Pearson Edexcel

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

Summer 2023 Exemplar
9ST0-03 A level Mathematics
Paper 03: Statistics in Practice
Senior Examiner's feedback on student responses

Contents

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Skip to Main Contents

About this booklet

This document has been produced to support mathematics teachers delivering the GCE Mathematics specification.

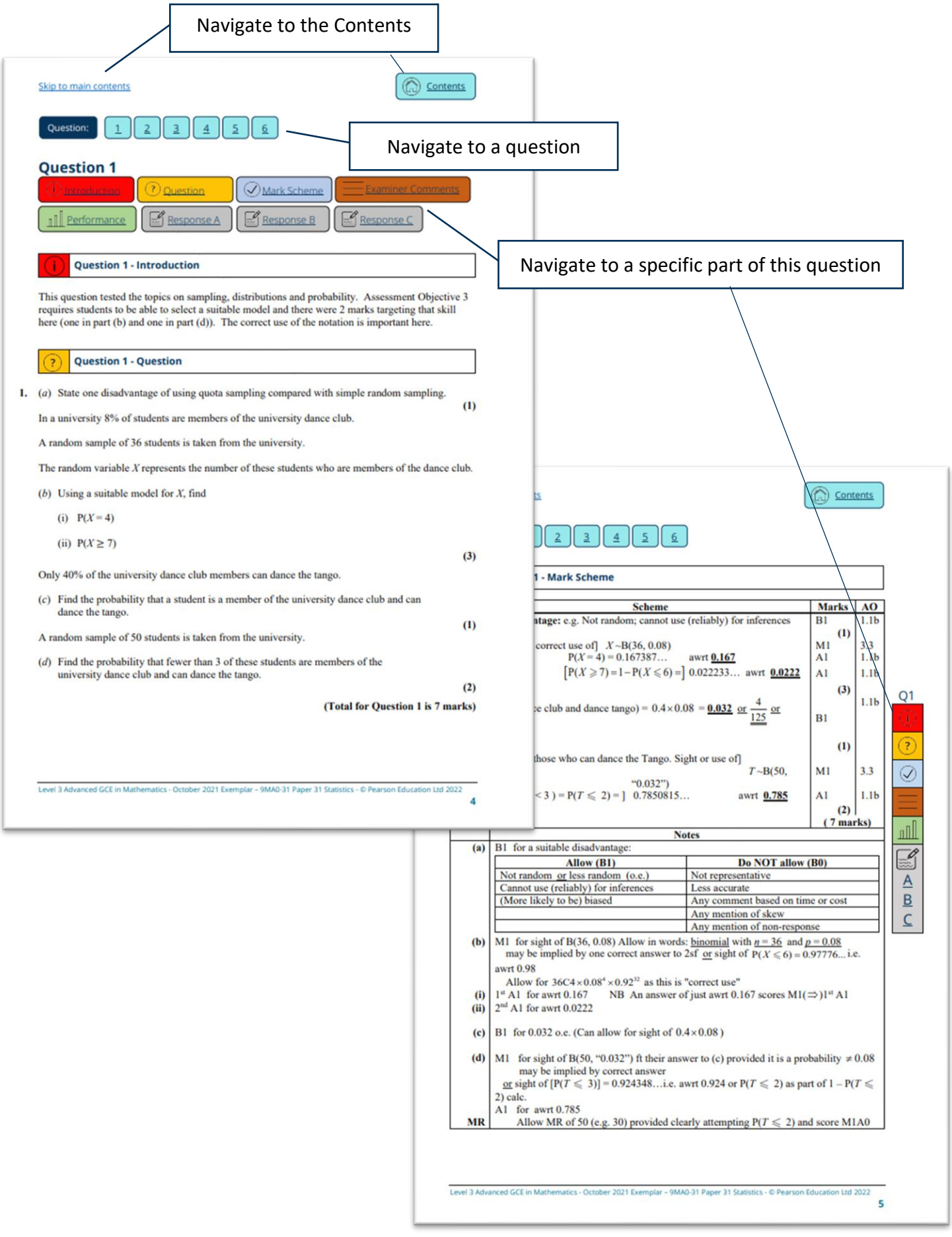
This document looks at questions from the 9ST0-03 A level Statistics Paper 3: Statistics in Practice Summer 2023 examination paper. It shows real student responses to these questions, and how the examining team follow the mark schemes to demonstrate how the students are awarded the marks. For the mark scheme notes and details of alternative methods please see the full mark scheme for this question paper on our website.

For 2023, the approach all exam boards have taken to grading was to return to pre-pandemic grading, while giving students protection against any impact of disruption. Results in summer 2023 therefore will be far more in line with summer exams that were sat in 2019, but lower than in 2022, when grades awarded were based on a mid-point between 2019 and 2021 outcomes. For more information please read our 'Understanding grade boundaries 2023' document.

* The question level performance data is there to give an indication only of how students performed, on each question, in the context of sitting the entire exam paper and is not an indication of how students may perform sitting a question in isolation.

Skip to Main Contents

How to use this booklet



General Examiner Feedback

In general candidates made good attempts at the calculations in this paper. A common area for improvement however is ensuring that explanations and assumptions for hypothesis tests and confidence intervals need to be given in the context of the question rather than simply being stated in general terms.

Question 1

Introduction | Question | Mark Scheme | Examiner Comments | Performance | Response A | Response B | Response C

Question 1 - Introduction

This question assesses content from the discrete random variable section of the specification as well as requiring candidates to know and use the central limit theorem.

Questions such as (a) or (c) will be given full marks for final answers alone, however candidates were given the answer to (b) and so had to communicate their working effectively in order to be awarded the mark.

Question 1 - Question

1 In a video game you can mine blocks of coal. The yield of a block, when mined, is determined randomly.

Mining a block of coal yields different numbers of pieces of coal, which follow the distribution shown in **Figure 1**.

The probabilities, p, of yielding 2, 3 or 4 pieces are equal.

Pieces of coal	1	2	3	4
Probability	0.4	p	p	p

[Source: https://minecraft.fandom.com/wiki/Fortune]

Figure 1

(a) Find the value of p

(1)

(b) Show the expected yield from mining one block of coal is 2.2 pieces of coal.

(1)

(c) Find the variance of the yield from mining one block of coal.

(1)

(d) Find the probability that the yield of a block is **one** piece of coal, given that the yield was an odd number of pieces of coal.

(2)

(e) State the central limit theorem.

(3)

(f) Estimate the probability that, when 60 blocks of coal are mined, the average yield per block is more than 2.25

(2)

(Total for Question 1 is 10 marks)

Question 1 - Mark Scheme

Question	Scheme	Marks	AO	Notes
1(a)	p=0.2	B1	1.1	
1(b)	$1 \times 0.4 + 2 \times 0.2 + 3 \times 0.2 + 4 \times 0.2 = 2.2$	B1*	1.1	AG
1(c)	$var(X) = 1.36$	B1	1.1	
1(d)	$\frac{0.4}{0.2 + 0.4}$ $\frac{2}{3}$	M1 A1	1.2 1.2	
1(e)	(Regardless of the distribution of X,) when n is large… the distribution of $\bar{X}$ is approximately normal… with mean equal to mean of X and variance = $\frac{\text{var}(X)}{n}$	E1 E1 E1	1.3 1.3 1.3	n≥ 30
1(f)	$\bar{X} \sim N\left(2.2, \frac{1.36}{60}\right)$ $P(\bar{X} > 2.25) = 0.3699$	B1 A1	1.2 1.2	Seen or used awrt 0.37
	Total	**10**		

Question 1 - Examiner Comments

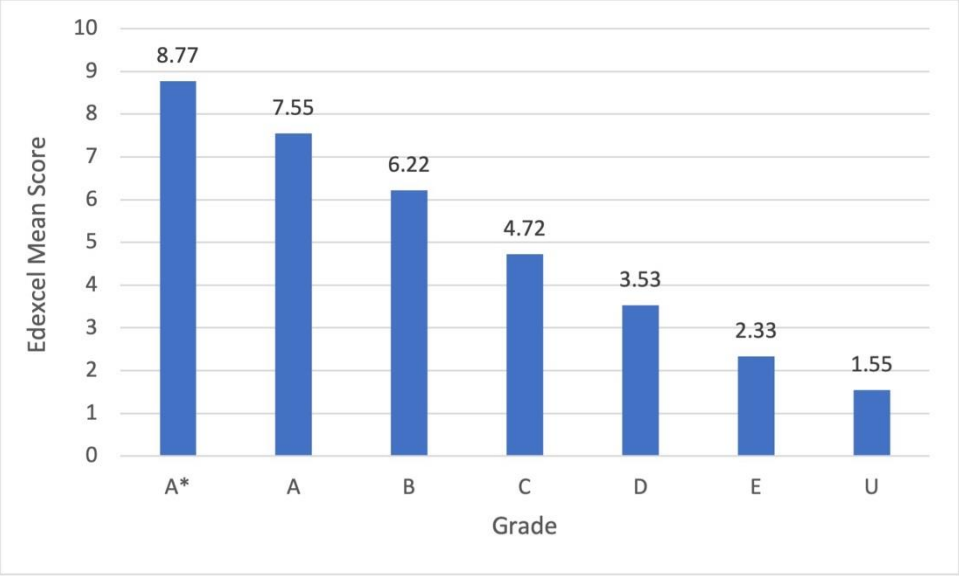
Q1(a) to (d) of this question were generally well answered however many candidates were unable to accurately **show** how to calculate the mean of a discrete random variable.

Q1(e) Very few candidates gained full marks in this part of the question with the majority only gaining the mark for knowing that the central limit theorem was for use with large samples. Very few candidates could state that the distribution of sample means was normally distributed and fewer still stated the variance of the distribution of sample means.

Q1(f) Many candidates were able to make a good attempt at this question although there were some who did not recognise the significance of finding the probability of an average value.

Question 1 - Performance

Mean score	Max score	Mean %	Edexcel averages: mean scored by candidates achieving grade:							
			ALL	A*	A	B	C	D	E	U
4.54	10	45	4.54	8.77	7.55	6.22	4.72	3.53	2.33	1.55



Question 1 - Response A

1 In a video game you can mine blocks of coal. The yield of a block, when mined, is determined randomly.

Mining a block of coal yields different numbers of pieces of coal, which follow the distribution shown in **Figure 1**.

The probabilities, p, of yielding 2, 3 or 4 pieces are equal.

Pieces of coal	1	2	3	4
Probability	0.4	p	p	p

[Source: https://minecraft.fandom.com/wiki/Fortune]

Figure 1

(a) Find the value of p

(1)

$1 - 0.4 = 0.6 \quad \frac{0.6}{3} = 0.2$

(b) Show the expected yield from mining one block of coal is 2.2 pieces of coal.

(1)

$1 \times 0.4 = 0.4$
$2 \times 0.2 = 0.4$
$3 \times 0.2 = 0.6$
$4 \times 0.2 = 0.8$
2.2

(c) Find the variance of the yield from mining one block of coal.

(1)

~~$Var = 1 \times (1-p)$~~
~~$Var = 1 \times (1-2.2) = 2.2$~~

$Var = np(1-p)$
~~$Var = 4 \times 2.2(1-2.2)$~~
$Var = 4 \times 0.55(1-0.55) = 0.99$

Q1

Question 1 continued

(d) Find the probability that the yield of a block is **one** piece of coal, given that the yield was an odd number of pieces of coal.

(2)

$X \sim B(2, 0.6)$

$P(X=1) = 0.48$

(e) State the central limit theorem.

(3)

The central limit theorem has to be stated as the data is insufficiently large as it is less than 30 variables

The central limit theorem is used when the data is insufficiently large which would be less than 30 so s_x will replace σ_x

(f) Estimate the probability that, when 60 blocks of coal are mined, the average yield per block is more than 2.25

(2)

$X \sim N(2.2, \frac{0.99}{60})$

$P(x > 2.25) = 0.348$ (3SF)

3 /10

Q1

A B C

Part (a)
B1: The candidate has correctly obtained the answer 0.2

Part (b)
B1: Although no "+" is present in the candidate's solution, the intent was considered clear enough to award the mark.

Part (c)
B0: This part of the question was not awarded the mark as 0.99 is not the variance. The student has confused the formula for the variance of a binomial distribution with the variance of a discrete random variable.

Part (d)

M0A0: There is no evidence of conditional probability or Bayes' theorem being used here and the answer is incorrect, so no marks were awarded.

Part (e)

E0E0E0: The first mark is not awarded here as the candidate has stated "insufficiently" large which is incorrect. There is no reference to the distribution of sample means or the standard error.

Part (f)

B1A0: The first mark has been awarded here as the candidate has clearly tried to use the correct distribution but using their variance from (c). Since the answer is not 0.37 however, they cannot be awarded the A mark.

Question 1 - Response B

1 In a video game you can mine blocks of coal. The yield of a block, when mined, is determined randomly.

Mining a block of coal yields different numbers of pieces of coal, which follow the distribution shown in **Figure 1**.

