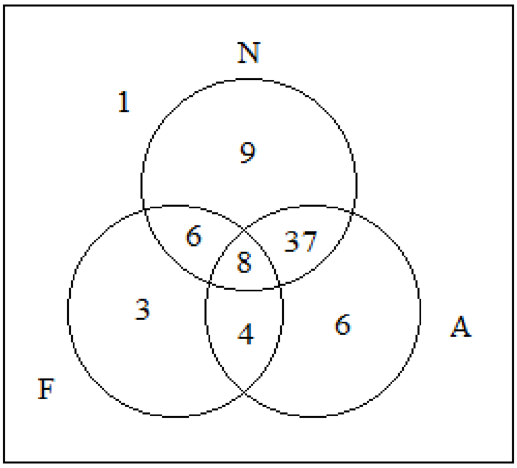
Question	Scheme	Marks	AO	Notes
6(a)	X is profit from one car boot sale X~N(412, 1568)	B1	1.2	Mean 412
	Standard deviation $= \sqrt{28^2 + 28^2} = 39.598$	B1	1.2	sd = 39.6 (3sf) (**or** variance = 1568)
	$z = \frac{(300 - 412)}{39.598} = -2.828$	M1	1.2	Correct attempt to find z = awrt $\pm$ 2.83 Also award for attempted z using continuity correction 299.995 giving $z = -2.8285\ldots$
	Prob not sufficient = 1 – 0.99765	M1dep	1.2	Dep previous M1 (may be awarded in (b) if not scored here).
	= 0.00235	A1	1.2	**or** 0.00234 from correct CC. Accept 0.0022 ~ 0.0024
6(b)	Normal with mean 412	B1	1.2	(Only award if not earned in (a))
	Standard deviation = 2 × 28 = 56	B1	1.2	56 cao (or variance = 3136)
	$z = \frac{(300 - 412)}{56} = -2$	M1	1.2	$z = \pm 2$ Also award for attempted use of continuity correction 299.99(5) used which gives all answers correct to 3sf
	Prob not sufficient = 1 – 0.97725 = 0.02275	A1	1.2	0.0227 or 0.0228
6(c)	Prob in (b) larger than in (a) but still small.	E1dep	2.1a	Dep answer (b) >

Question	Scheme	Marks	AO	Notes
6(c) continued	Probably OK to just hold one car boot sale.	E1dep	2.1b	answer (a)
6(d)	Examples include: - Many of the same people may attend both car boot sales. - Sales close together in time and customers may not have so much spare cash for the second sale. - The weather may affect both equally. -In this area car boot sales are particularly popular/unpopular. -Feedback from the first sale may affect attendance at the second sale.	E1	3.1a	Any sensible comment that may relate to a positive correlation or negative correlation between profits.
	Total	**11**		

Question	Scheme	Marks	AO	Notes
7(a)(i)	Using binomial $X \sim B(30, 0.65)$	M1	2.1a	May be implied Allow use of $Y \sim B(30, 0.35)$
	$P(X \geqslant 16) = 0.935$	A1	1.2	**or** $P(Y \leqslant 14)$
7(a)(ii)	Using binomial $X \sim B(30, 0.9)$	M1	2.1a	May be implied somewhere in (a). Allow use of $Y \sim B(30, 0.1)$
	$P(X \geqslant 27) = 0.647$	A1	1.2	**or** $P(Y \leq 3)$
7(b)	DVLA probabilities of numbers of cars per household applies in Dylan's town / probabilities constant.	E1	3.1a	p not the same for Dylan's homes/town.
	For example, the DVLA data may be for the whole of the UK but Dylan's town is poor/wealthy so would have different probabilities to those for the whole country. **or** Dylan's development may be for all very small or all very large homes so probabilities for his particular type of homes would not be as supplied by the DVLA.	E1dep	3.1a	In context Dep previous E1
	Suggest that Dylan obtains local information regarding the numbers of cars per household for the type of homes in the new development.	E1	3.1a	
	Total	**7**		

Question	Scheme	Marks	AO	Notes
8(a)(i)	Purpose is to ensure that no bias exists in assignment of a child to a treatment	E1	3.1a	Idea of no bias.
	Situations such as younger children being allocated the antibiotic/sicker children being allocated the antibiotic/pushy parents being allocated to the antibiotic will be avoided	E1	2.1a	Valid explanation in context.
8(a)(ii)	Purpose is to avoid expectations by children/parents or medical staff of a faster/slower recovery if nature of treatment is known	E1	3.1a	Expectations of children/parents the same.
	and to ensure that all children are treated the same regardless of whether an antibiotic or a placebo is given	E1	2.1a	Treatment of children by parents/medical staff the same.
8(b)	Other factors such as: Severity of infection, age of child, other medical conditions.	E1	3.1a	Examples of other factors (than the treatment allocated) that may affect outcome (possibly implied in table).
	May affect the outcome and so the most important of these should also be taken into account.	E1	3.1a	Reasons that these should be taken into account – in context.

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td rowspan="3">8(b) continued</td><td>Example
<table><tr><td></td><td colspan="3">Severity of infection</td></tr><tr><td></td><td>Mild</td><td>Moderate</td><td>Severe</td></tr><tr><td>Amox</td><td></td><td></td><td></td></tr><tr><td>Placebo</td><td></td><td></td><td></td></tr></table></td><td>B1</td><td>1.1</td><td>Table with a blocking factor incorporated produced.</td></tr>
<tr><td></td><td>B1</td><td>1.1</td><td>Fully labelled correct possible randomised block design.</td></tr>
<tr><td>A completely randomised design could result in all the children with mild infections (oe) being allocated to the amoxicillin for example. This would give a result that may not indicate the effectiveness of the amoxicillin.</td><td>E1</td><td>3.1a</td><td>Explanation in context.</td></tr>
<tr><td></td><td>Total</td><td>9</td><td></td><td></td></tr>
</table>

Question	Scheme	Marks	AO	Notes
9(a)	ξ 	M1 A1 A1 A1	1.1 1.1 1.1 1.1	3 intersecting "circles" drawn and labelled with box outside. 8 placed in N∩F∩A region. 1 placed outside 'circles'. All values correctly placed.
9(b)	$6 + 9 + 3 = 18$	M1 A1	1.2 1.2	Addition of numbers placed in correct areas. Correct answer.
9(c)	Attempt to find correct conditional prob: $P(N \cap A \mid F')$ $\frac{37}{53} = 0.698$ (3sf)	M1 M1 A1	2.1a 1.2 1.2	PI 53 used as denominator (PI) and reasonable attempt at numerator. Answer correct.
9(d)	$\frac{9 \times 8 \times 65}{74 \times 73 \times 72} \times 3 = \frac{195}{5402} = 0.0361$ (3sf)	M1 M1 M1 A1	1.2 1.2 1.2 1.2	Numerator correct. Denominator correct. $\times 3$ used. Answer correct.
	Total	**13**		

Question	Scheme	Marks	AO	Notes
10(a)(ii)	Method for finding correct least squares regression line.	M1	1.2	PI
	Slope = −0.3117576 Intercept = 636.3806061	B1	1.2	Both −0.32 ~ −0.31 and 636 ~ 637
	When x = 2010, predicted t = 9.7479	M1	1.2	For finding a predicted t
	Giving the predicted world record time = 3m 30s + 9.7	M1	1.2	For 3m 30s + their predicted t
	= 3m 39.7s	A1	1.2	cao
10(a)(ii)	**Either** "Extrapolation, so unreliable" **or** "Prediction is only just outside range **and** plot suggests a strong relationship, so reliable".	E1	3.1b	
10(b)	Record time has remained the same. Rate of decreasing times may be levelling off.	E1	2.1a	oe
	The prediction from the model for 2020 is likely to be too low.	E1	2.1b	
10(c)(i)	$r_s = -1$	B1	2.1b	cao
10(c)(ii)	Correlation does not imply causality.	E1	3.1b	For comment on causality. oe
	Examples of other acceptable comments: Data only refers to 1-mile runners not all athletes. **or** World record holders don't necessarily win all races. **or** Inflation not accounted for. **or** Newspaper reporting can sometimes be unreliable.	E1	3.1b	Any other relevant comment besides causality.
	Total	**11**		

Question	Scheme	Marks	AO	Notes
11(a)(i)	Extract 1 is presents the information in a simpler, more informal way. Extract 1 mentions class size and refers to total numbers but Extract 2 does not. Extract 1 exhibits political bias. Extract 2 quotes more detailed summarised percentages than Extract 1 Extract 2 includes an unexplained abbreviation whereas Extract 1 does not. Extract 2 breaks down unqualified teachers by school category but Extract 1 does not.	E1 E1 E1	1.1 1.1 1.1	For any three differences identified Other valid differences allowed

Issue 1 – September 2017 © Pearson Education Limited 2017

Question	Scheme	Marks	AO	Notes
11(a)(ii)	Extract 1 is targeted at a newspaper audience/people not in the education profession/Labour party members because, for example, Extract 1 keeps the information simple and reports findings from the Labour party. **or** Extract1 tries to grab attention with selected bits of information, for example, '572,000 kids were taught by staff without the formal qualification'.	B1 E1	2.1b 2.1b	Identifies target audience as not (education) professionals/as labour party supporters/ as newspaper readers. Any valid reasoning.
	Extract 2 is targeted at education professionals because, for example, Extract 2 assumes that the target audience has knowledge of the terms used relating to teachers QTS, FTE and the systems within which teachers work. **or** Extract 2 could be in a professional magazine.	B1 E1	2.1b 2.1b	Identifies target audience as people in (education) professions. Any valid reasoning.

