

DELEGATE BOOKLET TASKS

1ST0 – 1901

Question - Paper 1F (Q15)

15 A driving school has 40 cars.

The cars are either petrol cars or diesel cars.

The cars have either a manual gearbox or an automatic gearbox.

16 of the cars have an automatic gearbox.

10 of the petrol cars have an automatic gearbox.

There are 30 petrol cars.

One of the petrol cars is to be picked at random.

One of the diesel cars is to be picked at random.

Derek says,

“The probability that the petrol car has a manual gearbox is greater than the probability that the diesel car has a manual gearbox”.

Is he correct?

You must show working and justify your answer.

Mark Scheme - Paper 1F (Q15)

Question number	Answer	Additional guidance	Mark
15	<p>B1 for finding the total number of diesel cars (10) AND the number of manual petrol cars (20) OR for finding the number of manual petrol cars (20) AND the total number of manual cars (24) B1 for finding the number of manual diesel cars (4)</p> <p>M1 for finding the probability of a manual diesel car $\frac{14}{10}$ OR a manual petrol car $\frac{20}{30}$ A1 $\frac{4}{10}$ and $\frac{20}{30}$ oe depB1ft correct conclusion for their two probabilities</p>	<p>B1B1 may be scored in a table or frequency (tree) diagram. Values may be implied by relevant probabilities, e.g. manual diesel $\frac{4}{10}$ oe implies 4 manual diesel cars and 10 diesel cars; petrol manual $\frac{20}{30}$ oe implies 20 petrol manual cars Numbers alone are not sufficient – there must be an indication of class of car e.g. diesel 10, petrol manual 20</p> <p>Accept 0.66, 0.67 or better for $\frac{20}{30}$ Dependent on M1 scored.</p>	(5)

Student response 1

15 A driving school has 40 cars.

The cars are either petrol cars or diesel cars.

The cars have either a manual gearbox or an automatic gearbox.

16 of the cars have an automatic gearbox. — 24 manual
 10 of the petrol cars have an automatic gearbox. — 6 manual
 There are 30 petrol cars.

One of the petrol cars is to be picked at random.

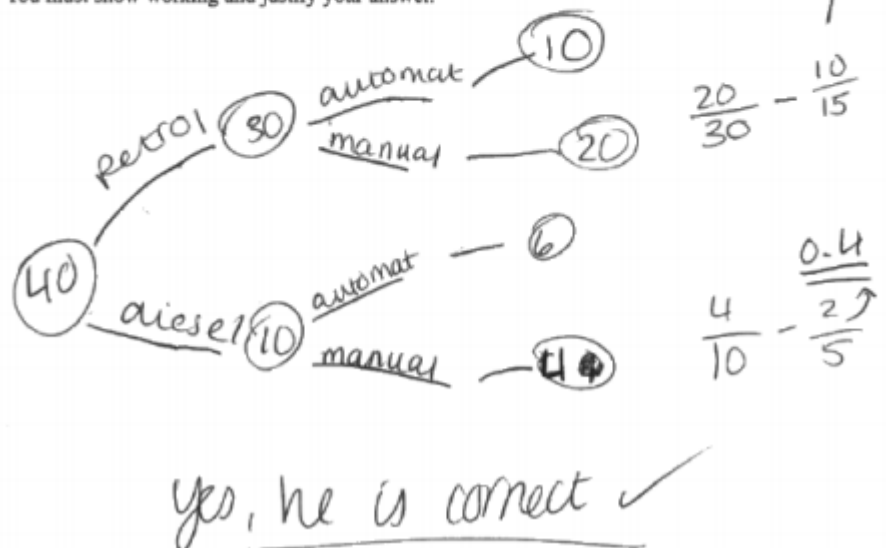
One of the diesel cars is to be picked at random.

Derek says,

"The probability that the petrol car has a manual gearbox is greater than the probability that the diesel car has a manual gearbox".

Is he correct?

You must show working and justify your answer.



Examiner comments

B1 B1 M1 A1 depB1ft

B1 for finding the total number of diesel cars (10) **and** the number of manual petrol cars (20).

B1 for finding the number of manual diesel cars (4).

M1 for finding the probability of a manual diesel car **or** a manual petrol car.

A1 for finding the probability of a manual diesel car ($\frac{4}{10}$) and a manual petrol car ($\frac{20}{30}$) correctly.

B1 for a correct conclusion for the two probabilities found.

Student response B

15 A driving school has 40 cars.

The cars are either petrol cars or diesel cars.