The probabilities, p, of yielding 2, 3 or 4 pieces are equal.

Pieces of coal	1	2	3	4
Probability	0.4	p	p	p

[Source: https://minecraft.fandom.com/wiki/Fortune]

Figure 1

(a) Find the value of p **(1)**

$1 - 0.4 = 0.6 \div 3 = 0.2$

(b) Show the expected yield from mining one block of coal is 2.2 pieces of coal. **(1)**

Expected yield = Average yield
Use ~~Stat~~ Menu 6: Statistics
1: 1-Variable.

$E(X) = \mu = 2.2$

(c) Find the variance of the yield from mining one block of coal. **(1)**

$Var(X) = \sigma^2 = 1.36$

(d) Find the probability that the yield of a block is **one** piece of coal, given that the yield was an odd number of pieces of coal. **(2)**

$0.4 + 0.2 = 0.6$ $P(X = 1 \text{ Odd}) = 0.6$

$P(X = 1 / \text{Odd}) = \frac{0.4}{0.6} = 0.667$

Q1

A B C

(e) State the central limit theorem.

The Central Limit Theorem states that sample means are approximately normally distributed, even if the underlying population is not, as long as the sample is large. (3)

6 /10

Q1

Part (a)

B1: No working is required here however it can be seen that the candidate has arrived at the correct answer with correct working.

Part (b)

B0: This part of the question was not awarded the mark as although the student has some understanding of expectation, the answer is given in the question so merely producing the correct answer is not sufficient for the mark.

Part (c)

B1: The candidate has most likely used the statistical functions on their calculator to find this answer as the answer was not given in the question.

Part (d)

M1A1: This part of the question was awarded both marks as the answer is correct and has not been obtained by an incorrect method. Strictly the answer is two thirds not 0.667 but it is correct to three significant figures and it is clear from the working what the candidate means.

Part (e)

E1E1E0: The candidate obtains the first mark from recognising that the central limit theorem is used when the sample is 'large'.
The second mark in this explanation is given due to the candidate's reference to the distribution of the sample means however the final mark cannot be awarded as they have not stated the change in standard deviation or variance.

Part (f)

B0A0: In this part of the question, the candidate was awarded no marks as neither of the two distributions quoted are the distribution of the sample means $\bar{X} \sim \mathrm{N}\left(2.2, \frac{1.36}{60}\right)$ and the answer is incorrect.

Question 1 - Response C

1 In a video game you can mine blocks of coal. The yield of a block, when mined, is determined randomly.

Mining a block of coal yields different numbers of pieces of coal, which follow the distribution shown in **Figure 1**.

The probabilities, p, of yielding 2, 3 or 4 pieces are equal.

Pieces of coal	1	2	3	4
Probability	0.4	p	p	p

[Source: https://minecraft.fandom.com/wiki/Fortune]

Figure 1

(a) Find the value of p **(1)**

$1 - 0.4 = 0.6$. $0.6 \div 3 = 0.2$

(b) Show the expected yield from mining one block of coal is 2.2 pieces of coal. **(1)**

$(1 \times 0.4) + (2 \times 0.2) + (3 \times 0.2) + (4 \times 0.2) = 2.2$

(c) Find the variance of the yield from mining one block of coal. **(1)**

$E(x) = 2.2^2$ $E(x^2) = (1^2 \times 0.4) + (2^2 \times 0.2) + (3^2 \times 0.2) + (4^2 \times 0.2)$

$= 6.2$

$\text{Var}(x) = E(x^2) - E(x^2) = 6.2 - 2.2^2 = 1.36$

(d) Find the probability that the yield of a block is **one** piece of coal, given that the yield was an odd number of pieces of coal. **(2)**

$P(1 \mid 1, 3) = 0.4 / 0.4 + 0.2 =$

$0.4 / 0.6 = 0.667$ (3sf)

(e) State the central limit theorem. **(3)**

If the number in a sample is large enough, there is no need to have a distribution applied. For example, a Binomial can be approximated to normal when n is larger than 10.

Q1

(f) Estimate the probability that, when 60 blocks of coal are mined, the average yield per block is more than 2.25

(2)

$X \sim N\left(2.2, \frac{\sqrt{1.36}}{\sqrt{60}}\right)$, $P(Y > 2.25) = 0.370$ (3s)

8 /10

Part (a)

B1: No working is required here, however it can be seen that the candidate has arrived at the correct answer with correct working.

Part (b)

B1: The answer is given here so it is important that candidates clearly communicate their working. In this case the candidate is clearly multiplying the number of pieces of coal by their respective probabilities and then adding up the results.

Part (c)

B1: The candidate has correctly worked this out by hand for the mark, it would be possible to simply use a calculator to give the right answer and also get the mark thus saving some time.

Part (d)

M1A1: Marks were awarded for a final answer of $\frac{2}{3}$ or 0.667 and although the initial fraction given here by the candidate could be read as $\frac{0.4}{0.4}+0.2$, it's clear from subsequent working what they meant.

Part (e)

E1E0E0: This part of the question was only given the first mark as the student knows that the central limit theorem applies to large samples however there is no reference to the distribution of means or the standard error.
It is worth noting that the reference to binomial has been disregarded in this case as the requirement for the first mark is only that the central limit theorem refers to large sample sizes.

Part (f)

B1A1: This part of the question was awarded both marks as the answer is correct and although X is stated both in the distribution and in the probability calculation (rather than $\bar{X}$), it is clear what the candidate has done.

The other inaccuracy in this solution is that the candidate has quoted $\frac{\sqrt{1.36}}{\sqrt{60}}$ inside their distribution bracket which is the standard deviation rather than the variance but as this has been used correctly, it has not been penalised.

It is worth noting that mixed variances such as $\frac{1.36}{\sqrt{60}}$ were not awarded the marks and resulted in incorrect final answers.

Q1

A
B
C

Question 2

Introduction | Question | Mark Scheme | Examiner Comments | Performance | Response A | Response B | Response C

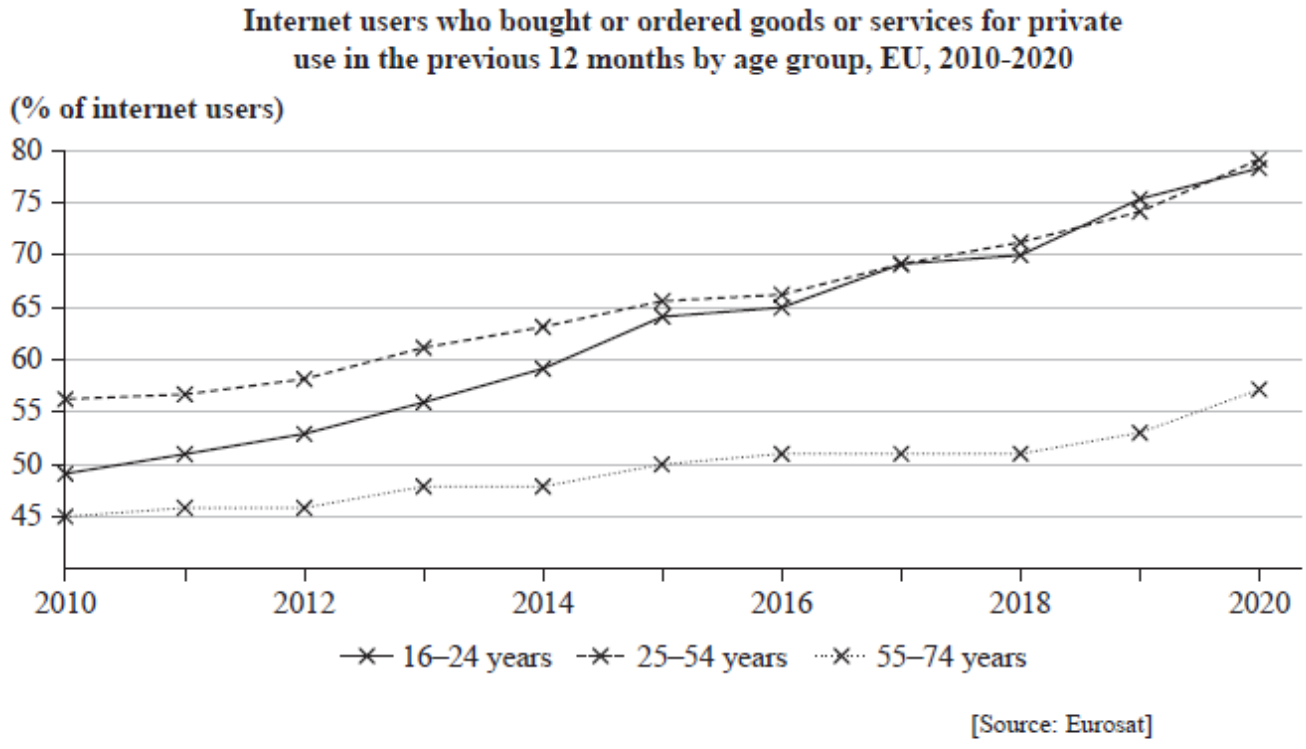
Question 2 - Introduction

This question asked candidates to interpret and justify meaning from a series of diagrams. It is important that candidates read these questions carefully and answer the question that is asked rather than what they think is being asked. Many candidates lost marks in (c) or (d) because they criticised the diagrams rather than trying to infer meaning from them.

Question 2 - Question

2 Brett runs a business based in the European Union (EU). He is planning on launching a website to sell products online and is carrying out research to see if this is a feasible idea.

He finds data on the online shopping habits of different age groups, shown in Figure 2.



[Source: Eurosat]

Figure 2

Question 2 – Question (Cont.)

Brett states that there has been an increase of more than 25 per cent in 55–74 year olds who have made an online purchase in the previous 12 months between 2010 and 2020.

His working is shown below.

$$\text{Difference: } 57 - 45 = 12$$

$$\frac{12}{45} = 0.267 \text{ (3 s.f.)}$$

$$= 26.7\%$$

Brett's colleague, Elias, states that the increase is only 12 per cent.

(a) Briefly discuss Brett's and Elias's statements.

(2)

Figure 3 shows the percentage of people within European countries that used the internet during 2020, and the percentage of people that used the internet to make an online purchase during 2020.

Percentage of internet use and online purchases, 2020
(% of individuals aged 16 to 74)

	Proportion of individuals who:	
	Used within the last 12 months	Purchased online within the last 12 months
EU	89	65
Belgium	92	73
Bulgaria	74	31
Czechia	89	72
Denmark	99	89
Germany	95	83
Estonia	90	68
Ireland	92	74
Greece	79	46
Spain	93	63
France	91	70
Croatia	80	55
Italy	81	44
Cyprus	91	47
Latvia	90	56
Lithuania	84	54
Luxembourg	99	79
Hungary	86	60
Malta	87	87
Netherlands	95	66
Austria	89	66
Poland	85	61
Portugal	79	61
Romania	85	45
Slovenia	88	63
Slovakia	91	62
Finland	97	76
Sweden	97	84
Iceland	99	83
Norway	98	85
Switzerland	97	80
Montenegro	79	23
North Macedonia	84	34
Serbia	79	38
Turkey	78	33
Bosnia and Herzegovina	74	28
Kosovo	97	46

Figure 3

(b) Considering the information in **Figure 3** only, suggest a country to Brett that would be useful to advertise in.

Explain your answer.

(2)

(c) Explain why Brett should **not** make his decision on where to advertise **only** using the information in **Figure 3**.

You should make **three** separate comments.

(3)

Also available is data on employment status, which is included with some additional data in **Figure 4**.