Question	Scheme	Marks	AO	Notes
11(b)	Extract 1 does not include information about school category (nursery, primary, secondary). In Extract 1 the average class size mentioned is not clarified as to exactly what figures have been averaged. Extract 2 does not include information about the actual numbers of teachers, only percentages. Extract 2 does not refer to the source of the data. (Extract 1 refers to Department of Education as the source) Both extracts only contain summarised results and no raw data.	E1 E1	3.1a 3.1a	Two drawbacks for Extracts 1 and/or 2 identified. Other valid drawbacks accepted.
	Total	**9**		

Write your name here

Surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number | Candidate Number

Statistics

Advanced
Paper 2: Statistical Inference

Sample Assessment Material for first teaching September 2017

Time: 2 hours

Paper Reference **9ST0/02**

You must have:
Statistical Formulae and Tables booklet, calculator

Total Marks

Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for algebraic manipulation, differentiation and integration, or 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.
- Inexact answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Statistical Formulae and Tables' is provided.
- There are 8 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

S58392A

©2017 Pearson Education Ltd.

1/1/1/

S 5 8 3 9 2 A 0 1 2 4

Pearson

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

1 An Italian restaurant, close to a university, was considering a new recipe for the topping used on one of its variety of pizzas.

A random sample of 8 students was selected and each student was asked to score the tastes of the original topping and the proposed new topping on a scale from 1 to 10

The restaurant wanted to investigate whether there was any difference in the average scores for the original and new toppings.

A Wilcoxon signed-rank test was performed at the 5% level of significance on the 8 pairs of scores. This gave a test statistic of $T = 5$

(a) State the null and alternative hypotheses for this test.

(1)

(b) Find the critical value for the test and state the conclusion reached from carrying out the test.

(3)

(Total for Question 1 is 4 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

2 An intervention scheme for disadvantaged Year 10 students is introduced by a Local Authority.

In order to investigate the impact of the scheme, a sample of n such students is obtained. The students are required to take a standardised test before the scheme starts.
After the scheme is completed, an equivalent standardised test is taken by the same group of students. The scores, before and after the scheme, are recorded for each student.
A higher score indicates a higher impact.

Suggest an analysis that could be carried out on the results in order to investigate the impact of the scheme.
You should include the name of test selected and the hypotheses used.
You should also state the assumptions necessary for your selected test to be valid.

(4)

(Total for Question 2 is 4 marks)

3 A company purchasing department wants to trial a new machine to be used for the manufacture of steel shafts. The company wants to produce shafts of length 35 millimetres.

The new machine is programmed to produce shafts of length 35 mm and is made available for the company to have on trial for one day.

Raluka, the quality control assistant for the company, takes a sample of shafts from the production of the new machine. The lengths, in mm, of the 10 shafts that Raluka obtained from the new machine are given below.

34.78 34.72 34.87 35.63 35.68 35.92 36.33 35.77 35.26 35.94

(a) Assuming that the lengths of shafts produced by the new machine may be modelled by a normal distribution, calculate a 95% confidence interval for the mean shaft length for those shafts produced by the new machine.

(4)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

Issue 1 – September 2017 © Pearson Education Limited 2017

(b) Raluka also has available the recorded information on the lengths of shafts sampled from those produced by the company's current machine on the same day.

These 36 recorded lengths had a mean of 34.90 mm and a standard deviation of 2.15 mm.

Calculate a 95% confidence interval for the mean shaft length based on Raluka's second sample.

(3)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(c) Explain why the confidence interval could be calculated in part (b) despite no information regarding the distribution of the shaft lengths from the current machine being given.

(2)

(d) Write a brief report for the purchasing department of the company, comparing the performances of the new and the old machines.

Nobody in the purchasing department is familiar with statistics.

(2)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

(e) The director of the company considers the confidence intervals constructed from Raluka’s data. He sends her an e-mail saying that he needs a more precise estimate of the mean shaft lengths produced by the new machine.

He still wants a 95% confidence interval, but suggests that Raluka should take a bigger sample because “the confidence interval is then certain to be narrower than the interval obtained from Raluka’s sample of 10 shafts”.

Explain why this is not necessarily the case.

(2)

(Total for Question 3 is 13 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

4 The Number 8 bus and the Number 700 bus both travel from a seaside village to the centre of a nearby town but they follow different routes. Each bus does not have a set timetable but claims to run every 10 minutes.

For each route, the journey time from the village to the town centre may be assumed to be an independently normally distributed random variable with standard deviation 1.8 minutes.

Denzil travels regularly from the village to the town centre and he takes whichever bus leaves first. He has free bus travel so the cost of the journey is not an issue but his friend, Hugo, suggests he should consider journey time as well.

Consequently Hugo persuades Denzil to record the journey times on a number of randomly selected days. On these days, there were 12 journeys on the Number 8 bus and 9 journeys on the Number 700 bus.

The sample mean journey time for each route is as follows.

Number 8 bus	Sample mean = 15.2 minutes
Number 700 bus	Sample mean = 17.5 minutes

(a) Use an appropriate test, at the 5% significance level, to determine whether the mean time taken to travel from the village to the town centre is the same for both bus routes.

(7)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(b) Suggest two improvements that Denzil might make to the way that he collected his data.

(2)

(c) Hugo wants to be able to give Denzil more precise advice about which bus to catch. What other information would you suggest that Hugo might usefully obtain? Explain your answer.

(2)

(Total for Question 4 is 11 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

5 Each year, from May to September, steam trains for tourists run along an old rural railway line in Devon. The operating company is interested in the extent to which the trains are used by Devon residents.

In 2015, during the busiest month of August, a survey of a random sample of 240 passengers showed that 36 of them were residents of Devon.

(a) Use an approximation to investigate, at the 5% level of significance, whether there is support for the assertion that more than 80% of passengers in August are tourists from outside Devon.

(7)

(b) (i) Give one reason why this survey would probably provide a biased view of the extent to which the train is used by Devon residents.

(1)

(ii) Do you think the survey would underestimate or overestimate usage of the train by Devon residents?

Give a reason for your answer.

(2)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

(c) It was suggested that, for a follow-up survey, it would be easier to collect data from all the passengers on a particular train.

Explain in context why it would not be appropriate to then apply the test that you used in part (a).

(2)

(Total for Question 5 is 12 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

BLANK PAGE

6 Roma, the manager of a warehouse store for trade members only, was interested in the lengths of time that customers spent shopping in the store.

Customers are required to swipe their club membership cards on entry and exit of the store.

Roma measured her time to walk at a steady pace around the store, visiting every aisle. Her time was 20 minutes. She suspected that, on average, her customers spend less than 20 minutes in the store.

Roma wants to investigate her customers' average times in her store because she believes that redesigning the layout of the store may result in customers staying longer and, hopefully, spending more money.

The time, x minutes, spent by each of a random sample of 250 customers was obtained from card swipe information and may be modelled by a normal distribution.

It was found that

$$\sum x = 4785 \text{ and } \sum x^2 = 98584$$

(a) Investigate, at the 1% level of significance, whether this sample provides evidence to support Roma's suspicion.

(6)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(b) Roma's assistant, Jurgen, is taking a course in Statistics.

He criticises the validity of the test carried out in part (a) because he says that the standard deviation of shopping times is not known.

Jurgen also criticises the 1% significance level used in the test carried out in part (a). He states that using a 5% significance level would make it more difficult to reject H_0 and this may then provide more convincing evidence to support Roma's suspicion.

Comment on both of Jurgen's criticisms.

(4)

(Total for Question 6 is 10 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

7 Patients who attend a particular clinic are routinely prescribed drug A to take over a long period of time. It was found that 98 of 220 randomly selected patients reported that they suffered from a dry mouth as a side effect of taking drug A.

The doctor who leads the clinic wants to investigate whether this side effect is just as likely to occur when a different, but equally effective, drug B is prescribed.

The doctor prescribes drug B to 150 of the patients attending his clinic during the following year. It was recorded that 86 of these 150 patients suffered from the side effect.

(a) Carry out a test to investigate whether there is a difference between the proportion of patients suffering from the side effect when taking drug A and when taking drug B. Use the 5% level of significance.

(8)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(b) It was later discovered that the doctor had explained to each patient that he saw in his clinic the purpose of prescribing the alternative drug. He then asked for volunteers to try drug B.

Explain how, if at all, this information might affect the conclusion that you made in part (a).

(2)

(Total for Question 7 is 10 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

8 Richard retired a few years ago and bought a small farm. On his farm he has three fields that each contain a similar number of apple trees.

When Richard was talking to his friend Krys, Richard said that he suspected there may be an association between the quality of apples and the field in which they grow.

Richard had picked the ripe apples from each of his fields that morning and had them in three sacks, one from each field. Krys has studied some statistics and suggested that he could test Richard's suspicion. Richard agreed and sorted the apples in each sack into those he considered to be of 'Excellent', 'Good' or 'Satisfactory' quality.