The cars have either a manual gearbox or an automatic gearbox.

16 of the cars have an automatic gearbox.

10 of the petrol cars have an automatic gearbox.

There are 30 petrol cars.

One of the petrol cars is to be picked at random.

One of the diesel cars is to be picked at random.

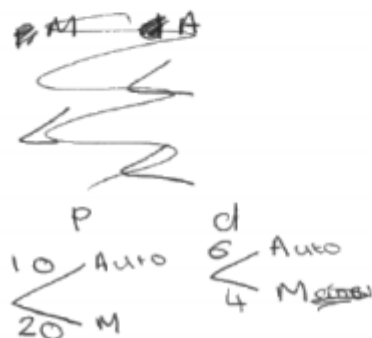
Derek says,

"The probability that the petrol car has a manual gearbox is greater than the probability that the diesel car has a manual gearbox".

Is he correct?

You must show working and justify your answer.

$\frac{20}{30}$ have a manual gearbox (petrol)
 $16 - 10 = 6$
 $\frac{6}{10}$ have a manual gearbox (diesel)
 he is correct.



Examiner comments

B1 B1 M1 A0 depB0ft

B1 for the total number of diesel cars (10) and the number of manual petrol cars (20). Implied by the correct denominator for the manual gearbox (diesel) probability and the correct numerator for the manual gearbox (petrol) probability.

B1 for finding the total number of manual diesel cars (4). Shown by 4 below 'd' on the manual branch of the frequency tree.

M1 for probability of a petrol car being a manual ($\frac{20}{30}$)

A0 did not find probability of a diesel car being a manual ($\frac{4}{10}$) (both $\frac{20}{30}$ and $\frac{4}{10}$ were required).

B0 conclusion based on incomplete working.

Question - Paper 2F (Q8b)

- 8 Mayokun measured and recorded the height, to the nearest cm, of each of the first 20 female students and of each of the first 20 male students to arrive at his college one morning.

He used statistical software to produce these diagrams and these summary statistics to help him compare the distributions.

Female students		Male students		Key: 13 4 represents 134 cm
13	4	14	6	
14	2 7 7	15	4 4 4 6 6 7 8 9	
15	1 1 2 6 9 9	16	2 4 4 5 8	
16	3 5 7 7 9 9	17	2 2 7	
17	1 4 8	18	2 5	
18	4	19	1	

	Median	Mean	Range	IQR
Females	161 cm	160 cm	50 cm	18 cm
Males	163 cm	165 cm	45 cm	16 cm

Mayokun chose to use stem and leaf diagrams rather than histograms.

- (a) Give one advantage of using stem and leaf diagrams rather than histograms for Mayokun's data.

(1)

Before collecting his data, Mayokun wrote down two hypotheses.

1. Males are taller than females.
 2. The heights of males vary more than the heights of females.
- (b) Using appropriate results from Mayokun's survey, discuss any conclusions that he might have made about his hypotheses.

You should comment on the reliability of the conclusions.

Mark Scheme - Paper 2F (Q8b)

Question number	Answer	Additional guidance	Mark
8 (a)	<p>B1</p> <ul style="list-style-type: none"> Any one from Can still see the original data values (in a stem & leaf) o.e. Can use to find (accurate) values for, e.g.: median/mode/mean/average/quartile(s)/range/I QR Can identify outliers 	<p>B1 for advantage of using a stem and leaf diagram Do not accept e.g. 'does not group data' to mean 'see original data' Condone 'it is more detailed/accurate' But do not accept, e.g. clearer / quicker or easier to plot/read/understand/interpret (all B0)</p>	(1)
(b)	<p>B1 B1 B1 B1 for four correct statements from</p> <ul style="list-style-type: none"> First hypothesis is supported / males are taller Males have a higher median (or mean) OR males are 2 cm (or 5 cm) taller on average, (Accept "163 > 161" or "165 > 160") Second hypothesis is not supported / female heights have greater spread Males have lower IQR (or lower range). (Accept 16 < 18 or 45 < 50) Median is more appropriate than mean (as males' data is positively skewed) <p>B1B1 any two comments from</p> <ul style="list-style-type: none"> conclusion(s) are not reliable (condone 'data' are not reliable) not a representative sample / only early arrivals small sample only (e.g. only/just used 20/40) quota sampling (or convenience sample) not random sample is only for college age / he didn't record age sample is for one area only / only his college 	<p>B1 for each of four statements from the options given, maximum 4 marks</p> <ul style="list-style-type: none"> statement supporting first hypothesis (condone 'is correct') correct supporting evidence (can ignore figs for 'median') (Comparison of e.g. tallest male/female alone is B0) statement refuting second hypothesis (condone 'incorrect') correct use of measure of dispersion. recognition of appropriate average due to skew <p>Note: for the first 4 marks it needs to be clear which hypothesis or gender their comment refers to.</p> <p>B1 for each of two statements from the options given, maximum 2 marks.</p> <p>Allow each bullet once only. Do not accept contradictory comments for any bullet point.</p>	(6)

