Course report 2025

Advanced Higher Statistics

This report provides information on candidates' performance. Teachers, lecturers and assessors may find it useful when preparing candidates for future assessment. The report is intended to be constructive and informative, and to promote better understanding. You should read the report with the published assessment documents and marking instructions.

We compiled the statistics in this report before we completed the 2025 appeals process.

Grade boundary and statistical information

Statistical information: update on courses

Number of resulted entries in 2024: 150

Number of resulted entries in 2025: 176

Statistical information: performance of candidates

Distribution of course awards including minimum mark to achieve each grade

Course award	Number of candidates	Percentage	Cumulative percentage	Minimum mark required
A	81	46.0	46.0	82
B	32	18.2	64.2	69
C	22	12.5	76.7	56
D	14	8.0	84.7	43
No award	27	15.3	100%	Not applicable

We have not applied rounding to these statistics.

You can read the general commentary on grade boundaries in the appendix.

In this report:

- ‘most’ means greater than or equal to 70%
- ‘many’ means 50% to 69%
- ‘some’ means 25% to 49%
- ‘a few’ means less than 25%

You can find statistical reports on the statistics and information page of our website.

Section 1: comments on the assessment

Question paper 1

Question paper 1 proved more demanding than expected.

The level of demand in questions 1(a), 1(d), 1(f)(iii), 2(c), and 2(e) was higher than intended. We took this into account when setting the grade boundaries.

Question paper 2

Question paper 2 performed as expected.

Section 2: comments on candidate performance

Question paper 1

Question 1(a)

Some candidates did not achieve any marks, mainly because they referred to the filler sounds themselves instead of the occurrence of filler sounds. Some candidates did not precisely describe the random variable under consideration and within context. Refer to notes 2 and 3 in the marking instructions.

Question 1(b)

Many candidates gained the 1 mark available.

Question 1(c)

Many candidates did not achieve any marks because they attempted to calculate the mean rate by doing $\frac{0+3+6+6+5+5+2+1}{8}=3.5$. These candidates did not process the information from the first two columns of Table 1, labelled x_i and O_i, correctly.

Question 1(d)

Many candidates did not achieve any marks because they used $\mathrm{P}(X=5)$ rather than $\mathrm{P}(X \geq 5)$, or they incorrectly stated $\mathrm{P}(X \geq 5)=1-\mathrm{P}(X<4)$ instead of $\mathrm{P}(X \geq 5)=1-\mathrm{P}(X \leq 4)$.

Question 1(e)

Many candidates did not achieve any marks because they did not count the estimation of the parameter value of 3.5 from the observed frequencies as an extra constraint. Only a few candidates gave well-written justifications for the stated number of degrees of freedom.

Question 1(f)(i)

Many candidates gained the 1 mark available.

Question 1(f)(ii)

Many candidates gained the 2 marks available. The commonly observed responses in the marking instructions outline all the accepted test statistic calculations that depended on the use of different values from earlier calculations.

Question 1(f)(iii)

Many candidates gained at least 2 of the 4 marks available.

Some candidates missed out on marks for stating incomplete hypotheses. Some candidates did not include the Poisson distribution parameter value of 3.5 in both H_0 and H_1. Some candidates did not phrase the conclusion to the test in terms of H_1. Refer to mark 16 of the illustrative scheme and note 7 in the marking instructions.

Question 1(g)

Most candidates did not achieve any marks. Most candidates did not state that the common assumption was that the populations were normally distributed.

Question 2(a)

Many candidates gained the 1 mark available.

Question 2(b)

Most candidates gained at least 1 of the 2 marks available. Many candidates did not gain mark 3 because they stated, ‘sprint times and hurdle times have a correlation’, which excluded the possibility that the correlation coefficient could be zero. Refer to notes 3 and 4 in the marking instructions.

Question 2(c)

Many candidates did not achieve any marks because they did not clearly comment on the validity of the two assumptions. Refer to marks 4 and 5 of the illustrative scheme in the marking instructions and the Understanding Standards document ‘Residual Plots Guidance’.

Question 2(d)

Most candidates gained at least 1 of the 2 marks available.

Question 2(e)

Many candidates did not achieve any marks.

Question 2(f)

Most candidates did not achieve any marks because they used unacceptable phrases to describe the meaning of a prediction interval. Refer to note 3 in the marking instructions.

Question 2(g)

Many candidates gained at least 1 of the 2 marks available.

Question 2(h)

Many candidates did not achieve any marks. Many candidates did not explain why it was not a concern that the individual value was not captured by the confidence interval.

Question paper 2

Question 1(a)

Most candidates gained all 3 marks.

Question 1(b)

Many candidates gained at least 3 of the 4 marks available. A few candidates included the incorrect formula $\text{SD}(X-Y)=\text{SD}(X)+\text{SD}(Y)$ in their solutions. Also, refer to note 2 in the marking instructions for assumptions that were not accepted.