He found that:

- From the 'Top' field there were 36 'Excellent apples, 36 'Good' apples and 21 'Satisfactory' apples.
- From the 'Middle' field there were 18 'Excellent apples, 50 'Good' apples and 31 'Satisfactory' apples.
- From the 'Bottom' field there were 30 'Excellent apples, 42 'Good' apples and 36 'Satisfactory' apples.

(a) Making any necessary assumptions, carry out a test on the available data to see if there is any evidence to support Richard's suspicion. Use the 5% level of significance.

(11)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Richard lives near a town where there is an annual Mistletoe Fair. He knows that mistletoe, a popular Christmas decoration, grows as a parasite on several varieties of host tree, including apple trees.

He has also read that apple trees that are used as mistletoe hosts produce fewer apples as a result.

Richard decides that, as an experiment, he will use the apple trees in one of his three fields as hosts for growing mistletoe.

(b) Using the data from the three sacks together with the analysis in part (a), which of Richard's three fields would you recommend that he should use for growing mistletoe?

Give a reason for your recommendation.

You should include a comment on the reliability of the conclusion of the test in part (a).

Also, suggest what further investigations Richard could undertake to improve the reliability of the recommendation of where to grow mistletoe, if at all.

(5)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(Total for Question 8 is 16 marks)

TOTAL FOR PAPER IS 80 MARKS

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Paper 2: Statistical Inference Mark Scheme

Question	Scheme	Marks	AO	Notes
1(a)	H_0 : population median difference = 0 H_1 : population median difference ≠ 0	B1	1.3	Both oe using η
(b)	Two-tailed test 5% CV = 4	B1	1.3	
	TS > CV so accept H_0	M1	2.1b	May be implied by correct conclusion.
	No significant evidence (at the 5% level) of a difference between average taste scores between new and original topping.	A1	2.1a	In context.
	Total	**4**		

Question	Scheme	Marks	AO	Notes
2	Paired t or z test (as standardised results)	M1	2.1b	Paired test.
	H0 $\mu_{difference} = 0$ H1 $\mu_{difference} < 0$	A1	1.3	Both hypotheses (in equivalent words gains marks). Difference is 'before – after'. oe
	Assuming differences in standardised scores are normally distributed and students selected at random.	E1 E1	2.1b 2.1b	Differences in scores ~ N Students selected at random.
	or			**or**
	Paired Wilcoxon signed-rank test	M1		Paired test.
	H0 $\mu_{difference} = 0$ H1 $\mu_{difference} < 0$ **or** H0 $\eta_{difference} = 0$ H1 $\eta_{difference} < 0$	A1		Both hypotheses.
	Assuming differences in scores symmetrically distributed and students selected at random.	E1E1		Differences in scores symmetrically distributed. Students selected at random.
	or			**or**
	Paired sign test	M1		Paired test.
	H0 $\eta_{difference} = 0$ H1 $\eta_{difference} < 0$	A1		Both hypotheses.
	Assuming students selected at random. No distributional assumptions necessary.	E1E1		Students selected at random. No distributional assumption needed.
	Total	**4**		

Question	Scheme	Marks	AO	Notes
3(a)	$\bar{x} = 35.49 \quad s = 0.5539 \quad n = 10$	B1	1.2	$\bar{x}$ and s correct (possibly implied by correct interval).
	$35.49 \pm 2.262 \times \dfrac{0.5539}{\sqrt{10}} =$	M1	1.3	Formula correct using t or z and $\dfrac{s}{\sqrt{10}}$ (possibly implied by correct interval).
		B1	1.3	$t = 2.26$ used (possibly implied by correct interval).
	(35.09, 35.89)	A1	1.3	awrt
3(b)	$34.90 \pm 1.96 \text{ (or } 2.03) \times \dfrac{2.15}{\sqrt{36}} =$	M1	1.3	Formula correct using(t **or**) z and $\dfrac{s}{\sqrt{36}}$ oe (possibly implied by correct interval).
		B1	1.3	$z = 1.96$ (or $t = 2.03$) (possibly implied by correct interval).
	(34.20, 35.60) using $z = 1.96$ (34.17, 35.63) using $t = 2.03$	A1	1.3	awrt
(c)	The sample size is sufficiently large.	B1	2.1b	
	In this case, the Central Limit theorem states that sample means are approximately normally distributed regardless of the distribution of the population that the sample is taken from.	E1	3.1a	CLT states sample means approximately normally distributed.

Question	Scheme	Marks	AO	Notes
(d)	New machine is off-target for the average length of shafts/mean length likely to be above 35mm. oe Variability of shaft lengths greater for old machine. **or** New machine produces shafts of more consistent length than old machine. (explanations must consider target audience - otherwise E0 E0)	E1 E1	2.1a 2.1a	Simple ref to means/averages. Simple ref to spread.
(e)	Because, with new sample values, the sample standard deviation might increase to such an extent that $t \times \frac{SD}{\sqrt{n}}$ is larger, making the interval wider.	E1 E1	3.1a 3.1a	Some consideration of standard deviation/variance /variation. Needs recognition that standard deviation may increase **and** clear acknowledgement that it may more than counteract the effect of increasing n (and decreasing t).
	Total	**13**		

Question	Scheme	Marks	AO	Notes
4(a)	$H_0: \mu_x = \mu_y$ $H_1: \mu_x \neq \mu_y$ $\bar{x} = 15.2 \quad n_x = 12 \quad \sigma_x = 1.8$ $\bar{y} = 17.5 \quad n_y = 9 \quad \sigma_y = 1.8$	B1	1.3	*X* represents times on Number 8 *Y* represents times on Number 700 Hypotheses correct.
	ts $z = \dfrac{(17.5 - 15.2) - 0}{1.8\sqrt{\dfrac{1}{12} + \dfrac{1}{9}}}$	M1	1.3	15.2 -17.5 used in formula attempt.
		M1	1.3	$\sqrt{\dfrac{1}{12} + \dfrac{1}{9}}$ in denominator.
	= 2.898 ts = 2.90 (3sf)	A1	1.3	awfw (2.89 - 2.90) Ignore sign.
	$cv = 1.960 \quad p = 0.00376$	B1	1.3	correct *p*-value awfw 0.0037 to 0.0038 implies M1A1
	2.898 >1.960 or 0.00376 < .05	M1	2.1b	*cv* correct and compared with ts. Allow −2.898 < −1.960 *p*-value compared to correct sig level.
	There is significant evidence (at the 5% level) to suggest that the mean times taken are not the same. The number 700 bus appears to take, on average, longer to make the journey.	E1dep	2.1a	Conclusion correct and in context dep test all correct.

Question	Scheme	Marks	AO	Notes
4(b)	Obtain information on times on the same number of occasions for each number bus (not 12 to 9). Ensure consideration of other (blocking) factors e.g. weather, day of week evenly for each bus route. Introduce a random system for selecting which number bus to use on each occasion. Have a larger sample size meaning that Denzil should undertake more bus journeys for his research.	E1 E1	3.1a 3.1a	Max 2 marks for relevant improvements in context.
4(c)	Useful to know when the previous 8 and 700 buses departed. If the quicker number 8 bus is due to leave less than 2.3 minutes after the number 700 bus, Denzil should arrive earlier if he waits for the number 8.	E1 E1	3.1b 3.1b	Some consideration of likely time to next bus. Recommendation it may be better to wait for the number 8.
	Total	**11**		

Question	Scheme	Marks	AO	Notes
5(a)	$H_0 : p = 0.80$ $H_1 : p > 0.80$	B1	1.3	Hypotheses. Alternatively can use $H_0 : p = 0.20$ $H_1 : p < 0.20$ and sample proportion of 0.15
	Using normal approximation to binomial.	M1	2.1b	
	Using proportions, $z = \dfrac{0.85 - 0.80}{\sqrt{\dfrac{0.80 \times 0.20}{240}}} = 1.936$	M1	1.3	General form for props or numbers; allow +/–, any n
	or Using numbers **without** CC $z = \dfrac{204 - 240 \times 0.8}{\sqrt{240 \times 0.8 \times 0.2}} = 1.936$	M1	1.3	Completely correct expression; allow 0.85 for 0.80 in denominator.
	$z = 1.94$ (3sf)	A1	1.3	1.93 ~ 1.94
	Alternative Using numbers **with** CC	M1	1.3	General form as above.
	$z = \dfrac{203.5 - 240 \times 0.8}{\sqrt{240 \times 0.8 \times 0.2}} = 1.8558$	M1	1.3	Completely correct expression; allow 0.85 for 0.8 in denominator.
	$= 1.86$ (3sf)	A1	1.3	1.85 ~ 1.86