Student response A

Before collecting his data, Mayokun wrote down two hypotheses.

1. Males are taller than females.
2. The heights of males vary more than the heights of females.

(b) Using appropriate results from Mayokun's survey, discuss any conclusions that he might have made about his hypotheses.

You should comment on the reliability of the conclusions.

- The median height for boys is 163cm which is 2cm taller than the girls median (161cm). This suggests that males are taller on average.
- The range for females is 50cm which is 5cm more than the range for males (45cm). This means that the heights of females vary more than the height of males. This means that Mayokun's hypothesis about height range is wrong.
- The data, however, may not be completely reliable, as Mayokun only asked the first 20 people of each gender to enter the college. This could mean that all of the first people to arrive were all in the same year as each other, so their heights would be similar.

Examiner comments	
(a)	B1 Clear answer indicating that the data is shown.
(b)	B1 B1 B1 B1 B1 B1 In the first three lines, the student made clear comparison of medians to reach the correct conclusion about the first hypothesis (bullets 1 and 2: B1 B1). In the next five lines, they made a clear comparison of ranges to reach the correct conclusion about the second hypothesis (bullets 3 and 4: B1 B1). The final explicit comment about the hypothesis was not required as they had already scored for the context statement about female heights varying more. In the final paragraph, they correctly stated that the data may not be reliable with the appropriate reasoning that it was a small sample (bullets 1 and 3: B1 B1).

Student response B

Before collecting his data, Mayokun wrote down two hypotheses.

1. Males are taller than females.
2. The heights of males vary more than the heights of females.

(b) Using appropriate results from Mayokun's survey, discuss any conclusions that he might have made about his hypotheses.

You should comment on the reliability of the conclusions.

Mayokun used Convenience sampling
So he just picked the first 20 females and males that arrived at his college in the morning. His hypothesis is correct. ~~to an extent~~. From looking at the table the median for males is higher than for females but the range for females is higher which says that females were more consistent than males.

Examiner comments	
(a)	B0 No correct reasoning given.
(b)	B1 B1 B0 B0 B1 B1 The first comment scored the two marks for considering reliability. Although the student did not state that the conclusions would be unreliable, they did recognise convenience sampling (bullet 4 from the second list) and the small sample size ('only the first 20 ...', bullet 3). The comment about his hypothesis being correct did not score, as the student did not state which hypothesis was meant. They then gave a correct comparison of medians (bullet 2 from the first list) and a correct comparison of ranges (bullet 4). Figures did not need to be stated for these comparisons. The comparison of medians was not interpreted in context to conclude for hypothesis 1. The student attempted to interpret the comparison of ranges for hypothesis 2 but did so incorrectly, stating that a higher range for females meant they were more consistent.

Question - Paper 2F (Q10)

10 At a university, 70% of students are undergraduates and 30% of students are postgraduates.

Amy and Robert want to do a survey.

Amy decides to use simple random sampling to collect a sample of 100 students.

She uses the university database as a sample frame and she numbers each student on the database.

She then generates exactly 100 random numbers and uses these random numbers to select her sample.

(a) Give **two** reasons why Amy's method may **not** produce a sample of 100 students.

.....

.....

.....

.....

(2)

Robert decides to use quota sampling to collect a sample of 100 students.

He plans to stand outside the main building until he has interviewed 70 undergraduates and 30 postgraduates.

(b) Give **two** advantages of using quota sampling.

.....

.....

.....

.....

(2)

(c) Explain why this quota sample is **not** a random sample.

.....

.....

