2024 NCEA Assessment Report

Mana Tohu Mātauranga o Aotearoa
New Zealand Qualifications Authority

Subject: Statistics

Level: 3

Achievement standard(s): 91584, 91585, 91586

General commentary

The format of all three examination papers remains consistent from year to year. This year, markers commonly reported that a significant proportion of candidates left question parts completely blank.

Candidates are encouraged to complete as many question parts as they are capable of, across all three questions in each paper. As in previous years, many Merit and Excellence level questions offered Achievement-level opportunities for partial solutions.

Most questions have two main parts, parts (a) and (b), and candidates should be prepared to restart a question in part (b) even if they have encountered challenges in part (a).

Candidates should endeavour to write clearly and neatly. Markers reported difficulty in reading some responses this year. The amount of space allocated in question booklets should be sufficient for an adequate, concise response. If candidates require additional space to complete responses, they should be aware of and use the extra pages at the back of the booklet. Answers written in this part of the booklet should be clearly labelled. Candidates are advised to write succinctly, directly responding to the question asked.

Written responses should be supported with statistical calculations, and sufficient working should be shown to indicate the candidate's thinking and reasoning. Only the final answer for each question should be rounded. For some question parts, particularly when working with probability distributions, the inclusion of sketches or diagrams may be helpful to support calculations. Final responses should be linked back to the question or statement that has been given and should include reference to the context of the question as appropriate.

Report on individual achievement standard(s)

Achievement standard 91584: Evaluate statistically based reports

Assessment

The examination included three questions, each in four to five parts, of which candidates were required to respond to all three questions. Candidates were provided with a resource booklet and an answer booklet.

The resource booklet contained four reports, one for Questions One and Two and two reports for Question Three. The reports were set in real-life contexts, two of which were New Zealand-based. The questions covered the requirements of the 2024 assessment specifications, which were to answer questions about statistically based reports. The questions required the candidate to evaluate claims or conclusions made in the report, including identifying and discussing potential sources of

error associated with statistical studies, calculating and interpreting margins of error, and considering study design and the type of sampling method.

Commentary

Candidates were required to assess the quality of reports using statistical methods indicated by the question, whether it was to do with the design of the study, or to identify potential issues with aspects reported in the study. Candidates needed to read the report and the questions carefully, and then consider what was being asked of them, perhaps by highlighting or underlining key words. Candidates should be aware of the blank pages at the back of the booklet to continue their responses should they run out of room in the allocated question space, or to request extra paper.

Candidates should avoid using generic, learned answers, without considering the context of the report or providing a necessary explanation to relate their observations back to the statistical reports.

Candidates are advised to avoid saying that there is a need to "eliminate bias". Where bias is being discussed, the term to use is "reduce", or words to that effect.

After calculating comparison confidence intervals, it was important to interpret these in context and appropriately discuss the underlying population. For example, if the confidence interval was [–1.5%, 7.5%], then candidates needed to interpret this correctly by saying something like "I am confident that the proportion of UK men and women who can confidently locate their rectum is somewhere between 1.5% less and 7.5% more than the proportion of UK men and women who can confidently locate their reproductive organs."

Candidates who then wrote a claim separately from the interpretation but related to the context were well-rewarded. For example, "Because this interval is both negative and positive, there is not sufficient evidence to support the claim that a higher proportion of UK men and women can confidently locate their rectum than their reproductive organs."

It should also be noted that overlap methods of confidence intervals are not appropriate for this assessment.

Candidates needed to be able to identify, describe, and discuss both experimental and observational studies and apply that knowledge. Further, they are encouraged to mention what type of study the report was, even if it was not clear from the question that it should be identified. Candidates must also realise that a causal claim can be inferred from an experimental study, but not from an observational study. A sample to population inference can be made from an observational study.

Candidates need to understand concepts such as extending the results appropriately, advantages and disadvantages of different survey types, and a wide range of different sampling methods.

Further, candidates need to be able to interpret a range of different statistical displays (graphs) and use numerical information from those displays to make comparative statements in context.