Internet users who bought or ordered goods or services for private use in the previous 12 months, EU[1], 2020
(% of individuals who used the internet in the previous 12 months)

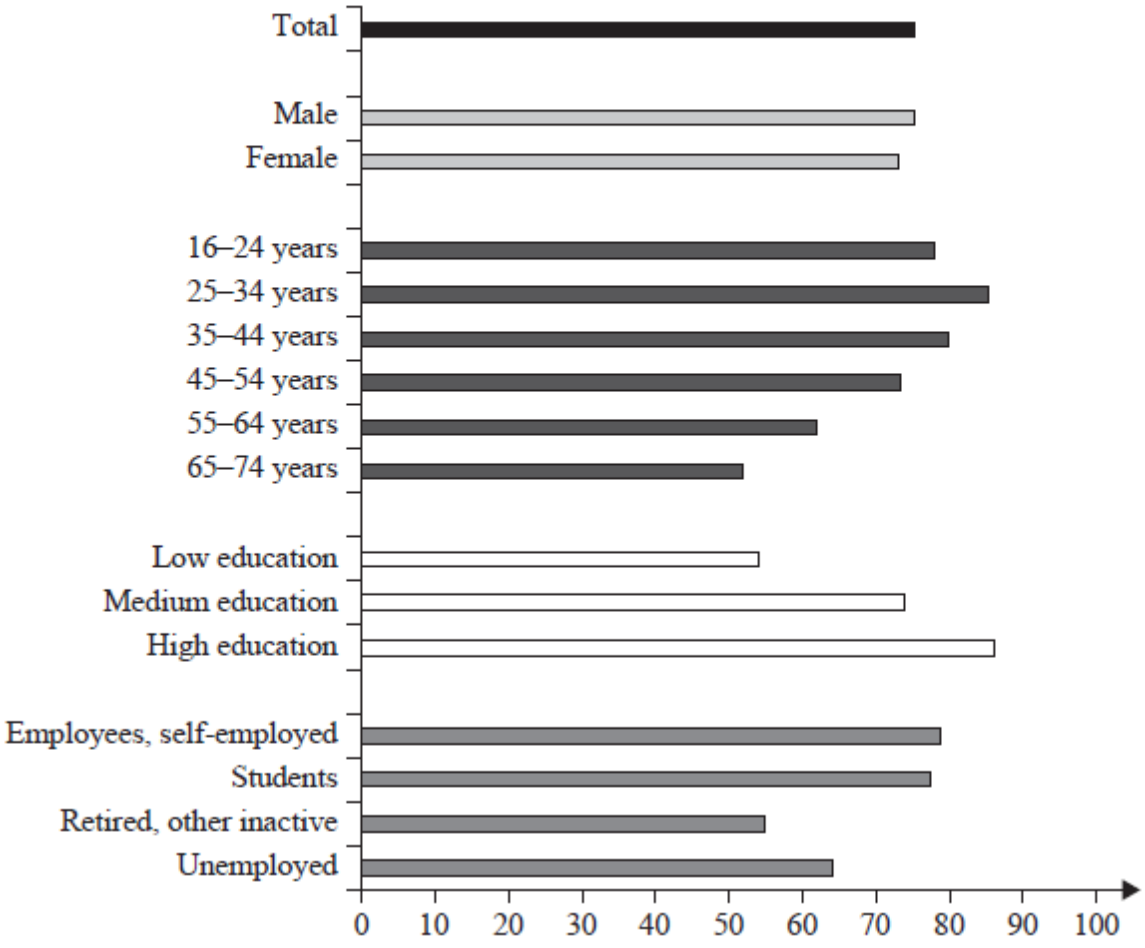


Figure 4

(d) Make **three** comments on the data in **Figure 4**.

(3)

(Total for Question 2 is 10 marks)

Question 2 - Mark Scheme

Question	Scheme	Marks	AO	Notes
2(a)	**Possible comments (not exhaustive)**			
	Elias is just looking at the difference of the percentage points.			
	Brett expresses the difference in the percentage points as a percentage.			
	Both the calculations are correct.			
	Brett's statement may be misleading.			**or** Elias's statement is easier to understand.
		E1, E1	3.1a, 3.1a	E1 for each sensible comment
2(b)	Most likely answers: • Denmark • Germany • Luxembourg • Netherlands • Sweden • Norway • Switzerland	B1	2.1a	**Accept** other countries above EU average in both categories • Belgium • Estonia • Ireland • France • Finland • Iceland
	This country would be useful to advertise in, as a high percentage of the population use the internet and have made a recent online purchase.	E1	2.1b	SC Malta shows all people using internet have purchased online B1E1

2(c)	**Possible explanations (not exhaustive)**			
	The population of a chosen country could be very small, with not many customers to advertise to.			
	Advertising costs could vary greatly by country.			
	It may not be practical to ship his products to all countries.			
	His products may not be popular in every country.			
	Pandemic might change habits			
		E1, E1, E1	3.1a, 3.1a, 3.1a	E1 for each sensible explanation (max E3)
2(d)	**Possible comments (not exhaustive)**			
	The percentage of the population who have used the internet in the last 12 months…			
	Age category			
	…is roughly the same for males and females.			**or** slightly higher for males.
	…is highest in the 25-34 age category.			**or** lowest in the 65-74 category.
	…decreases with age after the 25-34 age category.			

	Education category			
	…is highest in the 'High education' category.			**or** lowest in the 'Low education' category.
	…increases with level of education.			
	Employment category			
	…is highest for employed people [and students].			**or** lowest for retired people.
		E1, E1, E1	1.1, 1.1, 1.1	E1 for each sensible explanation (max E3)
		Total	**10**	

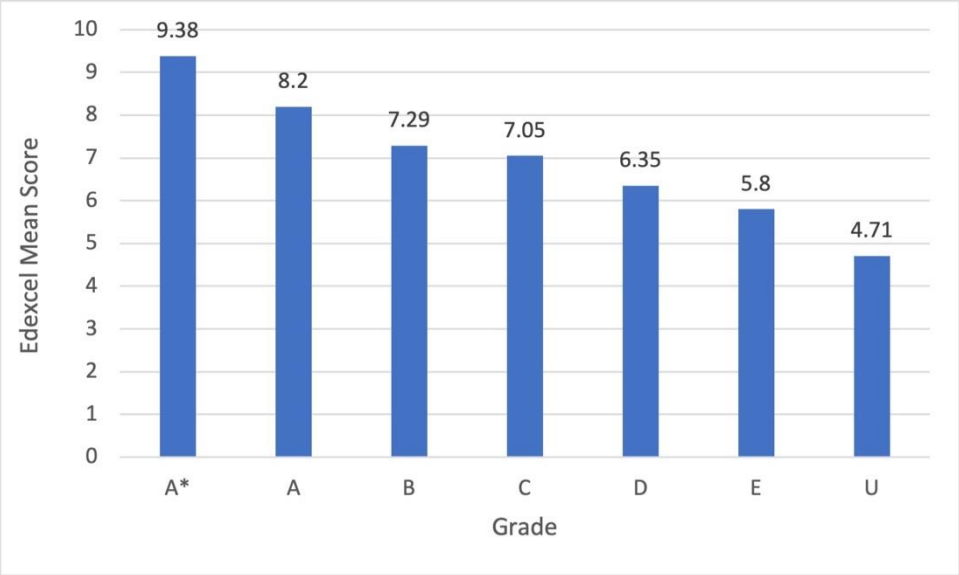
Question 2 - Examiner Comments

Q2(a) Many candidates were able to give good answers here recognising the difference between percentage point increase and percentage change of an amount. Candidates who did not obtain full marks in this question often only made a comment about one of the calculations and observed that the other one was incorrect.

Q2(b) to (d) were the best answered questions on the paper with many candidates able to identify features and good analyses from figures 3 and 4. Common errors observed were from candidates who criticised the representation of the data in figure 4 rather than commenting on the data displayed. In addition, candidates lost marks because they gave more than 3 reasons when only asked for 3, some of which were incorrect.

Question 2 - Performance

Mean score	Max score	Mean %	Edexcel averages: mean scored by candidates achieving grade:							
			ALL	A*	A	B	C	D	E	U
6.76	10	68	6.76	9.38	8.20	7.29	7.05	6.35	5.80	4.71



Question 2 - Response A

Brett states that there has been an increase of more than 25 per cent in 55–74 year olds who have made an online purchase in the previous 12 months between 2010 and 2020.

His working is shown below.

$$\text{Difference: } 57 - 45 = 12$$

$$\frac{12}{45} = 0.267 \text{ (3s.f.)}$$

$$= 26.7\%$$

Brett's colleague, Elias, states that the increase is only 12 per cent.

(a) Briefly discuss Brett's and Elias's statements.

(2)

Both Brett is wrong the difference is 12 which means it has increased by 12%. making Elias statement accurate

(b) Considering the information in **Figure 3** only, suggest a country to Brett that would be useful to advertise in.

Explain your answer.

(2)

Denmark as 99% use the internet and 89% of them made a purchase within the last 12 month

(c) Explain why Brett should **not** make his decision on where to advertise **only** using the information in **Figure 3**.

You should make **three** separate comments.

(3)

different Countries may prefer different products,

Some Countries may only use the internet to do shopping and VIS-versa

There is a large time frame as it is 12 months so the data could be unrepresentative

not all countries may have delivery services

(d) Make **three** comments on the data in **Figure 4**.

(3)

The bars for male and female add up to over 100%
there is also no x axis label

The bars are also larger than the total

The total is not even 100%

3 /10

Q2

Part (a)

E1E0: This candidate was awarded one of the two available marks for this question as they have correctly identified that 12% has come from the difference between the two percentages in 2010 and 2020 however, they haven't made a correct statement about the 26.7% being a percentage change from the number in 2010 and 2020.

Part(b)

B1E1: This candidate was given both marks as they have correctly identified Denmark and have justified it by quoting the high numbers for both usage and purchase.

Part (c)

E0E0E0: This candidate was not awarded any marks for this question.

- "Different countries may prefer different products" was not considered sufficient to award the mark for "Brett's product may not be suitable" in different countries.
- "Some countries may only use the internet to do shopping and vice-versa" was not relevant to the data
- "there is a large time frame as it is 12 months so the data could be unrepresentative" was a contradictory statement. Candidates were allowed to write that 12 months was not enough data to see how shopping trends changed but in this case the candidate has said that it is a big sample and that's not good which is false.
- "not all countries may have delivery services" was considered not relevant as all countries have had some internet shopping.

It is worth noting that this candidate has given four reasons when only three have been asked for. If a candidate gives three correct reasons and one or more incorrect one, then they will be awarded a maximum of 2 marks in this question.

Part (d)

E0E0E0: All of these comments are criticising the representation of this data. The question does not ask for critique of this and instead asks for candidates to comment on the data itself, so these responses are awarded no marks.

Question 2 - Response B

Brett states that there has been an increase of more than 25 per cent in 55–74 year olds who have made an online purchase in the previous 12 months between 2010 and 2020.

His working is shown below.

$$\text{Difference: } 57 - 45 = 12$$

$$\frac{12}{45} = 0.267 \text{ (3s.f.)}$$

$$= 26.7\%$$

Brett's colleague, Elias, states that the increase is only 12 per cent.

(a) Briefly discuss Brett's and Elias's statements. **(2)**

Brett is correct as although there has been a 12% increase of internet users, compared to 2010 results, there has been a 26.7% increase.

(b) Considering the information in **Figure 3** only, suggest a country to Brett that would be useful to advertise in. **(2)**

Explain your answer.

Denmark, as 99% of people use the internet, and a majority of those (89%) have made online purchases in the last 12 months.

(c) Explain why Brett should **not** make his decision on where to advertise **only** using the information in **Figure 3**.

You should make **three** separate comments.

(3)

The price of advertising may vary from country to country.

A country with predominantly more old people will make less purchases (as shown in Figure 2).

~~The income in these~~

~~The country may have a weak currency, leading to less purchases of goods from the website.~~

A country with a bigger population means that the advertisment will be seen by more people.

(d) Make **three** comments on the data in **Figure 4**.

(3)

The axis are unclear

Data is incorrect as it doesn't add up correctly ~~(Total ≈ 80%~~ (75% is female but males take up roughly 90% of the total as well)

~~It's just a stupid graph really.~~

Majority of people who used the internet had high educa

Q2

A
B
C

5 /10

Part (a)

E0E0: No marks were awarded for this part of the question as there is no obvious understanding of how the Brett has arrived at 26.7%. The statement about Elias is incorrect.

Part (b)

B1E1: The candidate has correctly identified Denmark and justified it with the two percentages. Another common answer was "Montenegro" which was underneath the EU average for both "users" and "purchasers" and therefore could not be awarded any marks in this part of the question.

Part (c)

E1E1E0:

- "The price of advertising" was considered worthy of a mark as this is on the mark scheme.
- "A country with predominantly more old people" was not considered worthy of a mark because it doesn't refer directly to the possible type of product Brett might be selling or the fact that the data only shows 16-74 year olds.
- "A country with a bigger population" was considered worthy of a mark as comments about population size rather than proportion were being awarded marks.

Part(d)

E1E0E0:

- "the axis are unclear" is not awarded a mark as that is a comment about how the data is displayed rather than the data itself
- "Data is incorrect" is not awarded a mark as the statement is wrong, the candidate has misunderstood that it is 75% of females have purchased rather than 75% of users of the internet are female
- "Majority of people who used the internet had high education" was awarded a mark because it is a valid comment about the data. It was decided that "users" was sufficient rather than "users who purchased items from the internet"

Question 2 - Response C

Brett states that there has been an increase of more than 25 per cent in 55–74 year olds who have made an online purchase in the previous 12 months between 2010 and 2020.

His working is shown below.

$$\text{Difference: } 57 - 45 = 12$$

$$\frac{12}{45} = 0.267 \text{ (3s.f.)}$$

$$= 26.7\%$$

Brett's colleague, Elias, states that the increase is only 12 per cent.

(a) Briefly discuss Brett's and Elias's statements.

(2)

Bretts calculation is correct. Elias's is not. This is because Elias has mistaken the difference of shoppers to be how much the entire percentage has gone up by. Yes the percentage has increased by 12 percent but that is not how much of a percentage overall has been changed from the orighal 45%.

(b) Considering the information in **Figure 3** only, suggest a country to Brett that would be useful to advertise in.

Explain your answer.