(1)

Mark Scheme - Paper 2F (Q10)

Question number	Answer	Additional guidance	Mark
10 (a)	<p>B1B1 Any two from:</p> <ul style="list-style-type: none"> • Repeated random numbers • Random numbers out of range/may not correspond to students' numbers • Selected students may not (want to) participate • Some students may have left the university 	<p>B1 for each bullet point up to a maximum of 2</p> <p>Accept each bullet point only once</p> <p>Students may have joined the university is B0.</p> <p>There may not be 100 students at the university is B0.</p> <p>Random numbers may not be whole numbers is B0.</p> <p>Database may not be up to date on its own is B0.</p> <p>Do not accept (random) numbers may be more than 100 for the second bullet point.</p> <p>Ignore extraneous non-contradictory comments.</p>	(2)
(b)	<p>B1B1 Any two advantages from:</p> <ul style="list-style-type: none"> • Easy/convenient/quick/efficient/cheap • Represents population (proportions) • Allows for comparison (between undergraduates and postgraduates) • No sample frame required 	<p>B1 for each bullet point up to a maximum of 2</p> <p>Accept each bullet point only once</p> <p>For 2nd bullet point allow e.g. 'fair number of each (group)' 'Unbiased' on its own is B0.</p> <p>Ignore extraneous non-contradictory comments.</p>	(2)
(c)	<p>B1 Any one from:</p> <ul style="list-style-type: none"> • Not every student has an equal(o.e.) chance of being selected • Only those in the main building can be selected/not every student has a chance of being selected • Robert is choosing the students 	<p>B1 for a reason which states or implies 'equal likelihood' of being selected or that Robert is doing the choosing</p> <p>Do not allow 'even' chance or 'its biased' for the first bullet point, but condone 'fair chance'.</p>	(1)

Student response A

10 At a university, 70% of students are undergraduates and 30% of students are postgraduates.

Amy and Robert want to do a survey.

Amy decides to use simple random sampling to collect a sample of 100 students.

She uses the university database as a sample frame and she numbers each student on the database.

She then generates exactly 100 random numbers and uses these random numbers to select her sample.

(a) Give **two** reasons why Amy's method may **not** produce a sample of 100 students.

one reason is a number may be produced multiple times, some students may not be in university that day, some people may not want to be interviewed

Robert decides to use quota sampling to collect a sample of 100 students.

He plans to stand outside the main building until he has interviewed 70 undergraduates and 30 postgraduates.

(b) Give **two** advantages of using quota sampling.

quota sampling will give a more fair representation of both groups, quota sampling is easier as you don't have to produce random numbers

(c) Explain why this quota sample is **not** a random sample.

Robert will decide who to interview, and the number of under and post graduates

Examiner comments	
(a)	B1 B1 The first comment showed an understanding of repeated (random) numbers being an issue. The further comment about non-participation was acceptable for the second mark.
(b)	B1 B1 The student gave two valid reasons: quota sampling was representative of the groups (bullet 2), and that it was easier (bullet 1), in this case with an explanation of why, although this was not required.
(c)	B1 The student recognised that Robert was choosing who to ask (bullet 3).

Student response B

10 At a university, 70% of students are undergraduates and 30% of students are postgraduates.

Amy and Robert want to do a survey.

Amy decides to use simple random sampling to collect a sample of 100 students.

She uses the university database as a sample frame and she numbers each student on the database.

She then generates exactly 100 random numbers and uses these random numbers to select her sample.

(a) Give **two** reasons why Amy's method may **not** produce a sample of 100 students.

As some numbers may occur more than once meaning less than 100 different numbers selected. Some numbers may be more than 100

Robert decides to use quota sampling to collect a sample of 100 students.

He plans to stand outside the main building until he has interviewed 70 undergraduates and 30 postgraduates.

(b) Give **two** advantages of using quota sampling.

easier to do and more presiceted

quicker information

(c) Explain why this quota sample is **not** a random sample.

as there looking for 70% and 30% so its not a random sample as they choose who to ask

Examiner comments	
(a)	B1 B0 The first comment showed a clear understanding of repeated (random) numbers being an issue. The second comment about being more than 100 was not appropriate.
(b)	B1 B0 The two answers of 'easier' and 'quicker' were both bullet 1, so scored only 1 mark.
(c)	B1 The first comment did not answer the question, but the student went on to say that participants were chosen (bullet 3), and so scored the mark.

Question - Paper 1H (Q11a)

- 11** Some students at a school walk home and some students go home by bus.

The times taken by the students at the school to walk home have a mean of 25 minutes and a standard deviation of 6 minutes.

The times taken to walk home can be modelled by a normal distribution.

- (a) Shanaya says,

“More than 80% of the students who walk home take between 19 and 37 minutes”

Use statistical calculations to assess Shanaya’s conclusion.