Question 2(a)

Most candidates gained at least 2 of the 3 marks available. A few candidates stated that because the lower fence was negative, it was actually zero, which is not correct. A few candidates extended their response to recommend removing outliers from the data set, which is not correct.

Question 2(b)

Most candidates gained at least 5 of the 6 marks available.

Question 3

Most candidates gained at least 3 of the 4 marks available. A few candidates did not name random variables appropriately, for example:

$\mathrm{P}(X > 20) \approx \mathrm{P}\left(X > \frac{20.5 - 14}{\sqrt{14}}\right) = \mathrm{P}(Z > 1.74)$ rather than the correct

$\mathrm{P}(X > 20) \approx \mathrm{P}\left(Z > \frac{20.5 - 14}{\sqrt{14}}\right) = \mathrm{P}(Z > 1.74)$.

Markers treated this as bad form.

Question 4(a)

Most candidates did not achieve any marks. Many candidates incorrectly stated, 'The variance is greater than the mean.' Most candidates who gained the mark structured their solution around a diagram of a normal distribution that they had appropriately labelled and scaled, with $0.00 clearly marked on it.

Question 4(b)(i)

Most candidates gained at least 2 of the 3 marks available. Some candidates omitted either the bar above the $\bar{X}$, or the communication of approximate normality, or both. A few candidates incorrectly applied continuity correction. A few candidates' notational usage was incorrect. Instead of writing $\mathrm{P}(Z > 2.0203) = 0.0217$, they wrote $2.0203 = 0.0217$.

Question 4(b)(ii)

Most candidates gained at least 1 of the 2 marks available.

Question 5

Most candidates gained at least 5 of the 6 marks available. Some candidates incorrectly used a critical value of 5, from looking up a 5% one-tailed test from page 15 of the Statistical Formulae and Tables.

Question 6

Most candidates gained at least 4 of the 5 marks available. Some candidates used μ_1 and μ_2, but did not clearly communicate which time period each of μ_1 and μ_2 referred to. Refer to note 2 in the marking instructions.

A few candidates stated that –4.2387 was further away from 0 than 2.58 (which was not acceptable logic for a one-tailed test) rather than $-4.2387 < -2.58$ or $|-4.2387| > 2.58$, or similar (which ensures that the sign of the test statistic is the same as the sign of the critical value and correctly uses an inequality symbol).

Question 7(a)

Many candidates gained at least 2 of the 3 marks available. Refer to notes 3 and 4 in the marking instructions for acceptable justification statements.

Question 7(b)

Many candidates did not achieve any marks. Some candidates stated an incorrect distribution and then attempted to justify it.

Candidates who gained marks described the contextual features, mentioned that the length of the leftover string could be any length from 0 to 8.0 centimetres, and stated that all lengths were equally likely. They then concluded that it would be the continuous uniform distribution, $\mathrm{U}(0, 8.0)$.

Question 8(a)

Many candidates gained at least 2 of the 3 marks available.

Question 8(b)

Many candidates gained at least 7 of the 9 marks available.

Some candidates used p_1 and p_2, when p_A and p_B would have been more appropriate for the context.

Note 1 in the marking instructions describes an alternative way that some candidates gained marks 1 and 2.

Question 8(c)

Many candidates gained the 1 mark available.

Question 9(a)

Most candidates gained the 2 marks available.

Question 9(b)

Few candidates gained at least 3 of the 4 marks available. Most candidates gave the commonly observed responses labelled as ‘Candidate B’ and ‘Candidate C’ in the marking instructions.

Question 10(a)

Many candidates gained at least 5 of the 7 marks available.

Some candidates incorrectly stated that spoonfuls were independent (when, in fact, masses of spoonfuls were independent).

Question 10(b)

Most candidates gained the 1 mark available.

Question 11(a)(i)

Many candidates gained all 3 marks. Some candidates set up the equation $0.821 = 0.92 - 2\sqrt{\frac{0.92 \times 0.08}{n}}$, solved for n and then used $\sqrt{\frac{0.92 \times 0.08}{30}}$ for the calculation of the two required limits. This was a valid approach, but longer than the expected strategy shown in the illustrative scheme in the marking instructions.

Question 11(a)(ii)

Many candidates gained the 1 mark available.

Question 11(b)

Many candidates gained the 2 marks available.

Question 12(a)

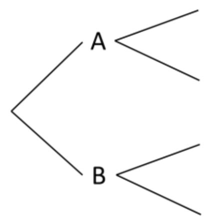
Many candidates gained at least 3 of the 5 marks available. Note 1 in the marking instructions contains example responses for the two assumptions required.

Question 12(b)

Many candidates gained the 2 marks available.

Question 13(a)

Some candidates gained the 2 marks available. Some candidates incorrectly labelled their tree diagram, as shown below:



Note 1 in the marking instructions shows a correctly presented tree diagram.

Question 13(b)

Some candidates gained the 2 marks available.

