GCE

# Statistics (MEI)

Advanced Subsidiary GCE **AS H132**

# Report on the Units

# June 2009

**H132/MS/R/09**

This report on the Examination provides information on the performance of candidates which it is hoped will be useful to teachers in their preparation of candidates for future examinations. It is intended to be constructive and informative and to promote better understanding of the syllabus content, of the operation of the scheme of assessment and of the application of assessment criteria.

Reports should be read in conjunction with the published question papers and mark schemes for the Examination.

OCR will not enter into any discussion or correspondence in connection with this Report.



Any enquiries about publications should be addressed to:

OCR Publications
PO Box 5050
Annesley
NOTTINGHAM
NG15 0DL

Telephone: 0870 770 6622
Facsimile: 01223 552610
E-mail: publications@ocr.org.uk

## CONTENTS

## Advanced Subsidiary GCE Statistics (H132)

## REPORTS ON THE UNITS

# G241 Statistics 1

### General Comments

The level of difficulty of the paper appeared to be entirely appropriate for the candidates. The more able candidates scored heavily on all questions and the weaker candidates often picked up some marks on all questions with question 7 on probability contributing significantly to their total.

Most candidates supported their numerical answers with appropriate explanations and working although some rounding errors were noted. The possible exception was in question 8 where the procedure for distinguishing between hypotheses did not always include specific comparisons with 10% and where the construction of the critical region was often sketchy. There was a surprising inability to use the given numerical data in question 3 to find the standard deviation.

Weaker candidates often scored a significant proportion of their marks from question 1, the first three parts of the probability question (question 7) and from the initial parts of question 8. Amongst lower scoring candidates, there was evidence of the use of point probabilities in question 8. Also in question 8, many candidates are still not meeting the requirement to define p in words.

There seemed to be no trouble in completing the paper within the time allowed and no obvious misinterpretations of the rubric, although a very small number of candidates ignored the instruction to use graph paper for the histogram. It would also be very helpful if candidates could write down the question numbers on the front of the question paper.

### Comments on Individual Questions

1) (i) The mode was usually correct, and most candidates also found the median correctly. However some candidates quoted locations rather than actual values and others thought that the median was 1 or 1.5. There were occasional errors such as thinking that there was a total of 180 (using $\sum fx$) rather than 102 cars in the survey. Some weaker candidates found the mean instead of the median.

   (ii) Most line diagrams were correct although a small number joined the lines in one manner or another. Some others forgot to label at least one of their axes.

   (iii) The majority identified the positive skewness of the distribution, but a significant number of candidates thought that the skewness was negative.

2) (i) Many totally correct answers were seen although candidates occasionally tried to use permutations.

   (ii) This part was rather less well answered, although a good number of fully correct answers were seen. The most common error was the use of addition instead of multiplication giving ${}^{14}C_3 + {}^{11}C_2$ and this occurred very frequently.

3) (i) Almost all candidates found the mean, but a large number of candidates did not know the formula for finding the standard deviation. Those who knew how to find $S_{xx}$ usually went on to complete part (i) successfully. However there were many incorrect attempts at $S_{xx}$ with common variations including $1582 – 10.5^2$ or $1582 – 12×10.5$. Others gave the standard deviation as 1582/11 or $\sqrt{(1582/11)}$ and some had no idea what to do with the

numbers they were given. Rather fewer candidates than in recent sessions divided by 12 rather than 11 and found the RMSD.

(ii) Almost all showed that Dwayne's monthly earnings were £1550. However the majority of candidates did not realize that all they needed to do was to multiply their standard deviation by 100, but instead tried to start again in finding the new standard deviation, almost always without success. Of those that did multiply by 100, a few then could not also resist the addition of 500.

(iii) The explanation regarding the means was usually correct but that for the standard deviations was either ignored or candidates failed to explain in context. A lack of context in explaining the means was condoned, but not in the case of the standard deviations.

4) (i) Almost all candidates correctly explained why $E(X) = 25$, although hardly any commented on the symmetry of the distribution, but instead calculated $\Sigma xp(x)$.

(ii) Very many correct answers were seen. Some candidates just found $E(X^2)$, failing to subtract $25^2$ to find the variance, and occasionally candidates found the correct answer but then went on to do further calculations. Several candidates tried to work out $10 \times 0.2^2 + 20 \times 0.3^2 + \ldots$, ie squaring the probabilities rather than the $x$ values.