Grade awarding

Candidates who were awarded **Achievement** commonly:

- commented on key features by referring to statistical evidence provided in the reports without including specific details
- calculated a confidence interval correctly
- identified and described a confounding variable
- described the concept of blinding in context
- identified a potential issue with extending the results of an experiment
- identified advantages and disadvantages of surveys conducted by interviews.

Candidates who were awarded **Achievement with Merit** commonly:

- calculated a confidence interval, but could only either correctly interpret it within the context, or justify why a claim was true in context
- identified advantages and disadvantages of survey interviews in context
- identified and described the importance of blinding and associated impacts on a study
- described the effect of post-weighting on survey results
- explained how random allocation will balance effects of confounding variables in context
- described the effect of independence on survey results.

Candidates who were awarded **Achievement with Excellence** commonly:

- understood what regional sample stratification meant with contextual reasoning
- calculated a comparison confidence interval and interpreted it in context, justifying a statistical claim using correct statistical language, including identifying the population
- used and correctly applied statistical language
- understood relevant reasoning for extending results, and could discuss this in context
- identified and described the importance of blinding and associated impacts on a study, including how this affects the response variable.

Candidates who were awarded **Not Achieved** commonly:

- did not refer to the statistical nature of the reports
- did not use the appropriate MOE to calculate the relevant confidence interval for a comparison confidence interval
- did not define or correctly use statistical terms and statistical terminology
- incorrectly focused on the size of samples to discredit studies.
- prefaced discussions with “eliminating bias”.

Achievement standard 91585: Probability concepts

Assessment

The questions covered the requirements of the assessment specifications, which were to calculate probabilities from formulae, a probability distribution table or graph, tables of counts or proportions, simulation results, or from written information.

Familiarity with the use of Venn diagrams, probability trees, and two-way tables of counts was required. It was necessary for candidates to clearly show the method they had used to calculate probabilities, and state any assumptions made.

Commentary

Candidates who attempted all three questions tended to achieve higher grades than those focusing on attempting only one or two questions. Candidates should read questions carefully to identify the necessary response, then write their responses in context, supporting these with necessary calculations.

Candidates need to be familiar with instructions such as “comment”, “interpret”, and “justify”.

Responses to such questions should include statistical calculations and sufficient working should be shown to indicate the candidate’s thinking and reasoning. Final responses need to be linked back to the question or statement that has been given.

While more candidates interpreted a likelihood ratio correctly compared to previous years, there is still a tendency to misinterpret the relative risk, confusing “times as likely” and “times more likely”.

Many candidates demonstrated that they require more practice in reading the question carefully to determine if it requires the use (or not) of conditional probability.

It is also useful for candidates to be familiar with the different tests for independent events, and use the one which is best suited to the information given.

Candidates’ ability to explain how to use a simulation and apply statistical reasoning with “true, theoretical, and experimental” probability in context is generally weak and an area that requires improvement.

Grade awarding

Candidates who were awarded **Achievement** commonly:

- chose and correctly used probability tools best suited to solve a problem for each part
- understood how to test for independence
- used evidence from graphs to answer questions
- processed statistical information to find a proportion
- explained that the experimental probabilities did not match the theoretical probability in a simulation
- recognised reason(s) why probabilities sourced from data may not be valid without justification.

Candidates who were awarded **Achievement with Merit** commonly:

- selected and used an appropriate test for independent events
- justified claims using calculations and / or statistical statements
- completed and used a Venn diagram to find a probability
- found the sample space for rolling two die
- identified the requirement for a conditional probability
- correctly calculated a conditional probability.

Candidates who were awarded **Achievement with Excellence** commonly:

- applied probability concepts using extended abstract thinking in solving problems
- explained the issues involved with extending the results
- calculated and explained relative risk in context
- explained independence fully in context
- demonstrated statistical reasoning using “true, theoretical, and experimental” probability in context
- recognised evidence for a biased die, and analysed the effect this had on the probability of getting a score from two dice
- recognised reasons why probabilities sourced from data may not be valid, and justified why they were not valid
- had a good level of statistical literacy, and were able to answer the question being asked.