Question 13(c)

Many candidates gained all 3 marks. Many candidates used the appropriate strategy in their responses to parts (a) and (b).

Section 3: preparing candidates for future assessment

Candidates' solutions should be legible and have a clear logical structure.

Candidates should be able to:

- Use the correct notation for random variables — this includes knowing when to use capital letters and when to use lower case letters and when to change to Z (a random variable with a standard normal distribution).
- Pay very close attention to what a test is measuring when they are writing contextual assumptions. Refer to the commentary in section 2 about paper 1 question 1(a) and paper 2 question 10(a).

Question paper 1 and question paper 2

The comments in the previous sections and those below can help teachers and lecturers to prepare future candidates for the Advanced Higher Statistics question papers.

Teachers and lecturers should:

- Ensure candidates know that values of fences can be positive or negative. Fence values are not data points themselves.
- Ensure candidates understand how to approach outliers. Candidates should not automatically delete outliers; they should view outliers as data points that require further investigation.
- Ensure candidates know how to calculate a mean from a frequency table.
- Ensure candidates practise constructing different diagrams (for example tree and Venn) for solving the same probability question. This helps them understand which diagram might be the most appropriate to use, and why.
- Encourage candidates to draw annotated diagrams of distributions.

- Ensure candidates practise:
  - defining an appropriate modelling distribution for a variety of situations
  - clearly describing the underlying assumptions of that distribution
  - clearly describing the reason or reasons for the choice of parameter values
- Ensure candidates know that the null and alternative hypotheses for a chi-squared goodness-of-fit test should include the parameter value for the distribution being tested.
- Ensure candidates know how to calculate the required number of degrees of freedom for a chi-squared goodness-of-fit test, when the parameter of the modelling distribution is calculated from observed data.
- Ensure candidates know that the central limit theorem describes the approximate distribution of the sample mean, $\bar{X}$, and that they can write it down using the correct notation.
- Encourage candidates to check that the parameter subscripts are logically connected to the data they are investigating when conducting two-sample hypothesis tests. Encourage candidates not to use the subscripts of '1' and '2' in every instance.
- Ensure candidates understand that a two-sample z-test requires normal distributions and the population variances to be known. A two-sample t-test requires normal distributions and for the variances of the two populations to be equal.
- Ensure candidates know how to describe the patterns observed in a residual plot. The Understanding Standards document '[Residual Plots Guidance]' can help with this.
- Explain the meaning and purpose of the text that accompanies the formulae in the Statistical Formulae and Tables to candidates. For example, candidates should know to refer to page 5 of the Statistical Formulae and Tables, and understand the implications of the sentence above the final two formulae, which says: 'If additionally $\varepsilon_i \sim \mathrm{N}(0,\sigma^2)$ then'.
- Ensure candidates know how to correctly describe the meaning of a confidence interval or a prediction interval using acceptable vocabulary.

Teachers and lecturers delivering the Advanced Higher Statistics course, and candidates taking the course, can consult the detailed marking instructions for the 2025 course assessment on our website. These illustrate the communication requirements for questions on, for example, prediction and confidence intervals, contextual description of assumptions, and writing conclusions in terms of the alternative hypothesis. Our website also contains the marking instructions from previous years.

The Understanding Standards website contains examples of candidate evidence with commentary.

Appendix: general commentary on grade boundaries

Our main aim when setting grade boundaries is to be fair to candidates across all subjects and levels and to maintain comparable standards across the years, even as arrangements evolve and change.

For most National Courses, we aim to set examinations and other external assessments and create marking instructions that allow:

- a competent candidate to score a minimum of 50% of the available marks (the notional grade C boundary)
- a well-prepared, very competent candidate to score at least 70% of the available marks (the notional grade A boundary)

It is very challenging to get the standard on target every year, in every subject, at every level. Therefore, we hold a grade boundary meeting for each course to bring together all the information available (statistical and qualitative) and to make final decisions on grade boundaries based on this information. Members of our Executive Management Team normally chair these meetings.

Principal assessors utilise their subject expertise to evaluate the performance of the assessment and propose suitable grade boundaries based on the full range of evidence. We can adjust the grade boundaries as a result of the discussion at these meetings. This allows the pass rate to be unaffected in circumstances where there is evidence that the question paper or other assessment has been more, or less, difficult than usual.

- The grade boundaries can be adjusted downwards if there is evidence that the question paper or other assessment has been more difficult than usual.
- The grade boundaries can be adjusted upwards if there is evidence that the question paper or other assessment has been less difficult than usual.
- Where levels of difficulty are comparable to previous years, similar grade boundaries are maintained.

Every year, we evaluate the performance of our assessments in a fair way, while ensuring standards are maintained so that our qualifications remain credible. To do this, we measure evidence of candidates' knowledge and skills against the national standard.

For full details of the approach, please refer to the Awarding and Grading for National Courses Policy.