5) (i) Although a number of fully correct histograms were seen, there were also many errors. Candidates should always draw a table to show the frequency densities even if such a table is not specifically asked for in the question. Common errors included a simple frequency diagram, frequency ÷ midpoint, frequency × class width, vertical axis not labelled correctly, 3.07 plotted as 3.7 and more rarely 0.665 plotted as 0.0665. The label on the vertical axis of the histogram was not always in agreement with the bars drawn; for example bars drawn at 360, 400, 153.5 and 33.25 were described as frequency density rather than frequency per 50 miles or sometimes as both. A horizontal scale consisting of inequalities was another common error.

(ii) In estimating the median, many candidates identified that the median was the 600th value or 600½ th value and then identified the correct interval from 50 to 100 but usually gave an answer of 75 rather than attempting any interpolation. Some got as far as 30 but then forgot to add on the 50. In many centres not a single candidate attempted interpolation, suggesting that this is a topic which centres should pay attention to. A small number decided to estimate a mean distance instead.

6) (i) (A) Marks scored on this question were surprisingly low. Errors of 0.36 (plants with one problem only) or 0.53 were very frequent in this part.

(i) (B) The correct answer of 0.13 was frequently seen. There was also a wide variety of incorrect answers, perhaps 0.17 being the most common of these.

(ii) Many candidates (including a significant number of very high scoring candidates) treated part ii) as if it were "with replacement" giving an answer of $0.53^3$. Another less but fairly frequent wrong answer was $1 - 0.47^3$. A small number interpreted it as being a binomial distribution of 100 trials.

7) (i) Almost all candidates answered this correctly.

(ii) Most candidates answered this correctly but some candidates chose to find P(delayed) first, meaning that lengthy calculations were needed.

(iii) Once again this was usually answered correctly.

(iv) Many candidates struggled with the conditional probability, making a variety of errors, including (0.329 × 0.388)/0.388, 0.388/0.329 or just 0.329.

(v) Many candidates attempted to use a conditional probability approach to this part, but then the majority of these gave answers such as 0.388/0.8, 0.176/0.8 or 0.235/0.8, rather than the correct 0.188/0.8 = 0.235. A good proportion of candidates calculated just the numerator (0.188) or miscalculated it as 0.176 (missing the triplet 0.8 × 0.15 × 0.1). Very few realized the direct methods available such as 1 – 0.9 × 0.85 = 0.235.

(vi) This was very well answered although candidates usually rounded their answer to 43. On this occasion this error was not penalized. A few candidates miscalculated 110 × 0.388 as 38.8 rather than 42.68.

8) (i) (A) This was usually answered correctly either by calculation or tables, with direct calculation being the more popular method..
(B) Again this was usually answered correctly, but some candidates made things difficult for themselves by calculating point probabilities and then either forgetting P(0) or including P(3) and with varying degrees of accuracy. Some used tables incorrectly finding $1 - P(X \leq 3)$, rather than $1 - P(X \leq 2)$.
(C) Once again this was usually correct but occasionally the mode was found rather than the expected number.

(ii) Many candidates correctly stated their hypotheses in symbolic form. However, much use of incorrect notation was also seen. The required notation is clearly given in the mark scheme and candidates should be trained to use this, leading to a straightforward two marks. Many candidates still do not realise the need to define the parameter ‘p’ and thus they lose a third mark, even if they have stated their hypotheses correctly. The reason for the form of the alternative hypothesis was not always well explained in context.

(iii) Some Centres do not seem to have taught how to find a critical region and candidates from such Centres often ignored the request for the critical region and went straight to the hypothesis test. Of those who did try to find the critical region, many made errors, including omission of probabilities, failure to compare the probabilities with 10%, confusion between $P(X \geq n)$ and $P(X > n)$, and even in a surprisingly large number of cases an attempt to do a two–tailed test despite having stated the correct alternative hypothesis. There are still a considerable number of candidates who attempt to use point probabilities for a hypothesis test. Although it is given in the mark scheme, it is worth repeating here the recommended method for comparing the probabilities with the significance level. Candidates should find the two upper tail (in this case) cumulative probabilities which straddle the significance level.
$P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.8358$ or $0.1642 > 10\%$
$P(X \geq 6) = 1 - P(X \leq 5) = 1 - 0.9389$ or $0.0611 < 10\%$
Irrespective of whether their critical region was correct, many candidates declined to use that information, but instead started again with $P(X \geq 7) = 0.0181 < 10\%$ and tackled the hypothesis test by that method. Those who did use their critical region sometimes did not make it clear that ‘7 lies in the critical region’. Candidates should also be advised that it is necessary not only to make a decision but give a conclusion in context.