Question	Scheme	Marks	AO	Notes
5(a) continued	Critical value is 1.6449 Reject H_0 at the 5% level	M1	2.1b	1.64 ~ 1.65 **or** p = 0.0264 (0.026 ~ 0.027) < 0.05 **or** p = 0.0317 (0.031 ~ 0.032) < 0.05 from correct CC
	There is significant evidence (at the 5% level) that more than 80% of customers in August are visitors from outside Devon.	E1dep	2.1a	Conclusion correct and in context dep test all correct
5(b)(i)	Survey only carried out in August (peak month).	E1	3.1a	
5(b)(ii)	Likely to underestimate…	B1dep	3.1a	Underestimate with valid effort at a reason.
	…because more tourists from further away expected in the busy holiday season (**or** during school holidays).	E1	3.1a	Devon residents likely to avoid busiest time of year or other valid reason.
5(c)	e.g. Not a random sample/trials not independent/populations different.	E1	3.1a	Relevant reason.
	e.g. Most are travelling in groups so are likely to come from the same place.	E1dep	3.1a	In context.
	Total	**12**		

Question	Scheme	Marks	AO	Notes
6(a)	$\bar{x} = 19.14 \quad s = 5.302$	B1	1.2	For 19.1 ~ 19.2 and s_{n-1} =5.30~ 5.31 or s_n = 5.291 (5.29 ~ 5.30)
	$H_0 : \mu = 20$ $H_1 : \mu < 20$	B1	1.3	Both hypotheses
	$z = \frac{19.14 - 20}{5.302/\sqrt{250}}$	M1	1.3	Completely correct formula attempt for z
	$= -2.56$ $p = 0.005 \sim 0.006$	A1	1.3	–2.57 ~ –2.56
	Critical value $z = -2.3263$ $-2.56 < -2.32$ or $p < 0.01$	M1	2.1b	Allow 2.56 > 2.32 or p compared with 0.01
	Reject H_0 at 1% level. There is significant evidence (at the 1% level) to support Roma's suspicion.	E1dep	2.1a	Conclusion correct and in context dep test all correct.

Question	Scheme	Marks	AO	Notes
6(b)	First criticism is not valid…	E1	2.1b	
	…because sample size is large so the sample standard deviation can be used in place of σ	E1dep	3.1a	Dependent on previous E1 and reference to sample size and sample standard deviation.
	Second criticism is not valid…	E1	2.1b	
	…because a 5% significance level would make it **more likely (easier)** to reject H_0 not 'more difficult'.	E1dep	3.1a	Dependent on previous E1.
	Total	**10**		

Question	Scheme	Marks	AO	Notes
7(a)	$H_0 : \pi_A = \pi_B$ $H_1 : \pi_A \neq \pi_B$	B1	1.3	Condone use of p Hypotheses correct.
	$\hat{p} = \frac{86+98}{370} = \frac{92}{185}$	M1	1.3	$\hat{p}$ effort
	Test statistic $= \frac{\frac{86}{150} - \frac{98}{220}}{\sqrt{\left\{\frac{92}{185} \times \left(1 - \frac{92}{185}\right) \times \left(\frac{1}{220} + \frac{1}{150}\right)\right\}}}$	M1	1.3	Attempt at formula with $p_A - p_B$ correct
		M1	1.3	$\sqrt{\frac{1}{220} + \frac{1}{150}}$ used Denominator of correct form with.
		M1dep	1.3	'their' $\hat{p}$ used
	= 2.42 (3sf)	A1	1.3	awrt 2.42 (ignore sign)
	cv = ± 1.960 2.42 > 1.960 p = 0.0157 (0.015 ~ 0.016) oe	M1	2.1b	Allow −2.42 < −1.960 p = 0.0157 < 0.05
	Reject H_0 There is significant evidence (at the 5% level) that there is a difference between the proportion of patients suffering the side effect/dry throat for drug A and for drug B.	E1dep	2.1a	Conclusion correct and in context dep test all correct.

Question	Scheme	Marks	AO	Notes
(b)	Patients who have suffered the side effect are more likely to volunteer than those who haven't. **or** Patients may 'expect' the side effect because the consultant has spoken to them about it. **or** The proportion of those with the side effect in those patients taking drug B may indicate a reduction if the sample is drawn mainly from a population of patients who have suffered the side effect when taking drug A.	E1 E1	3.1b 3.1b	E1 for each relevant comment.
	Total	**10**		

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td>8(a)</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>
<table>
<tr><th>Observed frequencies</th><th>Top Field</th><th>Mid Field</th><th>Bottom Field</th><th>Total</th></tr>
<tr><td>Excellent</td><td>36</td><td>18</td><td>30</td><td>84</td></tr>
<tr><td>Good</td><td>36</td><td>50</td><td>42</td><td>128</td></tr>
<tr><td>Satisfactory</td><td>21</td><td>31</td><td>36</td><td>88</td></tr>
<tr><td>Total</td><td>93</td><td>99</td><td>108</td><td>300</td></tr>
</table>
</td><td>M1</td><td>1.2</td><td>Data considered as a contingency table with correct totals for rows and columns.</td></tr>
<tr><td></td><td>
<table>
<tr><th>Expected frequencies</th><th>Top Field</th><th>Mid Field</th><th>Bottom Field</th><th>Total</th></tr>
<tr><td>Excellent</td><td>26.04</td><td>27.72</td><td>30.24</td><td>84</td></tr>
<tr><td>Good</td><td>39.68</td><td>42.24</td><td>46.08</td><td>128</td></tr>
<tr><td>Satisfactory</td><td>27.28</td><td>29.04</td><td>31.68</td><td>88</td></tr>
<tr><td>Total</td><td>93</td><td>99</td><td>108</td><td>300</td></tr>
</table>
</td><td>M1</td><td>1.2</td><td>Method for Exp Freqs.
(row)×(column)/ grand
Possibly implied by correct result.</td></tr>
<tr><td></td><td></td><td>A1</td><td>1.2</td><td>All correct to 1dp</td></tr>
<tr><td></td><td>H_0: no association between apple quality and field.
H_1: Association between apple quality and Field.</td><td>B1</td><td>1.3</td><td>Both hypotheses correct.</td></tr>
<tr><td></td><td>Individual contributions to χ^2:
<table>
<tr><th></th><th>Top</th><th>Middle</th><th>Bottom</th></tr>
<tr><td>Excellent</td><td>3.810</td><td>3.408</td><td>0.002</td></tr>
<tr><td>Good</td><td>0.341</td><td>1.426</td><td>0.361</td></tr>
<tr><td>Satisfactory</td><td>1.446</td><td>0.132</td><td>0.589</td></tr>
</table>
</td><td>M1</td><td>1.3</td><td>$\frac{(O-E)^2}{E}$
Attempted.</td></tr>
<tr><td></td><td>Test statistic = $\frac{(36-26.04)^2}{26.0}+.....+\frac{(36-31.68)^2}{31.68}$
$= 3.810 + 3.408 + 0.002 + 0.341 + 1.426 + 0.361 + 1.446 + 0.132 + 0.589$
$= 11.515$</td><td>M1</td><td>1.3</td><td>Effort to sum.</td></tr>
<tr><td></td><td>$\chi^2 = 11.5$ (3sf)</td><td>A1</td><td>1.3</td><td>awfw11.50-11.60</td></tr>
</table>

Question	Scheme	Marks	AO	Notes
8(a) continued	df = 4	B1	1.3	
	cv at 5% level = 9.488 **or** $p = 0.021(3)$	M1	2.1b	cao 11.515 > 9.488 or $p = 0.021 < 0.05$
	Reject H_0 at 5% level	A1	2.1b	
	There is significant evidence (at the 5% level) of an association between the quality of apples and the field they were grown in.	E1dep	2.1a	In context; dep all previous marks but condone missing H_0
8(b)	Most notable difference between fields is that the Top field has more Excellent quality apples than expected while the Middle field has fewer.	E1	2.1a	Correct identification of Top and Middle fields.
	Given that the field with mistletoe will have a reduced yield of apples: **Either** Recommend Middle Field as lower yield of worse quality apples seems sensible; **or** Recommend either of the other two fields if Richard wants to maintain a reasonable spread of apple quality across all his fields. **or** Recommend Top Field since it has the smallest yield of apples.	E1	2.1b	Recommendation with sensible reason.
	Comment on fact that the data is only from 1 day, so the conclusion may not be reliable. Could sample apples over more days	E1	3.1a	Comment on only 1 day.
	Examples of future investigations: Is quality of apple affected by the presence of mistletoe? Is quality of mistletoe affected by the field used? (Would need some in each field) Is yield of mistletoe affected by the field used? (Again, need some in each field) Investigate financial balance between possibly less income from apples with income from selling mistletoe at the annual Mistletoe Fair.	E1 E1	3.1a 3.1a	Any two sensible suggestions for further investigation .
	Total	**16**		

Write your name here

Surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number | Candidate Number

Statistics

Advanced
Paper 3: Statistics in Practice

Sample Assessment Material for first teaching September 2017

Time: 2 hours

Paper Reference **9ST0/03**

You must have:
Statistical Formulae and Tables, calculator
Insert: Question 6 – Figures 4, 5 and 6

Total Marks

Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for algebraic manipulation, differentiation and integration, or 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.
- Inexact answers should be given to three significant figures unless otherwise stated.

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

S58393A

©2017 Pearson Education Ltd.

1/1/

S 5 8 3 9 3 A 0 1 2 4

Pearson

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

1 As part of an investigation into the effect of protein on weight gain, twenty female rats were to be observed from the ages of four weeks to seven weeks.

Ten of them were given a high protein diet, while ten were given a low protein diet. For reasons not connected to diet, 5 rats were lost to the study leaving 6 in the high protein group and 9 in the low protein group.