(2)

Brett should look at using advertisement in malta as this has an equal amount of internet users and shoppers. This means that a majority of people use the internet soley for shopping or shopping as well as everything else. Malta also has one of the highest percentages of internet shopping percentages at 87%.

Q2

A B C

(c) Explain why Brett should **not** make his decision on where to advertise **only** using the information in **Figure 3**.

You should make **three** separate comments.

(3)

→ The picture does not state any population numbers so Brett cannot know how many people are actually buying off of and using the internet.

→ It is a very wide age range and doesn't show how many people from each group or what they are buying/ from. So you may not know what they want to purchase.

→ It also does not say how many purchases were ~~made~~ made and sold to customers. And how many were needed to be included within the shopping percentage.

(d) Make **three** comments on the data in **Figure 4**.

(3)

→ people in high education are more likely to be ordering things online as they have a higher percentage of shoppers compared to low & middle education

→ the category with the lowest purchasing percentage is 65-74 year olds. which is just over 50%.

→ more males purchase things online than females. with a percentage of just under 80%.

10 /10

Q2

A
B
C

Part (a)

E1E1: This part of the question was awarded both marks.

Comments about which person was correct was disregarded so long as the understanding about the methods was clear.

The candidate in this response can see that the 26.7% has arisen from a percentage change from the original amount whereas the 12% has arisen as the difference between the two percentages – so a percentage point rise – and therefore has been awarded both marks.

Part (b)

B1E1: This part of the question was awarded both marks. Malta was included in the mark scheme as a special case as it is reasonable to advertise here since everyone using the internet appears to use it to purchase items as both percentages are the same.

Part(c)

E1E1E1: All three marks were awarded here as the candidate has made valid comments about: proportion versus absolute number of people; the type of product being sold not necessarily being appropriate for the population of any country and the total number of purchases in the country rather than just the population size.

Part (d)

E1E1E1: All three marks were awarded for this part as the candidate has made valid comments about three distinct areas of the data.

Question 3

Introduction | Question | Mark Scheme | Examiner Comments | Performance | Response A | Response B | Response C

Question 3 - Introduction

This question was a fairly standard two factor analysis of variance. Many candidates demonstrated a lack of accuracy in their calculations, those seemingly using graphical calculators were generally very good. This question also required candidates to demonstrate their understanding of blocking factors, many of whom struggled to identify and carry out the second F test.

Question 3 - Question

3 Upulani works for a lawnmower manufacturing company.

The company is considering several new batteries to use in a cordless lawnmower.

Upulani measures the lifespan of a random sample of batteries made from four different materials at three different operating temperatures (15°C, 40°C and 65°C).

A longer battery lifespan is more desirable.

She records the lifespan of each battery in hours. Her results are shown in **Figure 5**.

		Temperature			
		15°C	40°C	65°C	Total
Material	A	10.5	6.9	2.2	19.6
	B	10.0	6.6	2.1	18.7
	C	13.7	9.0	2.9	25.6
	D	12.8	8.5	2.8	24.1
	Total	47	31	10	88

Figure 5

The data produced the following summary statistic

$$\sum_i \sum_j x_{ij}^2 = 831.7$$

Upulani claims that there is a difference between average battery lifespan for the four different materials.

(a) Name the technique Upulani should use to carry out this test.

(2)

Question 3 – Question (Cont.)

(b) State **two** necessary assumptions to carry out such a test.

(2)

(c) Carry out a hypothesis test to investigate Upulani's claim.
You may find it helpful to use the table provided.

(11)

Source	**Sums of squares**	**Degrees of freedom**	**Mean square**	**F ratio**
Rows				
Columns				
Error				
Total				

(d) Following your analysis in (c), make a recommendation to the lawnmower company.

(1)

(e) Complete a further analysis to determine whether the blocking factor was effective.

(4)

(Total for Question 3 is 20 marks)

Question 3 - Mark Scheme

Question	**Scheme**	**Marks**	**AO**	**Notes**
3(a)	Two-factor…	B1	1.3	
	ANOVA	B1	1.3	
3(b)	**Possible assumptions**			
	Lifespans are normally distributed The distributions from which the samples are taken have equal variances There is no interaction between the factors (no reason to believe any particular material works any better or worse at any particular temperature.			Context needed

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td></td><td></td><td>B1, B1</td><td>3.1a,
3.1a</td><td>Max 2 marks</td></tr>
<tr><td rowspan="8">3(c)</td><td>H_0: $\mu_A = \mu_B = \mu_C = \mu_D$
H_1: At least two means differ</td><td>B1</td><td>2.1a</td><td>oe</td></tr>
<tr><td>$SS_T = \Sigma\Sigma\, x_{ij}^2 - \frac{T^2}{mn} = 831.7 - \frac{88^2}{12}$
$= 186.36 \ldots$</td><td>M1</td><td>1.3</td><td>PI
SS Total
awrt 186.4</td></tr>
<tr><td>$SS_R = \Sigma\, \frac{R_i^2}{n_i} - \frac{T^2}{mn}$
$SS_R = \frac{19.6^2}{3} + \frac{18.7^2}{3} + \frac{25.6^2}{3} + \frac{24.1^2}{3} - \frac{88^2}{12}$
$= 11.34$</td><td>M1</td><td>1.3</td><td>PI
SS between materials</td></tr>
<tr><td>$SS_C = \Sigma\, \frac{C_i^2}{m_i} - \frac{T^2}{mn}$
$SS_C = \frac{47^2}{4} + \frac{31^2}{4} + \frac{10^2}{4} - \frac{88^2}{12}$
$= 172.16 \ldots$</td><td>M1</td><td>1.3</td><td>PI
SS between temperatures</td></tr>
<tr><td>SS_E = 186.37-11.34-172.17
=2.86</td><td>M1dep</td><td>1.3</td><td>PI
Dep previous two Ms (not negative)</td></tr>
<tr><td>
<table>
<tr><th>Source</th><th>Sums of Squares</th><th>Degrees of freedom</th><th>Mean square</th><th>F ratio</th></tr>
<tr><td>Rows</td><td>11.34</td><td>3</td><td>3.78</td><td>7.93</td></tr>
<tr><td>Columns</td><td>172.17</td><td>2</td><td>86.083</td><td>180.59</td></tr>
<tr><td>Error</td><td>2.86</td><td>6</td><td>0.4767</td><td></td></tr>
<tr><td>Total</td><td>186.37</td><td>11</td><td></td><td></td></tr>
</table>
</td><td></td><td></td><td></td></tr>
<tr><td rowspan="2"></td><td>B1</td><td>1.3</td><td>df correct, 3 and 6</td></tr>
<tr><td>M1</td><td>1.3</td><td>PI
MS=SS/df for between materials and error</td></tr>
<tr><td>$F = \frac{3.78}{0.4767} = 7.93$</td><td>A1</td><td>1.3</td><td>awrt 7.93
or
p = 0.0165
awfw 0.016 ~ 0.017</td></tr>
</table>

Question	Scheme	Marks	AO	Notes
				Correct table implies all previous marks
	Critical value $F_{3,6}(0.05) = 4.757$	B1	1.3	
	7.93 > cv so reject H_0	M1	2.1b	OR p=0.0165<0.05 **PLUS** correct conclusion
	There is significant evidence to suggest at least two average battery lifetimes differ for the four materials used.	E1 dep	2.1a	Conclusion in context Condone "difference between average lifetimes" Dep on correct cv and ts (p-value and sig)
3(d)	It is recommended material C or D is used	B1	2.1a	Condone just material C or material D recommended SC if Ho not rejected then "it doesn't matter so pick the cheapest"
3(e)	Identification of temperatures as the blocking factor	B1	1.3	PI
	F test for difference between temperatures gives *F* =180.6 which is highly significant (5% cv = 5.143)	M1	1.3	For attempting *F* test between temperatures
		A1	1.3	awrt 181 Condone for M1 MS between temperatures (86.083) is the largest (ie temperatures are the largest source of variation in lifespans) **or** uses only the totals/means for each temperature **or** p=4.36×10^{-6}
	So there seems to be a difference between at least two average lifespans for the three **temperatures** **or** Blocking factor was effective	E1 dep	3.1a	Dep some analysis completed
	Total	**20**		

Q3

Question 3 - Examiner Comments

Q3(a) Most candidates were able to identify this as analysis of variance however many didn't state "two factor" and therefore only obtained one mark.

Q3(b) Whilst many candidates knew the assumptions required to carry out an ANOVA test, the vast majority were not able to give the assumptions in the context of the question and therefore were not awarded full marks.

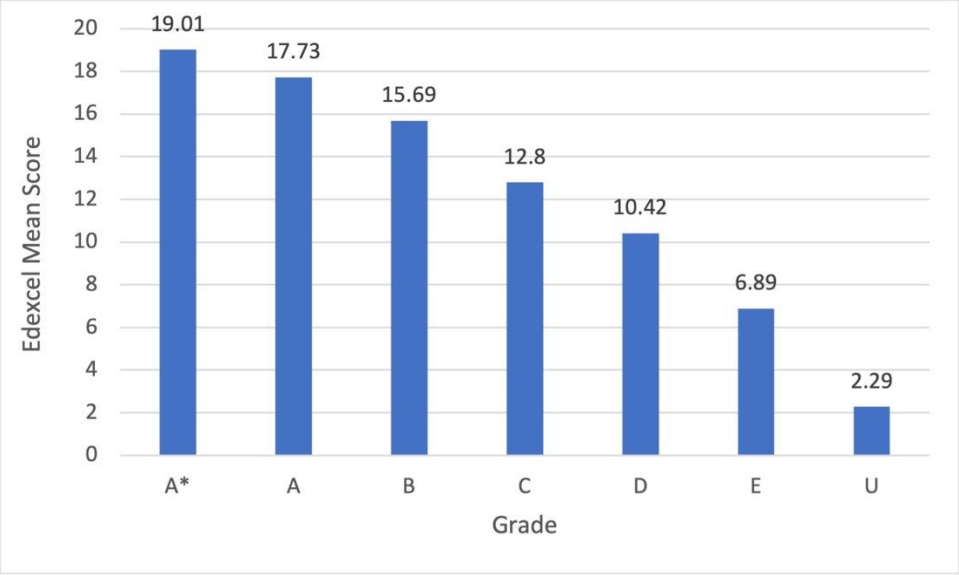
Q3(c) Many candidates were able to accurately carry out the hypothesis test however the most common errors were around stating the hypotheses and finding the correct critical value.

Q3(d) was generally well answered and credit was given to comments from candidates who incorrectly did not reject H_0 in (c) for practical suggestions about which material should be chosen.

Q3(e) This part of the question was poorly answered by many candidates with many responses left blank. Many candidates correctly identified temperature as the blocking factor but then did not make use of the F-ratio already calculated in (c) to make a decision about whether it was effective or not.

Question 3 - Performance

Mean score	Max score	Mean %	Edexcel averages: mean scored by candidates achieving grade:							
			ALL	A*	A	B	C	D	E	U
11.93	20	60	11.93	19.01	17.73	15.69	12.80	10.42	6.89	2.29



Question 3 - Response A

(a) Name the technique Upulani should use to carry out this test.

(2)

Organised Block Method

(b) State **two** necessary assumptions to carry out such a test.

(2)

- The data must be assumed ~~randomly~~ to be taken from a normal distribution
- The data

(c) Carry out a hypothesis test to investigate Upulani's claim.

You may find it helpful to use the table provided.

(11)

Source	Sums of squares	Degrees of freedom	Mean square	*F* ratio
Rows	11.34	2	$11.34/2 = 5.67$	$\frac{5.67}{0.475} = 11.94$
Columns	172.16	3	$172.16/3 = 57.39$	$\frac{57.39}{0.475} = 120.82$
Error	2.85	6	$2.85/6 = 0.475$	
Total	186.36	11		

$SS_T = 831.7 - \frac{88^2}{12} = 186.36$

$SS_R = \frac{19.6^2}{3} + \frac{18.7^2}{3} + \frac{25.6^2}{3} + \frac{24.1^2}{3} - \frac{88^2}{12} = 11.34$

$SS_C = \frac{47^2}{4} + \frac{51^2}{4} + \frac{10^2}{4} - \frac{88^2}{12} = 172.16$

H_0: $\mu_A = \mu_B = \mu_C = \mu_D$

H_1: At least one of the materials affects the average battery lifespan differently to the other materials.

Q3

A B C

T.S > C.V Reject Ho. Sufficient evidence to suggest that
11.94 > 9.12 at least one of the materials affected the
average battery life differently to the other
materials.

(d) Following your analysis in (c), make a recommendation to the lawnmower company.

with the highest totals overall and ~~individually~~ paired
together, material C at 15°C appears to make the ~~batt~~
average battery life last the longest. (1)

(e) Complete a further analysis to determine whether the blocking factor was effective.

~~Definite~~ The effect on the battery life was found to
be ~~definite~~ significant for the materials used,
making the blocking factor effective. (4)

Additionally, the blocking factor of the temperature lead
to an incredibly high F-Ratio, showing that it was a
very effective blocking factor.