# G242 Statistics 2

The fourth sitting of this AS Statistics module has seen another increase in the number of candidates yet the total entry remains small. The overall standard of entry was encouraging; a broad range of marks were seen, with the majority of candidates being suitably prepared for the examination.
It is encouraging to see that many candidates are now comfortably applying appropriate techniques and can present solutions using suitable phrases. However, in questions requiring candidates to choose a technique, or in less routine questions, only the better prepared succeeded. Candidates would benefit from familiarising themselves with the differences between the statistical distributions/tests used and when to apply them. Candidates must have a good understanding of the differences between one-tailed and two-tailed hypothesis tests and the corresponding differences when looking up critical values in tables. Currently, too many candidates seem unsure how to identify the correct number of degrees of freedom when needed, and get mixed up by allowing for degrees of freedom when not needed.

**Question 1 Chi-squared test for Association**

In part (i), some candidates mixed up the hypotheses - this sort of mistake is expensive, as any resulting conclusions are usually contradictory. It is also a requirement that the hypotheses are written 'in context' - simply stating 'no association/association' is not enough.

In part (ii), most candidates scored well. Mistakes in calculating expected frequencies were surprisingly common. Attempts at calculating 'contributions' were on the whole more successful. Most candidates managed to correctly identify the number of degrees of freedom and the corresponding critical value, then carry on to make a sensible comparison (using either an inequality or sketch) and conclusion. Again, context must be included in the conclusion - simply stating 'significant' is not enough for the final mark.

Part (iii) was poorly answered. The requirement was for candidates to look at the table of contributions and identify either large contributions (indicating evidence of association) or contributions close to zero (indicating evidence of independence). In cases where contributions are large, candidates should also indicate whether there was a greater or smaller observed number of grapes than expected.

**Question 2 Normal distribution and confidence interval**

Part (i)(A) was well answered - some candidates calculated the probability for the 'other tail'.
Part (i)(B) was poorly answered - many recognised the binomial nature of the situation but were unsure what calculation was needed.
Part (ii)(A) was well answered.
Part (ii)(B) poorly answered - frequent mistakes included use of an incorrect z value (not 1.96), and use of variance in place of standard deviation.
In part (ii)(C), candidates were expected to point out that the (correct) interval calculated earlier did not contain 25kg and realise that this was an indication that the true value of the mean was perhaps greater than 25kg.

**Question 3 Wilcoxon test**

Part (i) was quite well answered but common errors were seen. These include omitting 'population' when stating hypotheses, ranking actual (not absolute value of) differences, adding the differences instead of the ranks, and using '$n$ - 1' when looking up the critical value. Despite this, many candidates scored well on this question.

In part (ii), most recognised that a t test was suitable, quoting the fact that it was a small sample, but few also mentioned that the population variance being unknown was relevant too.

**Question 4 Chi-Squared test for goodness of fit**

This proved to be a difficult question.
In part (i)(A), most candidates realised that evidence of calculation was necessary and were on the whole successful. A surprising number stated that 'my calculator says that the mean is 2.6' or words to that effect, and scored no marks.
In part (i)(B) many candidates compared the mean with the standard deviation, rather than with the variance, and achieved no credit. In such questions candidates are required to calculate the variance (using the given sample standard deviation) to make their comparison clear.
Part (i)(C) proved to be difficult for all but the most able; most candidates seemed to miss the instruction to look at the 'pattern of observed frequencies' and made general comments about 'independence'
In part (ii), many seemed unsure how to calculate the missing expected frequencies - attempts at $X = 5$ were more successful than for $X = 6$ or more. Few candidates identified the correct number of degrees of freedom, not realising the need to allow for a restriction due to the mean being estimated from the sample.

**Question 5 t test**

In part (i) most provided a correct estimate for the population mean and variance - mistakes with variance were more common.
In part (ii) many good answers were seen. It was pleasing to see $\mu$ identified as representing the population mean in many scripts. Attempts at finding the test statistic were good, although it was not unusual to see variance used in place of standard deviation. '$n - 1$' seems to be the 'default setting' for looking up numbers in tables, for many candidates, and so in this case most were successful in identifying the correct critical value.
In part (iii) most candidates made sensible comments.

# G243 Statistics 3

Q1 (i) Only a few candidates gave their hypotheses in symbols. The use of the customary statistical notation in terms of $\mu$ for a population mean is likely to lead to correct statements of hypotheses, and is to be commended. Nevertheless, candidates may state their hypotheses in words without any symbols at all, but there is an absolute insistence that the word "population" appears. Many candidates went on to perform a fully correct hypothesis test. However some incorrectly performed a two sample t test. Others found the test statistic correctly but then quoted a wrong critical value (usually 2.262), thus losing several marks. Most candidates knew that normality was a required assumption, but none mentioned ‘normality of the population of differences’.