The table below gives the means and the standard deviations of the observed weight gains (in grams).

	High protein diet	**Low protein diet**
$\bar{x}$	12.1	10.3
s	2.02	1.95
n	6	9

A t-test for comparing the mean weight gains for this data is carried out and found to have a *p*-value of 0.11

(a) Calculate the value of Cohen's *d* for the effect of a high protein diet compared to a low protein diet on the weight gain of such rats. Assume that the variances of weight gain for the two diets are equal.

(3)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

(b) Describe what the information provided by the Cohen's *d* value evaluated in part (a), together with the given *p*-value, reveal about the possible effect of protein on weight gain in such rats.

You should include a comment explaining the importance of considering the effect size together with the *p*-value when carrying out this t-test.

(3)

(Total for Question 1 is 6 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

2 The diagram, Figure 1, illustrates results for the PISA (Programme for International Assessment) reading test taken in 2009 for some OECD (Organisation for Economic Co-operation and Development) countries.

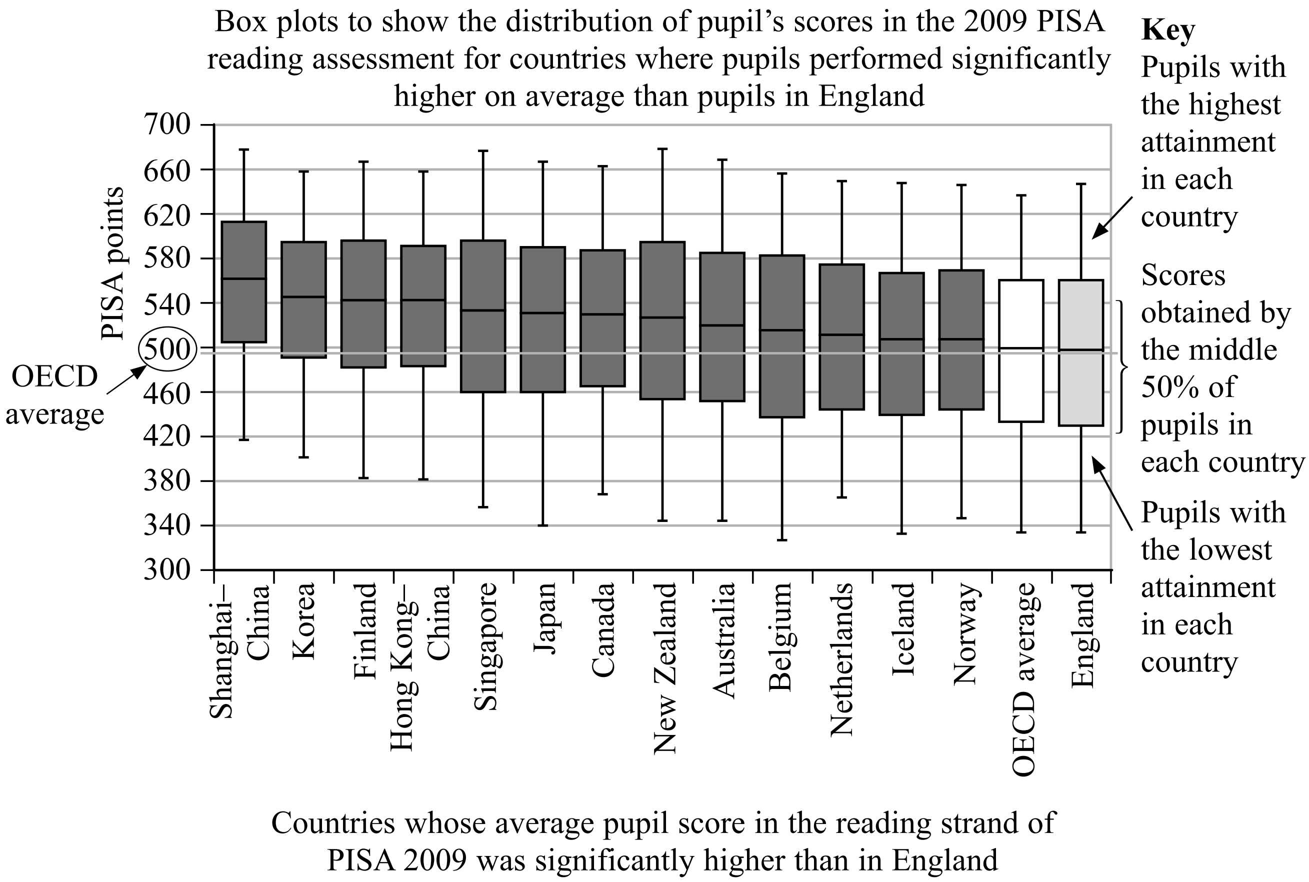


Figure 1

Discuss the distributions of PISA scores illustrated in Figure 1.
Make three different statements.

(3)

(Total for Question 2 is 3 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

3 Company trainees are required to take a series of selection tests that are designed to assess the core skills required in different parts of the company. Each test results in a score from 0 to 50.

The table below contains the scores for a random sample of 20 trainees on two of these selection tests.

Trainee

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Test A	33	25	17	31	23	29	14	37	35	33	30	21	32	43	24	37	33	37	28	37
Test B	35	36	37	20	28	24	36	19	19	16	28	27	23	10	23	14	21	15	24	14

Figure 1 contains output from part of a spreadsheet data analysis of these scores.

	Test A	**Test B**
Mean	29.95	23.45
Variance	54.8921	64.7868
Observations	20	20
Pearson Correlation	–0.8133	
Hypothesized Mean Difference	0	
df	19	
t Stat	1.9748	
P(T< = t) one-tail	0.0315	
t Critical one-tail	1.72913	
P(T< = t) two-tail	0.06301	
t Critical two-tail	2.09302	

Figure 2

(a) Using the information in Figure 2, state whether there is evidence, at the 5% level of significance, of an association between scores in test A and test B. Justify your answer.

(2)

(b) Using the information in Figure 2, state whether there is evidence at the 5% level of significance that, on average, scores in test A are higher than scores in test B. Justify your answer.

(3)

(c) New trainees have taken test A. Describe in simple terms to these new trainees how they may typically be expected to perform in test B relative to their performance in test A.

(3)

(Total for Question 3 is 8 marks)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

4 Large quantities of loose sand are delivered to a builders' merchant. The builders' merchant then uses a machine that automatically fills sand into 50 kilogram bags that are then sold to the general public.

Past data suggests that the mass of sand in a full bag may be modelled by a normal distribution with a standard deviation of 1.25 kg.

Over-filling and under-filling of bags could be problematic therefore it is important to maintain the target mean mass of 50 kg per bag.

Pasha is responsible for quality control at the builders' merchant and wants to check that their filling machine is operating properly. Consequently, she selects a random sample of 20 filled bags in order to investigate, at the 5% level of significance, whether the mean mass of sand in a bag has changed from 50 kg.

(a) State the name of the hypothesis test that you would carry out for this investigation.

(1)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

(b) (i) Determine the probability of a Type II error for a test of the hypothesis $H_0 : \mu = 50$ against $H_1 : \mu \neq 50$ at the 5% level of significance, based upon a random sample of 20 bags, in the case when, in fact, $\mu = 50.5$

(4)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(ii) Pasha has evaluated the probability of Type II error when $\mu = 50.5$ correctly and thinks that this probability is too high.

For future such checks on the mean of 20 bags, suggest a value for the significance level that should be used so that the probabilities of committing Type I and Type II errors are more acceptable to Pasha.

Justify your choice of significance level.

(2)

(Total for Question 4 is 7 marks)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

5 Ada, a researcher, is conducting an initial investigation into the effect of consuming alcohol on mental agility. She obtained 12 volunteers and divided them randomly into three groups of four.

One hour before being asked to solve a logic puzzle, the volunteers in Group 2 each consumed one measure of an alcoholic drink and those in Group 3 each consumed two measures of the alcoholic drink. The volunteers in Group 1 consumed no alcohol.

The time, x seconds, that it took each volunteer to solve the puzzle was recorded and the summarised results are given below.

	Group 1 (no alcohol)	**Group 2 (1 measure)**	**Group 3 (2 measures)**
	164	196	262
	126	188	198
	108	222	240
	174	144	190
Total	**572**	**750**	**890**

$$\sum\sum x^2 = 430040$$

You may assume that times taken to solve the logic puzzle may be modelled by a normal distribution.

(a) Carry out a hypothesis test to investigate for any difference between the mean times taken to solve the puzzle for the different amounts of alcohol consumed. Use the 5% level of significance.

Interpret fully your conclusion with advice, if appropriate, regarding the consuming of alcoholic drinks before attempting a logic puzzle.

You should also include a recommendation to Ada for the next stage of her investigation.

(12)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(b) Ada's assistant, Mehid, suggested that the time taken to solve a logic puzzle in such an investigation may also depend on the weight of the volunteer. To investigate this, he obtained twelve different volunteers who were ranked so that V1 represented the volunteer with lowest weight, up to V12 that represented the volunteer with highest weight.

They were assigned to four groups that were required to consume different measures of alcoholic drink, as shown in Figure 3. The times taken to complete the puzzle were then recorded for each volunteer.