Mark Scheme - Paper 1H (Q11a)

Question number	Answer	Additional guidance	Mark
11 (a)	<p>M1 for standardising $\frac{37 - 25}{6} (= 2)$ oe</p> <p>OR by inspection stating 37 minutes is 2 sd's above 25 minutes. [May be seen on a diagram]</p> <p>M1 for standardising $\frac{19 - 25}{6} (= -1)$ oe</p> <p>OR by inspection stating 19 minutes is 1 sd below 25 minutes. [May be seen on a diagram]</p> <p>M1 for either $0.95 \div 2 (= 0.475)$ or $0.68 \div 2 (= 0.34)$</p> <p>AI 0.815 or 81.5%</p> <p>depB1ft 0.815 > 0.8 so Shanaya is correct</p>	<p>M1 for working out 37 is 2 s.d. above the mean. Allow M1 for $25 + 2 \times 6 = 37$</p> <p>M1 for working out 19 is 1 s.d. below the mean. A Allow M1 for $25 - (1) \times 6 = 19$</p> <p>M1 for correct use of either awrt 0.67 or 0.68 or awrt 67% or 68% OR for the correct use of either 0.95 or 95% i.e. $95\% \div 2$ OR $67/8\% \div 2$</p> <p>A1 – for use of both 47.5% or 0.475 AND 34% or 0.34 AND adding them to achieve 81.5% or 0.815 Allow answers greater than 81% or less than and equal to 82%. Note: Value from calculator is 81.8%</p> <p>Sight of 0.815 or 81.5% with no working is M1M1A1</p> <p>Dependant on getting any one of the M marks. depB1ft for correct conclusion based on their evaluated probability between 0 and 100%. Follow through their value. Conclusion required for this mark.</p>	(5)

Student response A

11 Some students at a school walk home and some students go home by bus.

The times taken by the students at the school to walk home have a mean of 25 minutes and a standard deviation of 6 minutes.

The times taken to walk home can be modelled by a normal distribution.

(a) Shanaya says,

"More than 80% of the students who walk home take between 19 and 37 minutes"

Use statistical calculations to assess Shanaya's conclusion.

$$25 \pm 6 = 31 = 66.7\% \\ = 19$$

$$6 \times 2 = 12 \\ 25 \pm 12 = 37 = 98\% \\ = 13$$

$$\text{Conclusion } 66.7\% = 19 \leq x < 31 \\ \text{and } 98\% = 13 \leq x < 37$$

$$19 \leq x < 37$$

= 1 sd left and 2 sd right

$$(66.7/2) + (98/2) = 82.35$$

$\leq 82.35 > 80$ so Shanaya is correct.

(5)

Examiner comments	
(a)	<p>M1 M1 M1 A0 depB1ft</p> <p>The student found the area of -1 and +2 standard deviations. The use of 66.7% was allowed but 98% was incorrect. They failed to gain the accuracy mark as their percentage was incorrect. They interpreted their percentage correctly and had gained a method mark, so were awarded the dependent B1 follow-through mark.</p>

Student response B

11 Some students at a school walk home and some students go home by bus.

The times taken by the students at the school to walk home have a mean of 25 minutes and a standard deviation of 6 minutes.

The times taken to walk home can be modelled by a normal distribution.

(a) Shanaya says,

"More than 80% of the students who walk home take between 19 and 37 minutes"

Use statistical calculations to assess Shanaya's conclusion.

99% of people walking should take between 7 minutes and 43 minutes. 95% of people should be between 13 minutes and 37 minutes and 65% of people are between 19 and 31 minutes. This is because to work out the percentages you do $3 \times$ standard deviation \pm from mean. Her estimate is incorrect as ^{47.5}95% of people should have been able to walk home within 37 minutes when added to the mean but ^{32.5}65% of people should have been able to walk home within 19 minutes when minused from the mean.