Due to both blocking factor's effectiveness the Error
was ~~minimised~~ minimised and remained quite small.

Q3

8 /20

Part (a)
B0B0: The candidate has given neither of the correct statements for these marks.

Part (b)
B1B0: A single mark is awarded for a correct assumption, however this not given in context as in this case.

Part (c)
B1M1M1M1M1B0M1A0B0M1E0: The candidate has switched the degrees of freedom for rows and columns and, although the method is correct, has lost the B mark for the degrees of freedom, the A mark for the correct F ratio and the B mark for the correct critical value. They were able to be awarded the method for a correct comparison and statement of "reject H0" but not the final E as this is dependent on a correct critical value and test statistic.

Part (d)

E1: The candidate correctly identifies that C (or D or both) should be recommended.

Part (e)

B1M1A0E1: Whilst there are no calculations in this part of the question, the candidate has correctly identified temperature as the blocking factor and stated that the F value is very high. Candidates should be comparing test statistics to critical values to make this conclusion; in this case it was decided that candidates would be given credit for stating that the test would be rejected at "any sensible level of significance" without specifically stating any critical value.

Question 3 - Response B

(a) Name the technique Upulani should use to carry out this test.

(2)

ANOVA

(b) State **two** necessary assumptions to carry out such a test.

(2)

All battery lifespans are independent of one another.

All batteries are new.

(c) Carry out a hypothesis test to investigate Upulani's claim.

You may find it helpful to use the table provided.

(11)

Source	Sums of squares	Degrees of freedom	Mean square	*F* ratio
Rows	11.34	3	3.78	0.0439
Columns	172.16	2	86.08	
Error	2.86	2		
Total	186.36	7		

H_0: all the average battery lifespans for the four different materials are the same

H_1: at least two of the average battery lifespans for the four different materials are different

$\frac{19.6^2}{4}$ +

$\sum \frac{R^2}{n} = 492.505$

656.673

~~831.7~~ 831.7

0.0439 < 199.2

Reject H_0 sufficient evidence to suggest at least two of the average battery lifespans are different

HA067208748 11

Question 3 continued

(d) Following your analysis in (c), make a recommendation to the lawnmower company.

(1)

I would recommend battery C as it has the longest mean lifespan

(e) Complete a further analysis to determine whether the blocking factor was effective.

(4)

The blocking factor was not effective as H_0 was rejected and battery C has the best lifespan.

Q3

A B C

8 /20

Part (a)

B0B1: The candidate correctly stated ANOVA for the second B mark but did not identify the test as "two factor".

Part (b)

B0B0: The assumption that the candidate has tried for is that there is no interaction between the factors, but it was not sufficiently explained so did not earn the mark.

Part (c)

B1M1M1M1M1B0M1A0B0M0E0: Candidates should be giving the hypotheses as "means" rather than "averages" however this was condoned, so the hypotheses were given the marks. The calculations for the sum of squares were all correct but then the degrees of freedom were incorrect, the candidate then used a correct method for mean square but the final F value was incorrect so the A mark could not be awarded.

The correct critical value was not found however if the comparison had been done correctly for an incorrect critical value, that is correctly rejecting H0 or not, then the final method mark could be awarded. In this instance the candidate incorrectly said that their test statistic being less than a critical value meant that the null hypothesis would be rejected.

The final explanation mark is dependent on having the correct test statistic and critical value.

Part (d)

E1: The candidate correctly identifies that C (or D or both) should be recommended. Candidates who had incorrectly not rejected H0 in (c) could be given credit here if they said that the decision about which material should be used should be based on another factor such as cost.

Part (e)

B0M0A0E0: The candidate does not identify the temperature as a blocking factor and gives no working worthy of credit.

Question 3 - Response C

(a) Name the technique Upulani should use to carry out this test.

(2)

Anova / Analysis of variance

(two factor)

(b) State **two** necessary assumptions to carry out such a test.

(2)

Question 3 continued

(c) Carry out a hypothesis test to investigate Upulani's claim.

You may find it helpful to use the table provided.

(11)

Source	Sums of squares	Degrees of freedom	Mean square	*F* ratio
Rows	11.34	3	3.78	~~180.54~~ 7.93
Columns	172.16	~~2~~ 2	86.083	~~410.8~~ 180.54
Error	2.86	6	0.4766	
Total	186.36	11		

P
0.0164
~~0.0000043~~
~~0.0164~~ 0.0000043

for temperature

H0: All the means are the same

H1: At least one of the means are different

P < 0.05

~~0.0164~~ 0.0000043

reject ~~to~~ H0. Sufficient evidence to suggest one of the means for temperature is different and that operating temperature has an effect on lifespan

Q3

A
B
C

H0: All the means for materials are the same
H1: ~~the~~ ~~for~~ atleast one of the means are different

$0.0164 < 0.05$

reject H0. Sufficient evidence to suggest one of the means to materials is different and that the material has an effect on lifespan

Question 3 continued

(d) Following your analysis in (c), make a recommendation to the lawnmower company.

(1)

have temprature at 15°C and use material C as they both have the highest means

(e) Complete a further analysis to determine whether the blocking factor was effective.

(4)

$86.083 > 0.4766$
$3.78 > 0.4766$

Therefore ~~the blocking factors were effective~~ sufficient evidence to suggest the blocking factors ~~were effect~~ are effective.

Q3

14 /20

Part (a)

B1B1: Marks are awarded here for each of "ANOVA" and "two factor" independently of each other. "Two way" was also allowed.

Part (b)

B0B0: This was not attempted by the candidate here however it is worth noting that many candidates knew the assumptions but did not put them in context; thereby losing one of the marks. Candidates writing one conclusion without context were awarded the first B mark.

Part (c)

B1M1M1M1M1B1M1A1B1M1E1: The hypotheses here are not strictly correct but would be allowed the mark. The mark scheme states "at least two means differ" and candidates should be encouraged to write that as the most efficient way of getting the mark.

This candidate has used the p-value to compare which is fine, many students chose to use a critical value method.

Although conclusions for both factors were given here, it is clear which one is for materials, and which is for temperature and therefore the candidate was not penalised.

Part (d)

E1: The candidate correctly identifies that C (or D or both) should be recommended. Candidates who had incorrectly not rejected H0 in (c) could be given credit here if they said that the decision about which material should be used should be based on another factor such as cost.

Part (e)

B0M0A0E0: This candidate has not identified temperature as the blocking factor and whilst they have correctly compared and concluded about the temperature in (c), they could not be awarded the marks here because this did not show any understanding about blocking factors.

Question 4

Introduction | Question | Mark Scheme | Examiner Comments | Performance | Response A | Response B | Response C

Question 4 - Introduction

In this question the E marks for the justification of answers were dependent on having the correct model and correct assumptions that didn't directly lead to the model in question were not given credit.

Question 4 - Question

4 Anna is a statistician. She is employed to advise companies on suitable statistical models to use in a variety of scenarios.

Recently, she has been asked to model three scenarios.

(a) For each of the following scenarios, explain which model you would expect Anna to use, and justify your answers.

(i) The number of days in January on which it snows,

(2)

(ii) The number of patients attending A&E between 9pm and 10pm on a Monday evening,

(2)

(iii) The amount of time between successive visits to a company's website.

(3)

(b) For **one** of the scenarios in (*a*), give a reason, in context, why the distribution you suggested may **not** be appropriate to model the situation.

Clearly state which model you are referring to.

(1)

(Total for Question 4 is 8 marks)

Question 4 - Mark Scheme

Question	Scheme	Marks	AO	Notes
4(a)(i)	Binomial distribution (The number of days fixed, and) the probability of snow is likely to be similar in all days of January	B1 E1	2.1b 2.1b	
4(a)(ii)	Poisson distribution The visitors could be reasonably expected to come with a constant average rate	B1 E1	2.1b 2.1b	
4(a)(iii)	Exponential distribution Visits to a website are likely to occur at a constant average rate and therefore follow a Poisson distribution The time between consecutive Poisson events follows the exponential distribution	B1 E1 E1	2.1b 2.1b 2.1b	
4(b)	**Binomial model** Unlikely to be independent if it snows on one day to the next as it may be more likely to snow in periods of cold weather **Poisson model** If two people are involved in an accident together they will not attend independently **Exponential model** Visitors are unlikely to visit at a constant rate across the whole day If something goes viral on the website clicks are unlikely to be independent Any sensible reason attached to model	E1dep	3.1a	List not exhaustive **Dep on student using correct model in (a)**
	Total	**8**		

Q4

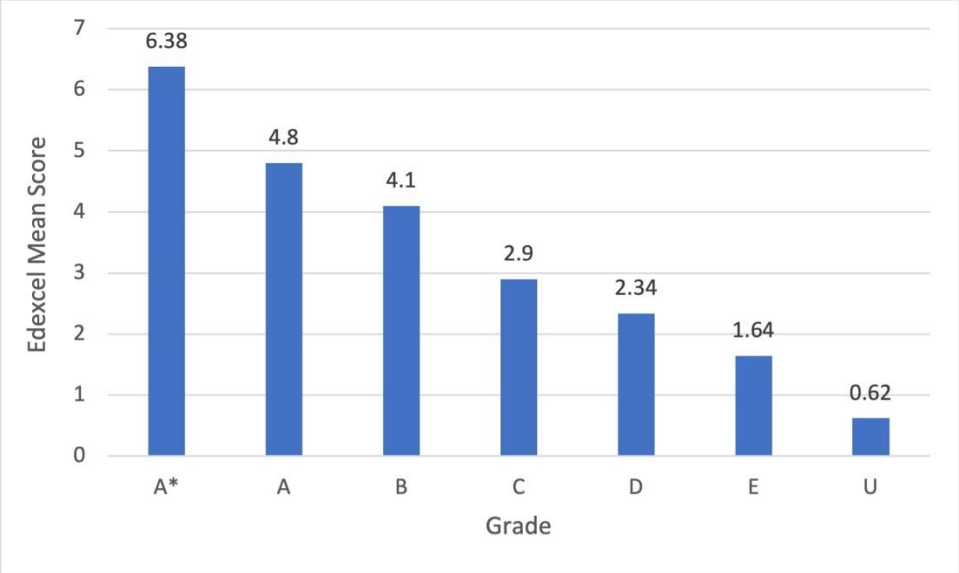
A
B
C

Question 4 - Examiner Comments

This proved to be the most challenging of the questions for many candidates and in many cases even where candidates were able to correctly identify the right models for each scenario, they were unable to articulate the reasons for their choice.

Question 4 - Performance

Mean score	Max score	Mean %	Edexcel averages: mean scored by candidates achieving grade:							
			ALL	A*	A	B	C	D	E	U
2.93	8	37	2.93	6.38	4.80	4.10	2.90	2.34	1.64	0.62



Question 4 - Response A

4 Anna is a statistician. She is employed to advise companies on suitable statistical models to use in a variety of scenarios.

Recently, she has been asked to model three scenarios.

(a) For each of the following scenarios, explain which model you would expect Anna to use, and justify your answers.

(i) The number of days in January on which it snows,

(2)

The chi-squared (X^2) model because it's a multiple variable table with observed values.

(ii) The number of patients attending A&E between 9pm and 10pm on a Monday evening,

(2)

The poisson distribution model because it involves several events within a given period of time and the data is discrete

(iii) The amount of time between successive visits to a company's website.

(3)

The exponential distribution model because it involves time inbetween events and the data is continuous.

Question 4 continued

(b) For **one** of the scenarios in (a), give a reason, in context, why the distribution you suggested may **not** be appropriate to model the situation.

Clearly state which model you are referring to.

(1)

The poisson distribution model may not be appropriate because it makes the assumption that the data is independent when the number of patients attending A&E between 9pm and 10pm on a Monday evening may not be independent.

2 /8

Q4

Part (a)(i)
B0E0: The candidate did not identify the correct model and so could not be awarded either mark.

Part (a)(ii)
B1E0: The candidate has correctly identified Poisson distribution but there is no justification in context.

Part (a)(iii)
B1E0E0: This is the same as (a)(ii) where there is no justification in the context of the question

Part (b)
E0: The mark was not awarded here because there is no explanation why patients may not arrive at A and E independently.

Question 4 - Response B

4 Anna is a statistician. She is employed to advise companies on suitable statistical models to use in a variety of scenarios.

Recently, she has been asked to model three scenarios.

(a) For each of the following scenarios, explain which model you would expect Anna to use, and justify your answers.

(i) The number of days in January on which it snows,

(2)

~~poisson~~ binomial

random only 2 outcomes, yes or no
independant
fixed ~~time~~ number of days

(ii) The number of patients attending A&E between 9pm and 10pm on a Monday evening,

(2)

poisson

independant
random
low p

(iii) The amount of time between successive visits to a company's website.

(3)

exponential
time between events
independant

(b) For **one** of the scenarios in (a), give a reason, in context, why the distribution you suggested may **not** be appropriate to model the situation.

Clearly state which model you are referring to.