(ii) Many candidates were able to explain why the student should not measure times taken in the same order for all 10 friends due to the learning effect which cannot be disentangled in the analysis.

Q2 (i) Many candidates were able to give suitable reasons, such as changes in products or in the economy.
(ii) Many candidates produced fully correct solutions. There were few candidates who did not try to rank the data other than those who thought that a t test was required. Others made an attempt at a paired sample Wilcoxon test, despite the sample sizes being different! A few candidates thought that as the test statistic was less than the critical value, they should accept rather than reject the null hypothesis.
(iii) Very few candidates were able to describe clearly how to use random number tables to choose a random sample of 8 stores. Hardly any mentioned that repeats should be ignored, and none mentioned the need for a random starting position in the table. Putting names into a hat and selecting 8 of them was as popular a method as the use of random number tables.

Q3 (i) Almost all candidates found the mean, but only a small number found the standard deviation correctly, despite this being part of the Specification for Statistics 1 (G241). A number found the variance but forgot to take the square root and others just found the value of $S_{xx}$.
(ii) Most candidates correctly explained that the sample is large.
(iii) The majority of candidates made a reasonable attempt at the test statistic, but often made errors such as using $s$ instead of $s^2$ for either males, or females, or both, in calculating the pooled standard deviation. The hypotheses were often correctly given, either in symbols, or in words, and unlike in question 1, most candidates did mention ‘population’ when giving hypotheses in words. The critical value was often correct, and most candidates who got this far then completed the test correctly.

Q4 (i) Most candidates completed the scatter diagram correctly.
(ii) Only a few candidates gave their hypotheses correctly in terms of $\rho$, and hardly any actually defined $\rho$. Candidates were more successful in carrying out the test, although trying to compare -0.5711 with a positive critical value was a fairly common error.
(iii) Many candidates noted either that the scatter diagram was not elliptical, or that there was an outlier. However many also discussed the ‘linearity’ of the scatter diagram, which is not really relevant to the question.
(iv) Many candidates found Spearman’s rank correlation coefficient correctly, although a number did not rank the results.
(v) Most candidates did not define their hypotheses correctly, describing correlation rather than association. A number however went on to correctly complete the test.
(vi) Candidates often correctly mentioned that the two tests had different results, but they rarely discussed the reasons for this.
(vii) Most candidates suggested sensible factors such as age and exposure to loud music.

(viii) Some candidates explained why simple random sampling would not be appropriate and rather more were able to suggest stratified sampling as an appropriate method. Systematic sampling and cluster sampling were popular wrong answers.

# Grade Thresholds

**Advanced GCE Statistics MEI (H132)**
**June 2009 Examination Series**

**Unit Threshold Marks**

| Unit | | Maximum Mark | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|---|
| **G241** | Raw | 72 | 60 | 53 | 46 | 40 | 34 | 0 |
| | UMS | 100 | 80 | 70 | 60 | 50 | 40 | 0 |
| **G242** | Raw | 72 | 56 | 48 | 41 | 34 | 27 | 0 |
| | UMS | 100 | 80 | 70 | 60 | 50 | 40 | 0 |
| **G243** | Raw | 72 | 52 | 45 | 38 | 32 | 26 | 0 |
| | UMS | 100 | 80 | 70 | 60 | 50 | 40 | 0 |

**Specification Aggregation Results**

Overall threshold marks in UMS (ie after conversion of raw marks to uniform marks)

| | Maximum Mark | A | B | C | D | E | U |
|---|---|---|---|---|---|---|---|
| **H132** | 300 | 240 | 210 | 180 | 150 | 120 | 0 |

The cumulative percentage of candidates awarded each grade was as follows:

| | A | B | C | D | E | U | Total Number of Candidates |
|---|---|---|---|---|---|---|---|
| **H132** | 8.5 | 23.4 | 36.2 | 61.7 | 78.7 | 100 | 48 |

For a description of how UMS marks are calculated see:
http://www.ocr.org.uk/learners/ums_results.html

Statistics are correct at the time of publication.

**OCR (Oxford Cambridge and RSA Examinations)**
**1 Hills Road**
**Cambridge**
**CB1 2EU**

**OCR Customer Contact Centre**

**14 – 19 Qualifications (General)**
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

**www.ocr.org.uk**

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored

**Oxford Cambridge and RSA Examinations**
**is a Company Limited by Guarantee**
**Registered in England**
**Registered Office; 1 Hills Road, Cambridge, CB1 2EU**
**Registered Company Number: 3484466**
**OCR is an exempt Charity**

**OCR (Oxford Cambridge and RSA Examinations)**
**Head office**
**Telephone: 01223 552552**
**Facsimile: 01223 552553**