Group 1 (no alcohol)	**Group 2 (1 measure)**	**Group 3 (2 measures)**	**Group 4 (no alcohol)**
V_1	V_4	V_7	V_{10}
V_2	V_5	V_8	V_{11}
V_3	V_6	V_9	V_{12}

Figure 3

(i) Identify two disadvantages of Mehid's suggested experimental design for an investigation of the effect of drinking alcohol on mental agility when weight is also taken into account.

(2)

(ii) Suggest a preferable experimental design for allocating Mehid's volunteers to 'no alcohol', 'one measure of an alcoholic drink' or 'two measures of an alcoholic drink' when weight rank is taken into consideration.

(2)

(Total for Question 5 is 16 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

You will need the extra document containing Figures 4, 5 and 6 with this question

6 Bryan is a senior manager at a large automotive company specialising in roadside assistance. He is interested in analysing data on breakdown incidents in the Greater London area.

Figure 4 is part of a spreadsheet with selected fields from the dataset *London Breakdown Records 2013-16*. It also contains the first 15 incidents (out of a total of 403663 incidents).

(a) Bryan wants to know what proportion of the total incidents involve a van.

Explain how you would use a spreadsheet program to calculate this proportion.

(2)

(b) Bryan is particularly interested in motorbike breakdowns where the only person at the vehicle is a teenage male. He wants to compare the response times of these incidents in different postcode districts.

Explain how you would use a spreadsheet program to obtain a list of '*Motorbike*' breakdowns, with '*One adult [teenage male]*', in postcode district '*SW1V*'.

(2)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

You will need the extra document containing Figures 4, 5 and 6 with this question

Figure 5 shows the twelve incidents in SW1V that satisfy the conditions in part (b).

Figure 6 shows the same data for another postcode district, SW1P.

The final column in **Figures 5** and **6**, '*Attendance time*', gives the time (in seconds) for the breakdown vehicle to arrive after the customer call is completed. When breakdown crews arrive at their destination, they record their arrival using an on-board computer.

(c) Bryan suspects that customers of this type who break down in SW1P are, on average, attended to more quickly than customers who break down in SW1V.

Bryan decides that incident IN00751414 in **Figure 5** is rather extreme and should be treated as an outlier and not used in any analyses of the data given in **Figure 5** and **Figure 6**.

(i) Give two possible explanations, in context, for this outlier.

(2)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

You will need the extra document containing Figures 4, 5 and 6 with this question

(ii) Justify Bryan's decision to remove the data on this incident before conducting any analyses of the data in given in **Figures 5** and **6**.

(2)

(iii) The attendance times in these two districts, excluding incident IN00751414 may be assumed to be random samples of times from independent normal distributions with equal variances.

Investigate Bryan's suspicion at the 5% level of significance.

(9)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 6 is 17 marks)

7 A vehicle emissions testing company wants to investigate whether its three testing machines, T_1, T_2 and T_3, are consistent with each other when testing a vehicle's nitrogen oxides (NO_X) emission level.

The lead engineer, Ugne, chooses to test three different makes of similar-sized hatchbacks.

She decides that she will take 15 cars of each make, selected randomly, straight from the factories. She suggests using a randomised block design for the investigation.

Ollie, another engineer, instead suggests that each car should be randomly assigned to a testing machine for the investigation.

(a) (i) State the name of Ollie's suggested experimental design.

(1)

(ii) Explain why Ugne's suggestion is preferable to Ollie's suggestion.

(1)

(iii) Suggest, with a reason, how Ugne's experimental design to test for the consistency of NO_X measurements by the three machines could be improved.

(2)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

(b) Using Ugne's randomised block design, the cars are tested on T_1, T_2, and T_3 and the levels of NO_X , X milligrams per kilometre, recorded.

The recorded data from all 45 tests are summarised in Figure 7.

Range of values	Frequency
$X < 18$	6
$18 \leqslant X < 19$	4
$19 \leqslant X < 20$	12
$20 \leqslant X < 21$	5
$21 \leqslant X < 22$	10
$22 \leqslant X < 23$	5
$X > 23$	3

Figure 7

Summary statistics obtained from the recorded data gave $\bar{x} = 20.2$ $s^2 = 4.09$

Test whether X can be modelled by a normal distribution with the same mean and the same standard deviation as the recorded data.

(11)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(c) The three testing machines are old model machines. Ugne has become aware that old machines may not be detecting all of a vehicle's NO_X emissions.

The manufacturer of a new model machine claims that it can detect more of a car's NO_X emissions

Ugne decides to randomly select 15 of the cars already tested on the old model machines, 5 of each make, and test their NO_X emissions using the new model machine.

She calculates the differences between the measurements on the old model and new model machines (new – old) and these differences are shown in Figure 8.

You may assume that these differences may be modelled by a normal distribution.

Car	**Difference (New model – old model)**
C_1	3.5
C_2	1.8
C_3	–2.1
C_4	–0.7
C_5	1.1
C_6	–0.2
C_7	–1.1
C_8	–0.5
C_9	1.2
C_{10}	2.7
C_{11}	–2.8
C_{12}	0.3
C_{13}	2.8
C_{14}	4.3
C_{15}	1.4

Figure 8

Investigate whether the data supports the manufacturer's claim.

(6)

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

(d) Ollie suspects that the car make may also be affecting the data.

Assuming that you had further time and resources, explain how you would change the experimental design and analysis used by the company in order to also investigate the effect that the make of a car has on the data.

(2)

(Total for Question 7 is 23 marks)

TOTAL FOR PAPER IS 80 MARKS

DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA DO NOT WRITE IN THIS AREA

Pearson Edexcel Level 3 GCE

Statistics

Advanced
Paper 3: Statistics in Practice

Sample Assessment Material for first teaching September 2017 **Question 6 – Figures 4, 5 and 6**	Paper Reference **9ST0/03**

Do not return the insert with the question paper.

Turn over ▶

S58393A

©2017 Pearson Education Ltd.

1/

S 5 8 3 9 3 A

Pearson

Question 6

Figure 4 – London Breakdown Records 2013–16

Incident_Number	Date_Of_Call	Vehicle_Category	Failure_Group	People_At_Vehicle	Postcode_District	Attendance_Time
IN00000113	01/01/2013	Car	Battery	One adult [female]	NW6	1515
IN00000413	01/01/2013	Car	Engine cooling system	One adult [male]	UB6	2142
IN00000513	01/01/2013	Car	Engine cooling system	Two adults [female/female]	W6	1978
IN00000213	01/01/2013	Motorbike	Wheels/tyres	One adult [male]	W1H	3866
IN00000313	01/01/2013	Car	Engine cooling system	Three or more adults [no males]	EC3R	3140
IN00000613	01/01/2013	Motorbike	Bodywork	Two adults [male/female]	NW10	2931
IN00000713	01/01/2013	Car	Engine cooling system	One adult [male] & child(ren)	NW6	4065
IN00001013	01/01/2013	Car	Electrical systems	One adult [female] & child(ren)	SW17	1361
IN00001313	01/01/2013	Motorbike	Electrical systems	One adult [female]	E7	2015
IN00001513	01/01/2013	Car	Battery	Two adults [male/female]	UB3	2486
IN00001713	01/01/2013	Motorbike	Engine cooling system	Two adults [male/male]	N22	2323
IN00001913	01/01/2013	Car	Electrical systems	One adult [teenage male]	SE1	2741
IN00002213	01/01/2013	Van	Electrical systems	Two adults [male/female] & child(ren)	N16	NULL
IN00002413	01/01/2013	Van	Electrical systems	Two adults [male/female] & child(ren)	SE1	NULL
IN00002513	01/01/2013	Van	Electrical systems	One adult [teenage male]	SE17	4510

Based on genuine data

Figure 5 – London Breakdown Records 2013–16

Incident_Number	Date_Of_ Call	Vehicle _ Category	Failure_Group	People_At_Vehicle	Postcode District	Attendance_ Time
IN00314813	08/01/2013	Motorbike	Electrical systems	One adult [teenage male]	SW1V	1824
IN00840113	20/01/2013	Motorbike	Electrical systems	One adult [teenage male]	SW1V	1806
IN00966413	23/01/2013	Motorbike	Electrical systems	One adult [teenage male]	SW1V	1488
IN00633114	16/01/2014	Motorbike	Electrical systems	One adult [teenage male]	SW1V	2804
IN00751414	19/01/2014	Motorbike	Electrical systems	One adult [teenage male]	SW1V	8231
IN00793914	20/01/2014	Motorbike	Electrical systems	One adult [teenage male]	SW1V	2169
IN03112014	14/03/2014	Motorbike	Electrical systems	One adult [teenage male]	SW1V	2831
IN15744114	23/11/2014	Motorbike	Electrical systems	One adult [teenage male]	SW1V	2895
IN06348715	23/05/2015	Motorbike	Electrical systems	One adult [teenage male]	SW1V	2995
IN17693515	26/12/2015	Motorbike	Electrical systems	One adult [teenage male]	SW1V	2568
IN11766716	01/09/2016	Motorbike	Electrical systems	One adult [teenage male]	SW1V	2232
IN12931016	23/09/2016	Motorbike	Electrical systems	One adult [teenage male]	SW1V	2750