(1)

poisson for A and E, multiple people may attend at one time from the same thing injuring both of them so each event may not be independant

4 /8

Part (a)(i) - B1E0
Part (a)(ii) - B1E0
Part (a)(iii)
B1E0E0: In the first three parts of the question the candidate has correctly identified each of the models but has given no context to the justifications so could not be awarded any of the E marks.

Part (b)
E1: This is a good example of a correct answer to this question.

Question 4 - Response C

4 Anna is a statistician. She is employed to advise companies on suitable statistical models to use in a variety of scenarios.

Recently, she has been asked to model three scenarios.

(a) For each of the following scenarios, explain which model you would expect Anna to use, and justify your answers.

(i) The number of days in January on which it snows,

(2)

Binomial - Fixed number of trials - 31 days in January. - Can be used to model providing Snow fall is independant from one another and has a fixed probability of it Snowing.

(ii) The number of patients attending A&E between 9pm and 10pm on a Monday evening,

(2)

Poisson due to the period being over a fixed ~~number of~~ period of time. Can be used providing Patients attend at a constant average rate which is likely assued as independant from one another which could also be likely due to illnesses during the night.

(iii) The amount of time between successive visits to a company's website.

(3)

Exponential distribution due to data being Continuous. Website visits likely to occur at a constant average rate and could be independant from one another.

Q4

A
B
C

(b) For **one** of the scenarios in (a), give a reason, in context, why the distribution you suggested may **not** be appropriate to model the situation.

Clearly state which model you are referring to.

(1)

For number of patients attending A & E, this wouldn't be unlikely to be best to model modelled by Poisson distribution as people are unlikely to attend singly in turn meaning it doesn't satisfy the condition for Poisson as A & E is busy.

7 /8

Q4

Part (a)(i)

B1E1: The candidate has correctly identified the binomial distribution and that the probability for snow remained constant.

Part (a)(ii)

B1E1 The candidate has correctly identified the Poisson distribution and the need for a constant average rate of patients.

Part (a)(iii)

B1E1E0: The candidate did not gain the second E mark in this question as they did not state that the times between patients arriving needed to for a Poisson distribution.

Part (b)

E1: The reason that people are unlikely to attend singly as A and E is busy as about the bare minimum for this mark. Candidates were not given the mark for simply stating that an assumption was unlikely be true, they needed to give a reason why the assumption was unlikely to be true.

A
B
C

Question 5

Introduction | Question | Mark Scheme | Examiner Comments | Performance | Response A | Response B | Response C

Question 5 - Introduction

This question required students to understand a section of the bivariate data topic and give most of the responses within the context of the question.

Question 5 - Question

5 Anja is investigating 'Speedcubing', a competition in which participants attempt to solve puzzle cubes in the fastest time possible.

One category in Speedcubing is 'Individual time'. Individual time considers the time taken to solve a single scramble of a cube puzzle.

Another category, 'Average time', considers the average time taken when solving 5 different scrambles of the same cube puzzle.

Anja finds the results from one Speedcubing competition.

For a random sample of participants in this competition, Anja calculates the least squares regression line between Individual time, x seconds, and Average time, y seconds in the form $y = a + bx$

(a) Explain why Anja should expect her value of b to be positive.

(1)

Anja's regression line has equation $y = -0.374 + 1.564x$

[Source: https://www.worldcubeassociation.org/]

(b) Estimate the Average time for a competitor with an Individual time of 9 seconds.

(2)

(c) Explain why there is no sensible interpretation of the value of -0.374 in Anja's model.

(1)

(d) Interpret the value of 1.564 in context.

(1)

Question 5 – Question (Cont.)

Anja finds a database of results of all Speedcubing competitions.

Each competitor has a unique ID, so that if there are two competitors with the same name they can identify the times.

Records containing the competitors and their unique IDs are listed in one table, and a competitor's best times in the Individual time category are listed in another table, also with their unique ID.

(e) Explain how Anja could obtain a list of best times, from her database, with each time shown beside the competitor's name instead of their unique ID.

(3)

(Total for Question 5 is 8 marks)

Question 5 - Mark Scheme

Question	Scheme	Marks	AO	Notes
5(a)	As participants who have a slower individual time will probably have a slower average time	E1	2.1a	
5(b)	-0.374 + (1.564 × 9)	M1	1.2	PI
	$y = 13.70$ seconds	A1	1.2	
5(c)	It is impossible to have an individual time of 0 seconds. It is impossible to have a negative average time.	E1	2.1a	Either
5(d)	When a competitor's individual time increases by 1 second their average time increases by 1.564 seconds.	E1	2.1a	
5(e)	**Query** the database to produce a report **Join** the tables together Using the **unique ID**	E1 E1 E1	1.1 1.1 1.1	

It is possible to see a SQL query here, e.g.

SELECT firstName, surname, uniqueID, bestTime FROM competitors, times WHERE competitors.uniqueID = times.uniqueID

Or a similar version with INNER JOIN

Sight of a SQL statement should be awarded the first mark.

A sensible INNER JOIN or WHERE statement would award the second and third if unique ID field was mentioned.

Total	**8**

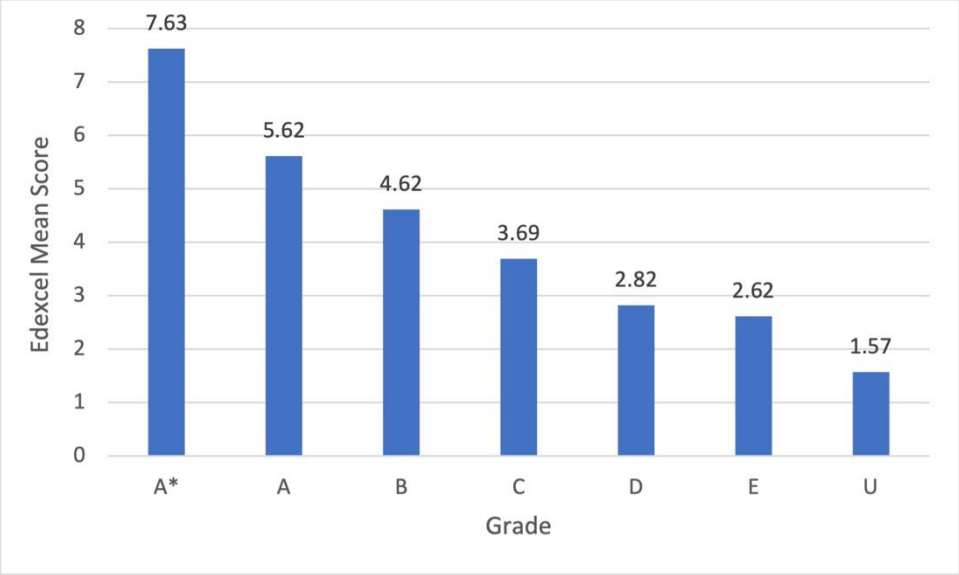
Question 5 - Examiner Comments

Q5(a) to (d) were well answered by many candidates although some candidates struggled to give an interpretation for the gradient in the context of the question.

Q5(e) This part of the question proved very challenging for many candidates. Many candidates tried to circumvent the use of a database by describing how a spreadsheet could be used to solve the problem or did not read the question fully and described how a list of fastest times could be found from a list of all times.

Question 5 - Performance

Mean score	Max score	Mean %	Edexcel averages: mean scored by candidates achieving grade:							
			ALL	A*	A	B	C	D	E	U
3.63	8	45	3.63	7.63	5.62	4.62	3.69	2.82	2.62	1.57



Question 5 - Response A

(a) Explain why Anja should expect her value of b to be positive.

(1)

There can't be a negative amount of participants and the time they take. The same individuals are doing both tests.

Anja's regression line has equation $y = -0.374 + 1.564x$ ← indiv sec.

↑ average [Source: https://www.worldcubeassociation.org/]

(b) Estimate the Average time for a competitor with an Individual time of 9 seconds.

(2)

$y = -0.374 + 1.564 \times 9\ (x)$

$= 13.702$ 13.7 seconds

(c) Explain why there is no sensible interpretation of the value of –0.374 in Anja's model.

(1)

There is a weak negative correlation between the ~~time~~ individual time and average time. There can't be a negative amount of individuals and the time they take.

(d) Interpret the value of 1.564 in context.

(1)

It's the individual time, how much it goes up by.

(e) Explain how Anja could obtain a list of best times, from her database, with each time shown beside the competitor's name instead of their unique ID.

(3)

She could ~~put~~ turn on the setting of lowest times (quickest) down to highest times (the longest) which would link to the ~~ID~~ unique ID's of the individuals.

Q5

A B C

3 /8

Part (a)
E0: This was the most common incorrect answer to this question where candidates did not understand the meaning of the gradient of the line in the context of the question.

Part (b)
M1A1: Although the "x" is still present in the substitution, the candidate's intent is clear and the answer is correct.

Part (c)
E1: Again this is a good answer to this question.

Part (d)
E0: The candidate probably understands that they are trying to describe the gradient of the line but needs to state that it is the amount the average time goes up by for each extra second of individual time.

Part (e)
E0E0E0: Again this candidate has tried to sort the data as though it was in a spreadsheet and has been awarded no marks.

Question 5 - Response B

(a) Explain why Anja should expect her value of b to be positive.

(1)

Because average time will always ~~be~~ increase as individual time goes up.

Anja's regression line has equation $y = -0.374 + 1.564x$

[Source: https://www.worldcubeassociation.org/]

(b) Estimate the Average time for a competitor with an Individual time of 9 seconds.

(2)

$y = -0.374 + 1.564 \times 9$

$y = 13.$

14.

(c) Explain why there is no sensible interpretation of the value of -0.374 in Anja's model.

(1)

Because with both factors measured being time when the y intercept is at 0 seconds the individual time wouldn't make sense to be -0.374.

(d) Interpret the value of 1.564 in context.

(1)

1.564 is the gradient of the line, meaning each time individual time increases, but average time will increase by 1.564.

(e) Explain how Anja could obtain a list of best times, from her database, with each time shown beside the competitor's name instead of their unique ID.

(3)

- Filter by competitors names.
- Sort by "time" in descending order.

Q5

A B C

4 /8

Part (a)
E1: This was considered an acceptable answer to this question.

Part (b)
M1A0: Despite the fact that the candidate crossed out their working, it is possible to see the correct substitution but not a correct answer so the M mark was awarded.

Part (c)
E1: This was about the bare minimum for this mark although it would have been better for the candidate to say that any negative time did not make sense.

Part (d)
E1: This answer is a good answer for this part of the question.

Part (e)
E0E0E0: The candidate here has confused the question, thinking that the tables are in a spreadsheet so that they can be sorted which was not awarded any marks.

Question 5 - Response C

(a) Explain why Anja should expect her value of b to be positive.

(1)

If time taken to solve a single scramble increases, the average time taken when solving 5 different scrambles of the same cube puzzle is likely to increase.

Anja's regression line has equation $y = -0.374 + 1.564x$

[Source: https://www.worldcubeassociation.org/]

(b) Estimate the Average time for a competitor with an Individual time of 9 seconds.

(2)

average time $= -0.374 + 1.564 \times 9 = 13.702$ seconds

average time $= 13.702$ seconds

(c) Explain why there is no sensible interpretation of the value of -0.374 in Anja's model.

(1)

Time taken to solve a scramble of cube puzzle must be greater than zero.

(d) Interpret the value of 1.564 in context.

(1)

For every second to solve a single scramble of a cube puzzle, the average time taken will be 1.564 when solving 5 different scrambles of the same cube puzzle.

(e) Explain how Anja could obtain a list of best times, from her database, with each time shown beside the competitor's name instead of their unique ID.

(3)

Query to join these two tables by the competitors' unique IDs.
~~Sort~~ Filter results of each competitor using the shortest ~~fastest~~ individual time
~~Sort those shortest individual time in ascending order.~~
Only show competitor's name and shortest individual time.

~~SELECT table name, another table name~~

8 /8

Q5

Part(a)
E1: This is a well explained description of why the gradient of the line should be positive.

Part (b)
M1A1: This calculation has been performed accurately and answers given to more than 3 significant figures were ok.

Part (c)
E1: This was the most common correct answer to this question.

Part (d)
E1: It would be preferable for the candidate to have said that "for every second extra to solve a single … the average time taken would be 1.564 extra …" but the candidate's intent was considered clear enough for the mark.

Part (e)
E1E1E1: All the marks for this question were awarded from the first sentence in this solution. Many candidates attempted to explain how they would sort the data but this was not required for the question.

Question 6

Question 6 - Introduction

This was the largest question on the paper and required candidates to demonstrate their knowledge of a wide range of different topics. Candidate needed to be able to explain different aspects of modelling and hypothesis testing in the context of the question as well as carry out various calculations accurately.

Question 6 - Question

6 In field hockey, some fouls result in the awarding of a penalty corner. A player then has the opportunity to score a goal from the penalty corner.

A study of 84 games in the 1998 Field Hockey World Cup found that

- the mean number of penalty corners awarded per game was 2.976
- the standard deviation of penalty corners awarded per game was 0.98
- 114 penalty corners resulted in a goal, and 136 did not.