Examiner comments	
(a)	<p>M1 M0 M1 A0 depB0ft</p> <p>95% implied the two standard deviations. 65% was incorrect as it should be 68%. The student found 47.5% from $95\% \div 2$. They did not find the correct percentage for the accuracy mark. They did not compare their percentage with 80%.</p>

Question - Paper 2H (Q5)

- 5 The figures below show the amount, in £ millions, of UK aid given to Ethiopia for each of the years 2009 to 2015

219 263 344 265 329 321 338

The table gives a summary of the amount, in £ millions, of UK aid given to Pakistan for the years 2009 to 2015

Mean	Standard deviation	Greatest amount
243.7	79.5	374

(Source: *dfid.gov.uk*)

Compare the amount of UK aid given to Ethiopia and the amount of UK aid given to Pakistan for the years 2009 to 2015

You may use $219^2 + 263^2 + 344^2 + 265^2 + 329^2 + 321^2 + 338^2 = 631217$

Mark Scheme - Paper 2H (Q5)

Question number	Answer	Additional guidance	Mark
5	<p>B1 (Mean =) $297 \sqrt{\frac{631217}{7} - 297^2}$</p> <p>M1</p> <p>A1 44.326...</p> <p>B1ft e.g. The mean for E(Ethiopia) > mean for P(akistan) or The s.d. for E(Ethiopia) < s.d. for P(akistan)</p> <p>B1ft On average, more aid/money(o.e.) is given to Ethiopia and the amount of aid/money(o.e.) given to Pakistan varies more.</p>	<p>B1 for correct mean (must be seen outside of s.d. formula)</p> <p>M1 for attempt at standard deviation (allow f.t. on their mean)</p> <p>A1 for awrt 44.3 (Ignore calculations of standardised scores)</p> <p>B1ft for a correct comparison of means or standard deviations (may be implied by one correct interpreted comparison of means or standard deviation) Do not allow this mark if it clearly comes from a comparison of greatest values.</p> <p>B1ft for correctly interpreted comparison of means and standard deviations</p>	(5)

Student response A

- 5 The figures below show the amount, in £ millions, of UK aid given to Ethiopia for each of the years 2009 to 2015

219 263 344 265 329 321 338

The table gives a summary of the amount, in £ millions, of UK aid given to Pakistan for the years 2009 to 2015

Mean	Standard deviation	Greatest amount
243.7	79.5	374

(Source: dfid.gov.uk)

Compare the amount of UK aid given to Ethiopia and the amount of UK aid given to Pakistan for the years 2009 to 2015

You may use $219^2 + 263^2 + 344^2 + 265^2 + 329^2 + 321^2 + 338^2 = 631217$

$$\text{Standardised Score Pakistan} = \frac{374 - 243.7}{79.5} = 1.29$$

$$\text{Mean Ethiopia} = 219 + 263 + 344 + 265 + 329 + 321 + 338 \div 7 = 297$$

$$s = \sqrt{\frac{631217}{7} - 297^2} = 44.3$$

$$\text{Greatest amount} = 344$$

$$SS \text{ Ethiopia} = \frac{344 - 297}{44.3} = 1.06$$

Uk have given more aid to Pakistan from 2009 to 2015 than Ethiopia.

Examiner comments

B1

Correct mean.

M1 A1

Correct method and answer.

B0ft

The student gave an incorrectly interpreted comparison of the means and did not compare the standard deviations.

B0ft

The student only gave one interpreted comparison and it was incorrect.

Student response B

- 5 The figures below show the amount, in £ millions, of UK aid given to Ethiopia for each of the years 2009 to 2015

219 263 344 265 329 321 338

The table gives a summary of the amount, in £ millions, of UK aid given to Pakistan for the years 2009 to 2015

Mean	Standard deviation	Greatest amount
243.7	79.5	374

(Source: *dfid.gov.uk*)

Compare the amount of UK aid given to Ethiopia and the amount of UK aid given to Pakistan for the years 2009 to 2015

You may use $219^2 + 263^2 + 344^2 + 265^2 + 329^2 + 321^2 + 338^2 = 631217$

$$\text{mean} = 297$$

$$219 + 263 + 344 + 265 + 329 + 321 + 338 = 2079$$

my pen ran out

$$2079 \div 7 = 297$$

$$\text{Greatest amount} = 344$$

$$\text{Standard deviation} = \sqrt{\frac{1}{8} \times (219 - 297 \dots)^2}$$

UK aid to Ethiopia = 106.1
 have a higher greatest amount, \leftarrow rounded to 1 d.p.
 but UK aid to Pakistan have a higher mean and standard deviation. UK gave roughly the same amount to each.

(Total for Question 5 is 5 marks)

Examiner comments

B1

Correct mean.

M0 A0

The student used the wrong formula for standard deviation.

B0ft

The student gave an incorrect statement about the mean 'Pakistan have a higher mean' and an incorrect follow-through comparison of the standard deviations.

B0ft

The student gave no correct contextual interpretations. Both would be required for this mark.