Candidates who were awarded **Not Achieved** commonly:

- made no attempt to use a table or probability diagram to organise and display the information given
- did not recognise that probabilities need to be between 0 and 1
- did not use probabilities or numerical evidence to solve problems or support answers
- did not multiply decimals accurately, and did not accurately read scientific notation from calculators

- did not complete a Venn diagram
- struggled with understanding the context of the question
- made no attempt to answer the question.

Achievement standard 91586: Apply probability distributions in solving problems

Assessment

Candidates needed to select and use methods, demonstrate knowledge of concepts and terms, communicate their thinking using appropriate statements from discrete and continuous probability distributions, find the mean and standard deviation of a random variable, and compare distribution of model estimates of probabilities.

Candidates were required to identify which of the Uniform, Normal, Triangular, Binomial and Poisson Distributions were appropriate to use to model observed data. Candidates should be familiar with the features of each distribution and be able to discuss the appropriateness of these distributions and their parameters for modelling observed data.

Tools such as those on https://www.stat.auckland.ac.nz/~fergusson/prob_dist_explorer/fit/ are helpful in assisting candidates in this area.

Being able to identify the random variable or event being discussed in a question is key to success in this assessment.

Some candidates were unclear about the conditions of each probability distribution model, or when discussing the condition of a particular probability distribution, failed to give sufficient evidence or linking to the context to achieve.

Commentary

Candidates with strong calculation skills performed well in this assessment, but those with the ability to calculate probabilities and describe / compare distributions and discuss the appropriateness of these to the context of the question gained higher grades.

Candidates often wrote answers that did not include sufficient working or failed to link their calculations to the context of the question. A question often involves multiple steps. If a minor error in working is made but the final answer is consistent with that error, it is often possible to award a grade. However, if insufficient working or no working is shown the grade for that question will be Not Achieved.

Premature rounding or incorrect rounding continues to be a problem. Candidates should ensure that they do not round their work to less than four decimal places until their final answer.

The ability to calculate relative frequencies from graphs and use these to discuss claims and appropriateness of models is an advantage in this standard.

Grade awarding

Candidates who were awarded **Achievement** commonly:

- identified which distribution to use, and calculated a probability for that distribution
- identified the correct parameters needed to solve a probability distribution problem
- sketched a distribution accurately
- calculated the mean and standard deviation for the distribution of a discrete random variable in table form

- calculated an observed proportion from a frequency graph, and used it to support a claim
- explained the variation in a simulation model
- explained how well a given model matched a bar graph of results.

Candidates who were awarded **Achievement with Merit** commonly:

- completed multi-step problems across a range of distributions
- calculated a conditional probability
- identified reasons why a particular probability distribution model was appropriate to the context of a problem
- discussed the suitability of the parameters of a given probability distribution model to the given context of a problem
- communicated thinking using appropriate statements and calculations
- explained why the standard deviation of two different random variables might be different, in context
- calculated the probability of a Poisson distribution requiring a change in lambda.

Candidates who were awarded **Achievement with Excellence** commonly:

- showed a depth of understanding across a range of distributions, appropriately linking statistical and contextual information
- discussed the appropriateness (or inappropriateness) of a probability distribution model and its parameters by considering features of the probability distribution, statistical evidence, and / or the context of the situation
- proposed and justified the use of an alternative model that could be appropriate for modelling a random variable, with the new parameters identified and justified in context
- used conditional probability correctly in the context of a normal distribution problem, and discussed the suitability of the model parameters in terms of the context
- compared the results of a simulation model with the original observed data and a proposed probability distribution model.

Candidates who were awarded **Not Achieved** commonly:

- did not identify which distribution and parameters they were using when calculating probabilities
- did not calculate a given probability for a normal, binomial, Poisson, or triangular distribution
- did not calculate the mean or standard deviation from a table showing the probability distribution of a random variable
- made calculation errors or rounded prematurely
- failed to show working or link responses to the context of the problem
- did not calculate an observed proportion from a frequency graph and use it to support a claim.