[Source: https://www.researchgate.net/figure/Data-of-successfulpenalty-corners_tbl2_233698670]

Douglas plays for a field hockey team.

Sam, a sports data analyst, decides to analyse the average number of penalty corners awarded per game for Douglas's team.

Sam studies data from a random sample of 15 games played by Douglas's team over the last three years.

Sam found that

- the mean number of penalty corners awarded per game was 3.25
- $s = 1.2$

(a) Making any necessary assumptions, calculate a 95% confidence interval for the mean number of penalty corners awarded per game for Douglas's team.

(3)

(b) State the necessary distributional assumption required for the confidence interval found in (*a*) to be valid.

(1)

(c) By making reference to the confidence interval found in (*a*), comment on the claim that, over the last three years, the number of penalty corners awarded to Douglas's team per game, on average, differs to that awarded to teams at the 1998 Field Hockey World Cup.

(2)

Question 6 – Question (Cont.)

Douglas believes that a **lower** proportion of penalty corners that his team is awarded result in goals compared to that for teams playing in the 1998 Field Hockey World Cup.

Last year, his team was awarded 35 penalty corners resulting in 7 goals.

Douglas decides to test the following hypotheses:

H_0: $\pi = 0.456$

H_1: $\pi < 0.456$

(d) Show how Douglas obtained the value of 0.456 for his hypothesis test.

(1)

(e) Explain why the critical region for Douglas's test is $X \leqslant 10$, where X represents the number of goals scored from 35 penalty corners.

(2)

(f) Complete Douglas's hypothesis test.

(2)

(g) Give one advantage of using a critical region instead of a p-value when carrying out hypothesis tests.

(1)

(h) Describe, in context, the meaning of a Type I error for Douglas's test.

(2)

(i) Write down the exact probability that Douglas makes a Type I error when carrying out his test.

(1)

(j) If the probability that, for Douglas's team, a penalty corner results in a goal is actually 0.15, find the power of his test.

(3)

(k) State **two** necessary assumptions that you made when selecting the distribution to use in (e).

You should explain whether you believe **each** to be valid or not.

(4)

During the 1998 Field Hockey World Cup, specialist penalty corner players were brought on to substitute for other players when a penalty corner was awarded.

In 1999, a new rule was introduced that specialist penalty corner players could not be brought on, unless a player was injured.

(l) Explain how this information may affect the validity of the conclusion you made in (f).

(2)

(Total Marks for Question 6 is 24 marks)

Question 6 - Mark Scheme

Question	Scheme	Marks	AO	Notes
6(a)	($\bar{x}$ = 3.25 s = 1.2 n = 15)			PI
	$3.25 \pm 2.1448 \times \frac{1.2}{\sqrt{15}}$	M1	1.3	Use of $\frac{1.2}{\sqrt{15}}$
				PI
		B1	1.3	t_{14} = 2.14(5) or z=1.96 used
				awrt
				SC (2.64, 3.86) if z=1.96 used scores M1B1A0
	CI=(2.59, 3.91)	A1	1.3	SC 15 df used t = 2.131 (2.58, 3.91) M1B1A0
6(b)	Number of penalty corners awarded per game is normally distributed.	E1	2.1a	Context required
6(c)	2.976 is within the CI…	B1ft	2.1b	oe comparison of 2.976 with their CI
	…so there is not enough evidence to suggest Douglas' team were awarded a different number of penalties per game, on average, than those awarded at the 1998 Field Hockey World Cup.	E1dep	2.1a	oe Dep on previous B1
6(d)	$\frac{114}{114+136} = 0.456$	B1	2.1a	AG
6(e)	Because $P(X \leq 10) = 0.03024$ and $0.03024 < 0.05$	E1	1.3	
	But $P(X \leq 11) = 0.06361$ and is $0.06361 > 0.05$	E1	1.3	
6(f)	H_0: $\pi = 0.456$ H_1: $\pi < 0.456$			Not needed, given before (d) but may be seen
	7< 10, Reject H_0	M1	2.1b	Or 0.00149 < 0.05

Question	Scheme	Marks	AO	Notes
	Evidence to support Douglas's belief (his team score fewer goals from penalty corners than expected)	E1dep	2.1a	Dependent on M1
6(g)	You can test the same hypotheses again without further calculation	E1	1.3	
6(h)	H_0 is rejected in error so…	E1	1.3	Explanation mark Implied by correct contextual answer
	…concluding, for his team, penalty corners resulted in fewer goals than expected when actually they result in at least as many as expected.	E1	2.1a	Context mark
6(i)	0.03024	A1ft	1.3	ft their (e) awrt 0.03
6(j)	X~B(35,0.15)	M1	2.1a	
	P(X>10)=0.01098	A1	1.2	SC P(X≥10)=0.0292 scores M1A1A0
	Power = 1-0.01098=0.98902	A1	1.2	awrt 0.99
6(k)	**Constant probability** is unlikely to be valid because… …different players may have different chance of scoring and it's unlikely the same player is always involved. …different teams may be better at defending	E1	3.1a	Constant probability
		E1	2.1a	Any sensible reason
	Independence is unlikely to be valid because… …decisions of tactics for taking a penalty corner may be impacted by the previous penalty corners. …confidence and form may be impacted by scoring/missing previous penalty corner.	E1	3.1a	Independence
		E1	2.1a	Any sensible reason

Q6

A B C

Question	Scheme	Marks	AO	Notes
6(l)	Douglas's team might not have specialist corner takers. **or** Data from the 1998 Field Hockey World Cup predates the change in rule but Douglas's team's data is after the rule change.	E1	3.1b	oe
	So the conclusion may not be valid because Douglas's team's record of scoring from penalties is not being compared with the same circumstances (in terms of specialist players being able to try to score goals from penalties).	E1	3.1b	
	Total	**24**		

Question 6 - Examiner Comments

Q6(a) was well attempted by most candidates however common errors observed were from candidates using values from the normal distribution rather than the t-distribution as well as many candidates using one-tailed values rather than two tailed ones.

Q6(b) was poorly answered by most candidates, many of whom knew that the distributional assumption was that the data needed to be normally distributed but could not put this in the context of the question.

Q6(c) The most common error in this part of the question was that candidates concluded too definitely. Candidates should be reminded that conclusions from confidence intervals, much like for hypothesis tests, most not be too certain as confidence intervals only provide evidence about the parameters of populations.

Q6(d) was well answered by the majority of candidates.

Q6(e) Some candidates were able to give a well-reasoned argument to answer this question but many could not identify what they needed to do in order to show that $X \leq 10$ was the critical region. Some candidates attempted to use the inverse function on their calculators but very few of these were then able to write the notation or argument to show the correct critical region.

Q6(f) This question in general was well answered however many candidates misunderstood the question and tried to use a test statistic of 10 rather than 7. Many candidates used a p-value in this part of the question which was perfectly suitable if more time consuming than directly comparing the test statistic and critical value.

Q6(g) Very few candidates were able to articulate that a critical value allowed repeated tests to be carried out without need for a p-value calculation.

Q6(h) This part was generally well answered with the most common loss of marks being due to a lack of context within their explanations.

Q6(i) Although some candidates were able to show the critical region of $X \leq 10$ accurately in (e) many could not interpret their values as the probability of a type I error with the most common wrong answer being 0.05.

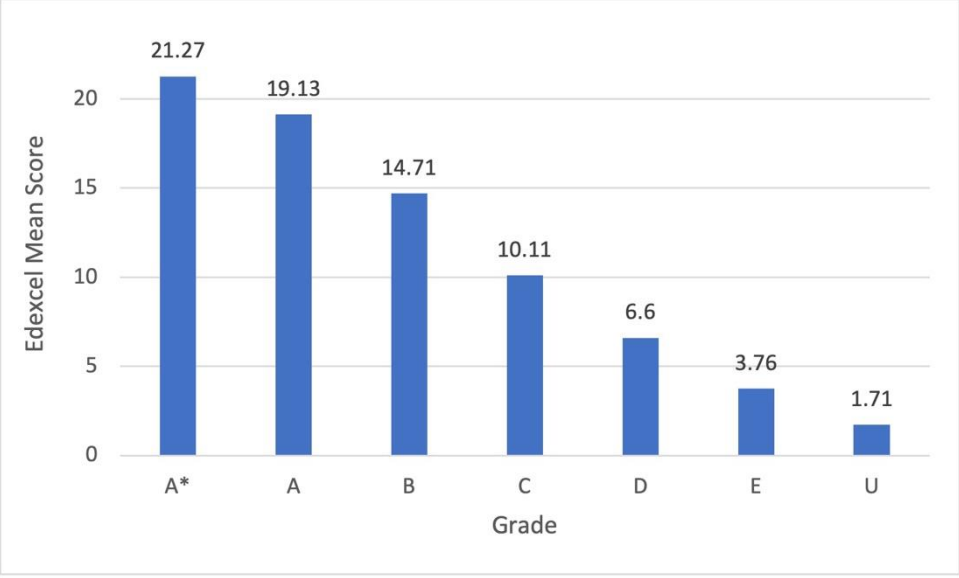
Q6(j) Pleasingly many candidates were able to give full solutions for this part of the question however some candidates confused the probability of a type II error with the power of the test.

Q6(k) Many candidates were not able to identify the requirements for a binomial distribution here with a common error being that candidate had not read the question fully and commented on the use of a normal distribution in (a).

Q6(l) This was generally well answered by candidates who were able to identify that since the conditions for scoring had changed, the conclusion was no longer valid.

Question 6 - Performance

Mean score	Max score	Mean %	Edexcel averages: mean scored by candidates achieving grade:							
			ALL	A*	A	B	C	D	E	U
9.84	24	41	9.84	21.27	19.13	14.71	10.11	6.60	3.76	1.71



Question 6 - Response A

(a) Making any necessary assumptions, calculate a 95% confidence interval for the mean number of penalty corners awarded per game for Douglas's team.

(3)

$t = 1.7613$

$\mu = 3.25$

$s = 1.2$

$n = 15$

$= 3.25 \pm 1.7613\left(\frac{1.2}{\sqrt{15}}\right)$

$= (2.704, 3.796)$ (3dp)

penalty corners awarded per game for Douglas's team.

(b) State the necessary distributional assumption required for the confidence interval found in (a) to be valid.

(1)

- normally distributed

(c) By making reference to the confidence interval found in (a), comment on the claim that, over the last three years, the number of penalty corners awarded to Douglas's team per game, on average, differs to that awarded to teams at the 1998 Field Hockey World Cup.

(2)

The mean number of penalty corners awarded per game was 2.976, which lays within the confidence interval. However, 3.25 is greater than 2.976 so they do differ so the claim is true.

Q6

A B C

Douglas believes that a **lower** proportion of penalty corners that his team is awarded result in goals compared to ~~that for~~ teams playing in the 1998 Field Hockey World Cup.

Last year, his team was awarded 35 penalty corners resulting in 7 goals.

Douglas decides to test the following hypotheses:

$H_0: \pi = 0.456$

$H_0: \pi < 0.456$

(d) Show how Douglas obtained the value of 0.456 for his hypothesis test.

(1)

$\frac{114}{(114+136)} = 0.456$

(e) Explain why the critical region for Douglas's test is $X \leq 10$, where X represents the number of goals scored from 35 penalty corners.

(2)

Because $P(X \leq 10) = 0.925$ so X being less than 10 will be at least 92.5% probability

$P(X \leq 10) = 0.925$

$\frac{7}{35} = 0.2$

$X \sim B(35, 0.2)$

(f) Complete Douglas's hypothesis test.

(2)

$P(X \leq 10) = 0.0302$ $0.286 > 0.0302 < 0.456$

X is in critical region $= 0.286$

TS = TS =

So reject H_0, there is significant evidence to suggest that a lower proportion of penalty corners that his team is awarded result in goals compared to teams in 1998 Field hockey world cup.

(g) Give one advantage of using a critical region instead of a p-value when carrying out hypothesis tests.

(1)

Critical regions can show where the test statistic lays, more accurate.

Q6

A B C

(h) Describe, in context, the meaning of a Type I error for Douglas's test.

(2)

A Type I error is where the ~~fact is~~ $H_0: \pi = 0.456$ is rejected when it is actually true.

(i) Write down the exact probability that Douglas makes a Type I error when carrying out his test.

(1)

~~0.8~~ P(Type I error) = 0.5

(j) If the probability that, for Douglas's team, a penalty corner results in a goal is actually 0.15, find the power of his test.

(3)

~~0.15~~ ~~7 x 35~~ = 5.25

~~P(X~~ $\frac{7}{35}$ = 0.2

35 x 0.15 = 5.25 = 1.75 difference

~~0.2 x 0.15 = 0.03~~

0.856 x 100 = 85.6%

N = 35

p = 0.15

Lower = -999999

Upper = 7

Q6

(k) State **two** necessary assumptions that you made when selecting the distribution to use in (e).

You should explain whether you believe **each** to be valid or not.