Based on genuine data

Figure 6 – London Breakdown Records 2013–16

Incident_Number	Date_of_ Call	Vehicle_ Category	Failure_Group	People_at_Vehicle	Postcode_ District	Attendance_ Time
IN05213913	26/04/2013	Motorbike	Electrical systems	One adult [teenage male]	SW1P	1334
IN10978213	10/08/2013	Motorbike	Electrical systems	One adult [teenage male]	SW1P	1597
IN15482213	08/11/2013	Motorbike	Electrical systems	One adult [teenage male]	SW1P	1125
IN16183413	23/11/2013	Motorbike	Electrical systems	One adult [teenage male]	SW1P	2577
IN00224615	06/01/2015	Motorbike	Electrical systems	One adult [teenage male]	SW1P	1543
IN02552716	02/03/2016	Motorbike	Electrical systems	One adult [teenage male]	SW1P	1579
IN10485116	09/08/2016	Motorbike	Electrical systems	One adult [teenage male]	SW1P	2949

Based on genuine data

Paper 3: Statistics in Practice Mark Scheme

Question	Scheme	Marks	AO	Notes
1(a)	Pooled $s^2 = \frac{5\times 2.02^2 + 8\times 1.95^2}{13}$	M1	1.2	Correct formula used (possibly implied).
	$= 3.90(938)$	A1	1.2	**or** s =1.9772 (1.97 ~ 1.98)
	Then Cohen's $d = \frac{12.1-10.3}{1.9772}$ $= 0.910$	A1ft	1.2	awrt 0.91 (ft their s^2)
1(b)	The p-value, > 0.10, suggests there is no significant evidence (at the 10% level) of a difference in mean weight gain for rats on a low or high protein diet …	E1	2.1a	oe; interpretation of p-value with strength described as something less than 'some' (**or** not significant at α = 0.10,0.05,0.01).
	…whereas the Cohen's d value suggests the effect (or the difference between means, which equals 12.1 – 10.3 = 1.8) is 'large' thus indicating that a high protein diet resulted in a higher mean weight gain.	E1	2.1a	oe; interpretation of d (based on guidance that $d > 0.8$ is 'large').
	Considering only the p-value in this investigation would lead to the large 'effect' on weight gain of the high protein diet being missed. **or** The small sample sizes influenced the size of the t-statistic – investigations should continue with a larger sample size.	E1	3.1b	Could lead to the large effect of the high protein diet being overlooked because the p-value was large. (small sample sizes influenced the size of the t-statistic). Sample size was very small.
	Total	**6**		

Question	Scheme	Marks	AO	Notes
2	England average score is about same/just below the average (500) for OECD countries. The highest median score is about 60 higher than that for England. Shanghai has the highest median score. The spread/range of scores for England is slightly greater than the average OECD spread. Shanghai-China, Korea, Finland, Hong Kong-China, Canada, Netherlands and Norway have a narrower spread/range of scores than England (or than the OECD average). New Zealand has the widest spread/range of scores. Singapore, Japan, Australia and Iceland have a similar spread/range of scores to England (or the OECD average). Shanghai- China, Singapore and New Zealand have the highest scoring pupil(s). Netherlands, Iceland and Norway have similar levels of high scoring pupil(s) (about 10 above the OECD average). Belgium has the lowest scoring pupil(s).	E1 E1 E1	2.1a 2.1a 2.1a	For three comments regarding distributions.
	Total	**3**		

Question	Scheme	Marks	AO	Notes
3(a)	5% critical value for r = ± 0.4438	B1	1.3	For (±) 0.4438
	Since r = −0.8133 < −0.443 (oe) there is evidence of an association between the scores for tests A and B.	E1	2.1a	For concluding there is an association from a correct comparison.
3(b)	Comparing one-tailed *p*-value (= 0.0315) with 0.05	B1	1.3	Use of 0.05 **or** 1.729
	or ts = 1.97 with one-tailed critical value (=1.729)	M1	2.1b	Correct comparison.
	and $\bar{x}_A > \bar{x}_B$	M1dep	2.1b	Dep previous M1. For considering direction of difference.
	So there is significant evidence (at the 5% level) that on average, scores on test A are higher than those on test B.	E1dep	2.1a	Dep first M1 only.
3(c)	On average trainees will score lower on test B.	E1	2.1a	Comparison of mean scores.
	However relative to other trainees, if (s)he had relatively high/low score in test A then can expect a relatively low/high score in test B.	E1	2.1a	Correct interpretation of negative correlation in context.
	Total	**8**		

Question	Scheme	Marks	AO	Notes
4(a)	A one-sample z-test on a population mean.	B1	2.1b	Just z-test needed. "One-sample" and "mean" can be implied by ts.
4(b)(i)	Type II error P(Accept H_0|H_0 false)			
	x values corresponding to z critical values are 49.4522, 50.5478 (may appear in part (a)).	M1	1.3	Using/finding critical values for x
	Use of $\mu = 50.5$	M1	1.3	Possibly implied.
	$P(49.4522 < \bar{x} \leq 50.5478 \mid \mu = 50.5)$	M1dep	1.3	Probability of acceptance effort.
	$= 0.568$	A1	1.3	awrt 0.57
4(b)(ii)	Recognise that if P(Type II error) is to decrease then P(Type I error) must increase (assuming n is unchanged).	E1	3.1a	Demonstrate awareness of trade-off between types of error.
	Suggest sig level $0.05 < \alpha < 0.15$	B1	2.1b	α not too high.
	Total	**7**		

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td>5(a)</td><td>H0: $\mu_1 = \mu_2 = \mu_3$

H1: at least 2 of the means differ.
5% 1 tail
<table>
<tr><th>Grp 1</th><th>Grp2</th><th>Grp3</th></tr>
<tr><td>164</td><td>196</td><td>262</td></tr>
<tr><td>126</td><td>188</td><td>198</td></tr>
<tr><td>108</td><td>222</td><td>240</td></tr>
<tr><td>174</td><td>144</td><td>190</td></tr>
</table>

$T_1 = 572$ $T_2 = 750$ $T_3 = 890$
$n_1 = 4$ $n_2 = 4$ $n_3 = 4$

$T = 2212$ $\sum\sum x^2 = 430040$ $N = 12$</td><td>B1</td><td>1.3</td><td>Hypotheses stated correctly.</td></tr>
<tr><td></td><td>Total ss = $430040 - \frac{2212^2}{12}$
= 22294.7</td><td>M1</td><td>1.3</td><td>ss Total (condone small slip).</td></tr>
<tr><td></td><td>Groups ss
$\frac{572^2}{4} + \frac{750^2}{4} + \frac{890^2}{4} - \frac{2212^2}{12}$
= 12700.7</td><td>M1</td><td>1.3</td><td>ss Between groups (condone small slip).</td></tr>
<tr><td></td><td>22294.7 – 12700.7 = 9594</td><td>M1dep</td><td>1.3</td><td>dep first two M's
ss Error
(allow small slip- not if negative).

Dep previous M's.</td></tr>
</table>

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td>5(a) continued</td><td>
<table>
<tr><th></th><th>ss</th><th>df</th><th>ms</th></tr>
<tr><td>Groups</td><td>12700.7</td><td>2</td><td>6350.5</td></tr>
<tr><td>Error</td><td>9594</td><td>9</td><td>1066</td></tr>
<tr><td>Total</td><td>22294.7</td><td>11</td><td></td></tr>
</table>
</td><td>B1

M1dep</td><td>1.3

1.3</td><td>df between groups and error – both.

ms Error ss and Between ss divided to get ms.
Dep previous M’s</td></tr>
<tr><td></td><td>$F = \frac{6350.5}{1066} = 5.96$</td><td>M1dep</td><td>1.3</td><td>F between groups divided by error dep previous M’s [5.9 ~ 6.0 or $p = 0.0225$].</td></tr>
<tr><td></td><td>$F_9^2 = 4.256$</td><td>B1</td><td>1.3</td><td></td></tr>
<tr><td></td><td>5.957 > 4.256

Reject H_0</td><td>M1dep</td><td>2.1b</td><td>or $p = 0.0225 < 0.05$
Correct comparison and conclusion (can be implied later in question) dep F_9^2 or 0.05 used.</td></tr>
<tr><td></td><td>There is significant evidence (at the 5% level) to suggest a difference between the mean time to solve the puzzle for at least two of the amounts of alcohol.