(4)

Assumed ~~normal~~ each goal was independent of one another, this is not valid as penalties may be caused by goals due to fouls.

Assume only two outcomes. Valid as a penalty results in a goal or no goal.

During the 1998 Field Hockey World Cup, specialist penalty corner players were brought on to substitute for other players when a penalty corner was awarded.

In 1999, a new rule was introduced that specialist penalty corner players could not be brought on, unless a player was injured.

(l) Explain how this information may affect the validity of the conclusion you made in (f).

(2)

Some goals may not be scored by injured players so there may be evidence to support Ho. Validity assumed all goals were scored by a player of choice.

Q6

A B C

6 /24

Part (a)

M1B0A0: The candidate has used the value for a 90% confidence interval rather than for a 95% confidence interval so only scores the first M mark for a correct standard error.

Part (b)

E0: There is no context in this answer so it cannot be awarded a mark.

Part (c)
B1E0: The candidate has correctly stated that the mean lies in the confidence interval however they have concluded incorrectly so have only been awarded the B mark.

Part (d)
B1: This answer was exactly as the mark scheme required.

Part (e)
E0E0: The candidate's response to this question uses the wrong binomial distribution and could not be awarded any marks.

Part (f)
M0E0: The candidate has confused the null hypothesis value with the significance level and although they have concluded correctly from here could not be awarded the mark.

Part (g)
E0: This was a common wrong answer to this question. Critical regions are no more accurate than p-values.

Part (h)
E1E0: There was no context in this answer so it could not be awarded the second mark.

Part (i)
A0: Only answers of 0.0302 or a correct follow through from a probability in (e) could be awarded a mark in this part of the question.

Part (j)
M0A0A0: There is no work of any merit in this part of the question.

Part (k)
E0E0E1E0: Two outcomes is a requirement of the binomial distribution but it could not be given as an answer in this part of the question as it is not an assumption: a penalty can either be scored or not. The third mark was awarded as the candidate was given the benefit of the doubt that they had identified that the penalties needed to be independent of each other however they have misunderstood the context and their reasoning was incorrect.

Part (l)
E1E0: This was considered to be an incomplete explanation about why the test is less valid. Ideally candidates would have stated that the conditions were not the same in each of the two situations being compared and this reduced the validity.

Question 6 - Response B

(a) Making any necessary assumptions, calculate a 95% confidence interval for the mean number of penalty corners awarded per game for Douglas's team.

(3)

$\mu = \bar{x} \pm z \times \frac{s}{\sqrt{n}}$

$3.25 \pm 1.959 \times \frac{1.2}{\sqrt{15}}$

$= (2.643, 3.857)$

(b) State the necessary distributional assumption required for the confidence interval found in (a) to be valid.

(1)

Sam's data was taken from a normally Distributed population.

(c) By making reference to the confidence interval found in (a), comment on the claim that, over the last three years, the number of penalty corners awarded to Douglas's team per game, on average, differs to that awarded to teams at the 1998 Field Hockey World Cup.

(2)

2.976 is inside the confidence Interval in (a) so I dont think the average number of penalty corners awarded has changed since 1998

Douglas believes that a **lower** proportion of penalty corners that his team is awarded result in goals compared to that for teams playing in the 1998 Field Hockey World Cup.

Last year, his team was awarded 35 penalty corners resulting in 7 goals.

Douglas decides to test the following hypotheses:

$H_0: \pi = 0.456$

$H_0: \pi < 0.456$

(d) Show how Douglas obtained the value of 0.456 for his hypothesis test.

(1)

114 + 136 = 250 penalty corners

114 ÷ 250 = 0.456

(e) Explain why the critical region for Douglas's test is $X \leq 10$, where X represents the number of goals scored from 35 penalty corners.

(2)

X ≤ 10 is the highest number at which the probability is >10% which I will assume is his significance level

P(x ≤ 10) = 0.146 ← above 10%
P(x ≤ 11) = 0.075

(f) Complete Douglas's hypothesis test.

(2)

0.146 < 0.2 ∴ reject H_0, significant evidence to show that the proportion of penalty corners to penalty corners scored is less than 0.456.

Q6

A
B
C

(g) Give one advantage of using a critical region instead of a *p*-value when carrying out hypothesis tests.

(1)

you can't have half a goal so when doing binomial you must have a Critical Region.

(h) Describe, in context, the meaning of a Type I error for Douglas's test.

(2)

Rejecting H_0 when H_0 is true

Saying the proportion of penalty corners scored is less than 0.456 when it is not.

(i) Write down the exact probability that Douglas makes a Type I error when carrying out his test.

(1)

14.6%

(j) If the probability that, for Douglas's team, a penalty corner results in a goal is actually 0.15, find the power of his test.

(3)

lower = 10
Upper = 99999
Num = 35
p = 0.15

$P(X \leq 10) = 0.029$

$1 - 0.029 = 0.971$

1 - P(Type 2 error) = Power

Q6

(k) State **two** necessary assumptions that you made when selecting the distribution to use in (e).

You should explain whether you believe **each** to be valid or not.

(4)

- data was taken from a random sample which i believe to be valid
- bivariate data, they either score or they dont so thats valid.

During the 1998 Field Hockey World Cup, specialist penalty corner players were brought on to substitute for other players when a penalty corner was awarded.

In 1999, a new rule was introduced that specialist penalty corner players could not be brought on, unless a player was injured.

(l) Explain how this information may affect the validity of the conclusion you made in (f).

(2)

It would make it invalid as it could mean each team has less chance of scoring from penalty corners now.

11 /24

Q6

A B C

Part (a)

M1B1A0: This answer was covered by the special case in the mark scheme. The candidate has incorrectly calculated a *z* confidence interval rather than a t confidence interval and was awarded the first 2 marks.

Part (b)

E0: There is no context in this explanation so it cannot be awarded a mark.

Part (c)

B1E1: Although no technical language was used for the conclusion here, it was considered that the candidate had shown sufficient "uncertainty" about their conclusion to award the E mark.

Part (d)
B1: This a better response than the student response C.

Part (e)
E0E0: Neither mark was awarded here as although the candidate had attempted to calculate both $P(X \leq 10)$ and $P(X \leq 11)$, the probabilities were wrong and so was the reasoning about why it was $X \leq 10$ it's "above 10%".

Part (f)
M0E0: The candidate, like many, has used their value from (e) to complete the hypothesis test showing a misunderstanding between the test statistic, 7, and the critical value, 10 and could not be awarded any marks.

Part (g)
E0: The candidate has not understood the repeatability of critical regions.

Part (h)
E1E1: This answer is essentially exactly what the mark scheme is looking for.

Part (i)
A1: Although this is incorrect, candidates were given marks if they correctly followed through from their value in (e).

Part (j)
M1A1A0: Although the candidate has stated $P(X \geq 10)$ they have clearly calculated $P(X \leq 10)$ and have the correct answer for this so they were awarded the first two marks as per the special case in the mark scheme.

Part (k)
E0E0E0E0: Although it is required that the sample is random, this is not an assumption in this scenario as we are told that it is random in the question so it cannot be awarded a mark.

Part (l)
E1E0: This was considered to be an incomplete explanation about why the test is less valid. Ideally candidates would have stated that the conditions were not the same in each of the two situations being compared and this reduced the validity.

Question 6 - Response C

(a) Making any necessary assumptions, calculate a 95% confidence interval for the mean number of penalty corners awarded per game for Douglas's team.

(3)

$\mu = \bar{x} \pm t_{n-1} \times \frac{s}{\sqrt{n}}$

$\mu = 3.25 \pm 2.145 \times \frac{1.2}{\sqrt{15}}$

$= (2.59, 3.91)$

(b) State the necessary distributional assumption required for the confidence interval found in (a) to be valid.

(1)

assumes data comes from a normal distribution

Question 6 continued

(c) By making reference to the confidence interval found in (a), comment on the claim that, over the last three years, the number of penalty corners awarded to Douglas's team per game, on average, differs to that awarded to teams at the 1998 Field Hockey World Cup.

(2)

$\mu = \bar{x} \pm 2 \times \frac{\sigma}{\sqrt{n}}$

$\mu = 2.976 \pm 1.96 \times \frac{0.98}{\sqrt{84}}$

$= (2.77, 3.19)$

so this is inside the confidence interval for douglas team and so is mean of 2.976 so insufficient evidence to show the number of penalty corners awarded to douglas team differ to that awarded to teams at 1998 field hockey world cup

Douglas believes that a **lower** proportion of penalty corners that his team is awarded result in goals compared to that for teams playing in the 1998 Field Hockey World Cup.

Last year, his team was awarded 35 penalty corners resulting in 7 goals.

Q6

A
B
C

Douglas decides to test the following hypotheses:

H_0: $\pi = 0.456$

H_0: $\pi < 0.456$

(d) Show how Douglas obtained the value of 0.456 for his hypothesis test. **(1)**

$\frac{114}{250} = 0.456$

(e) Explain why the critical region for Douglas's test is $X \leq 10$, where X represents the number of goals scored from 35 penalty corners. **(2)**

$X \sim B(35, 0.456)$

$P(X \leq ?) = < 0.05$

$= P(X \leq 10)$ as this $= 0.0302$

but if you do $X \leq 11 = 0.0636$
which isnt less than 0.05
so its ≤ 10

(f) Complete Douglas's hypothesis test. **(2)**

$H_0: p = 0.456$
$H_1: p < 0.456$

$0.00149 < 0.05$

$X \sim B(35, 0.456)$
$P(X \leq 7) = 0.00149$ so reject H0
sufficient evidence to show a lower proportion of penalty corners that his team is awarded results in goals compared to teams playing in 1998 field hockey world cup

(g) Give one advantage of using a critical region instead of a p-value when carrying out hypothesis tests. **(1)**

critical regions are more accurate [illegible]

Q6

A

B

C

(h) Describe, in context, the meaning of a Type I error for Douglas's test.

(2)

a type 1 error is where you reject H0 and you shouldn't have so in this case you are saying there is evidence of a lower proportion of penalty corners resulting in goal when this isn't true

(i) Write down the exact probability that Douglas makes a Type I error when carrying out his test.

(1)

0.05

(j) If the probability that, for Douglas's team, a penalty corner results in a goal is actually 0.15, find the power of his test.

(3)

$X \sim b(35, 0.15)$
$p(X \leq 10) = 0.989$

power = 0.989

(k) State **two** necessary assumptions that you made when selecting the distribution to use in (e).

You should explain whether you believe **each** to be valid or not.

(4)

~~because the number~~
assume number of trial is fixed which is true as it's the amount from 35 penalty corners and assume probability is constant which may not be true as if they score the previous one they may have more confidence so they are more likely to score the next one

Q6

(l) Explain how this information may affect the validity of the conclusion you made in (f).

(2)

because if they could sub players in the world cup they would obviously score more with specialist penalty corner players but if this wasnt allowed it may not be the case so the proportion for dougla team may not actually be lower

(Total Marks for Question 6 is 24 marks)

18 /24

Q6

Part (a)
M1B1A1: The candidate correctly identified this as a t distribution confidence interval and accurately carried it out.

Part (b)
E0: The candidate was not awarded the mark for this question as there is no context in their answer.

Part (c)
B1E1: Although the candidate had unnecessarily calculated another confidence interval, they have been given the benefit of the doubt that they had correctly identified that the mean is within the original confidence interval and concluded in a non-definite way. Many candidates lost the E mark here where they concluded very certainly.

Part (d)
B1: This is just about the bare minimum for this mark. This question requires candidates to "show that" 0.456 is the proportion and ideally should demonstrate each step otherwise they risk losing the mark. Here the ideal answer would show where 114 had come from.

Part (e)
E1E1: This is well answered by the candidate who has stated the probabilities for both $X \leq 10$ and $X \leq 11$

Part (f)
M1E1: The candidate has used a p-value method here which is perfectly ok for the marks.

Part (g)
E0: This was a common wrong answer to this question. Critical regions are no more accurate than p-values.

Part (h)
E1E1: The candidate has correctly defined a Type I error. It is worth noting that the context of this question was marked very leniently due to the double null hypothesis in the question.

Part (i)

A0: This was a common wrong answer to this question. The candidate has simply stated the significance level which would be correct for a continuous distribution.

Part (j)

M1A1A1: Typically candidates work out P(Type II error) and then find 1 – P(Type II error) to calculate power but this answer is correct with no incorrect working so the candidate was awarded all of the marks.

Part (k)

E1E1E0E0: The first mark here is awarded as the candidate has correctly identified that the probability needs to be assumed to be constant in order to carry out the hypothesis test. Their explanation about this assumption's validity is more about the independence of the penalties however this was considered sufficient for the second mark as the candidate has tied it to the probability. The comment about there being a fixed number of trials cannot be awarded a mark as this is not an assumption at this point, it is given as true in the question.

Part (l)

E1E0: This answer was considered to be a reasonable attempt to explain the change in validity of the hypothesis test. Ideally candidates would have stated that the conditions were not the same in each of the two situations being compared and this reduced the validity.