There is significant evidence (at the 5% level) to suggest that it takes considerably longer, on average, to complete the puzzle when 2 measures of an alcoholic are consumed one hour before starting than it does when no alcohol is consumed.</td><td>E1dep</td><td>2.1a</td><td>For either of these conclusions.
Dep whole test correct.</td></tr>
<tr><td></td><td>To complete a logic puzzle in the fastest time do not consume alcohol beforehand.</td><td>E1</td><td>2.1a</td><td></td></tr>
<tr><td></td><td>For further investigations Ada could consider:
• obtaining/using a larger sample
• the routine alcohol consumption of volunteers
• age/sex etc of volunteers</td><td>E1</td><td>3.1a</td><td>A valid suggestion for Ada.</td></tr>
</table>

Question	Scheme	Marks	AO	Notes
5(b)(i)	Alcohol allocation is not 'fairly' distributed amongst weight categories.	E1	3.1a	Allocation not 'fair'(possibly implied).
	All lightest and all heaviest are only tested when having consumed no alcohol. **or** 'Medium weight' people are the only ones to be tested after consuming alcohol.	E1	3.1a	Explanation in context.
5(b)(ii)	Rank volunteers (as Mehid did) and define 4 weight categories (lightest to heaviest) with 3 volunteers in each.	E1	3.1a	Each weight category to be represented in all alcohol groups.
	Randomly assign one person from each weight category to each of the 3 groups 'no alcohol', '1 measure of alcohol' and '2 measures of alcohol'.	E1	3.1a	Fully explained in context and random assignment mentioned.
	Total	**16**		

Question	Scheme	Marks	AO	Notes
6(a)	Enter a formula to COUNT the number of records with 'Van' in the 'Vehicle_Category' column.	E1	1.1	Must see 'count' oe COUNTIF Vehicle_Category = "Van" **or** 'filter' seen and used for selecting 'Van' from 'Vehicle_ Category' column. oe
	Count the **total** number of records, then divide the 'Van' count by the total count.	E1	1.1	Must see both.
6(b)	Filter…	E1	1.1	"Filter" seen.
	…'Vehicle_Category' to 'Motobike, 'People_At_Vehicle' to 'One adult [teenage male]', and 'Postcode_District' to 'SW1V'	E1	1.1	All three correct. Field names must be specified.
6(c)(i)	Possible reasons for extremely *large* value (i.e. slow attendance time) include: • Traffic problems. • Weather problems. • Breakdown of rescue vehicle. • Recording error. • Crew failed to register attendance on computer.	E1, E1	2.1a 2.1a	Any 2 sensible explanations in context.

Question	Scheme	Marks	AO	Notes
6(c)(ii)	Outliers can greatly affect the analysis and/or conclusion… **or** It may be from a different population of times to the other values in Figure 5…	E1	3.1b	Valid reason.
	…so, yes, it should be removed if it is likely to be an error of recording (but not just for convenience).	E1dep	3.1b	cao. Requires previous E1.
6(c)(iii)	$H_0 : \mu_V = \mu_P$ $H_1 : \mu_V > \mu_P$	B1	1.3	
	$\bar{x}_V = 2396.55$ $s_V = 519.28$	B1	1.2	Both correct to 3sf.
	$\bar{x}_P = 1814.86$ $s_p = 677.16$	B1	1.2	Both correct to 3sf.
	Pooled $s^2 = \dfrac{10 \times 519.28^2 + 6 \times 677.16^2}{16}$ $= 340\,486.9$	M1F	1.3	Correct formula used for pooled s^2
	Test stat $t = \dfrac{2396.55 - 1814.86}{\sqrt{340486.9 \times \left(\dfrac{1}{11} + \dfrac{1}{7}\right)}}$ $= 2.06$	M1F A1	1.3 1.2	Correct formula used for t answer 2.05-2.10
	df = 16, 1-tailed cv = 1.746	B1	1.3	**Alternatively**, comparing p-values where p = 0.0279 (0.025~0.030) oe (Correct p implies previous 6 marks).
	2.06 >1.746 Significant evidence (at the 5% level) to reject H_0	M1	2.1b	$p < 0.05$ Conclusion from correct comparison, their values.

Question	Scheme	Marks	AO	Notes
6(c)(iii) continued	There is significant evidence (at the 5% level) to support Bryan's suspicion.	E1dep	2.1a	oe in context. Dep whole test correct.
	Total	**17**		

Issue 1 – September 2017 © Pearson Education Limited 2017

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td>7(a)(i)</td><td>Completely randomised design.</td><td>B1</td><td>1.1</td><td></td></tr>
<tr><td>7(a)(ii)</td><td>Blocking minimises the chance of experimental error due to the nuisance factor, make of car.
If a difference between
or
the testing machines exists, then it is more likely to be detected if blocking is used.</td><td>E1</td><td>3.1a</td><td>oe</td></tr>
<tr><td>7(a)(iii)</td><td>Test every car on every machine.</td><td>E1</td><td>3.1a</td><td></td></tr>
<tr><td></td><td>This will minimise another nuisance factor, which car.</td><td>E1</td><td>3.1a</td><td>Any sensible explanation.</td></tr>
<tr><td>7(b)</td><td>Method for finding probabilities from a normal distribution with mean 20.2 and variance 4.09
The full set of probabilities are:
<table>
<tr><th>Range of values</th><th>Probability</th></tr>
<tr><td>$C < 18$</td><td>0.1383</td></tr>
<tr><td>$18 \leqslant C < 19$</td><td>0.1381</td></tr>
<tr><td>$19 \leqslant C < 20$</td><td>0.1841</td></tr>
<tr><td>$20 \leqslant C < 21$</td><td>0.1932</td></tr>
<tr><td>$21 \leqslant C < 22$</td><td>0.1595</td></tr>
<tr><td>$22 \leqslant C < 23$</td><td>0.1036</td></tr>
<tr><td>$C > 23$</td><td>0.0831</td></tr>
</table></td><td>M1

A1

A1</td><td>1.1

1.1

1.1</td><td>May not be seen if, as is expected, answers are found from the normal dist. functions on a calculator. May be implied by one correct probability.

Any two correct probabilities correct to 3sf.

All probabilities correct.</td></tr>
<tr><td></td><td>H_0 : X can be modelled by a normal dist.
H_1 : Normal dist cannot be used to model X</td><td></td><td></td><td>Do not have to be stated.</td></tr>
</table>

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td>7(b) continued</td><td>For using χ^2 test and sensible sig level (1, 5, 10%)</td><td>M1</td><td>2.1b</td><td>Selection of appropriate test and sig level.</td></tr>
<tr><td rowspan="2"></td><td rowspan="2"><table>
<tr><th>NO_X</th><th>O</th><th>E</th></tr>
<tr><td>X < 18</td><td>6</td><td>6.22</td></tr>
<tr><td>18 ≤ X < 19</td><td>4</td><td>6.21</td></tr>
<tr><td>19 ≤ X < 20</td><td>12</td><td>8.28</td></tr>
<tr><td>20 ≤ X < 21</td><td>5</td><td>8.69</td></tr>
<tr><td>21 ≤ X < 22</td><td>10</td><td>7.18</td></tr>
<tr><td>X ≥ 22</td><td>8</td><td>8.40</td></tr>
</table></td><td>M1</td><td>1.3</td><td>Correct expected.
ft their probs × 45
(Allow 1 dp accuracy for E's)</td></tr>
<tr><td>M1</td><td>1.3</td><td>Last two classes combined.</td></tr>
<tr><td></td><td>$\sum\frac{(O-E)^2}{E} = 5.16$ (3sf)</td><td>M1</td><td>1.3</td><td>ts $\sum\frac{(O-E)^2}{E}$ effort correct.</td></tr>
<tr><td></td><td></td><td>A1</td><td>1.3</td><td>5.1 ~ 5.2
sc no pooling
$\sum\frac{(O-E)^2}{E}=5.31$
M1 M1 only</td></tr>
<tr><td></td><td>df = 6 – 1 – 2 = 3</td><td>B1</td><td>1.3</td><td>cao</td></tr>
<tr><td></td><td>5% cv for χ_3^2 is 7.815 > 5.16</td><td>B1</td><td>1.3</td><td>Acceptable cv. Alternatively, comparing p-values where p = 0.16(0) ~ 0.165 > 0.05 (or 0.01 or 0.10).
Correct p implies previous M1A1.</td></tr>
<tr><td></td><td>No significant evidence (at the 5% level) that the normal distribution is not a suitable model for NO_X emissions</td><td>E1dep</td><td>2.1a</td><td>Correct conclusion in context
dep whole test correct.</td></tr>
</table>

Question	Scheme	Marks	AO	Notes
7(b) continued				Condone use of 1% or 10% significance level 1% cv = 9.210 **or** 10% cv = 6.251
7(c)	$H_0 : \mu_d = 0$ $H_1 : \mu_d > 0$	B1	1.3	Both correct.
	$\bar{d} = 0.78$ $s_d = 2.05155$	M1dep	1.3	Mean and standard deviation of differences found. Dep M1.
	$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} = \frac{0.78}{\frac{2.052}{\sqrt{15}}}$	M1dep	1.3	ts method correct. Dep previous M's.
	$= 1.47$	A1	1.3	awrt 1.47
	$cv = 1.761 > 1.47$	M1	2.1b	Correct comparison Alternatively, comparing p-values where $p = 0.0185 < 0.05$ oe Correct p implies previous 3 marks.
	Accept H_0 There is no sufficient evidence (at the 5% level) to support the manufacturer's claim.	E1dep	2.1a	Correct conclusion in context. All correct dep whole test correct
7(d)	Separate data by car make.	E1	3.1a	
	Use two-factor ANOVA	E1	3.1a	
	Total	**23**		

edexcel

For information about Edexcel, BTEC or LCCI qualifications
visit qualifications.pearson.com

Edexcel is a registered trademark of Pearson Education Limited

Pearson Education Limited. Registered in England and Wales No. 872828
Registered Office: 80 Strand, London WC2R 0RL
VAT Reg No GB 278 537121

ISBN 978-1-4469-4613-8