2022 NCEA Assessment Report

Mana Tohu Mātauranga o Aotearoa
New Zealand Qualifications Authority

Subject: **Statistics**

Level: **3**

Standards: **91584, 91585, 91586**

Part A: Commentary

Candidates need to ensure that they complete as many question parts as they can, across all three questions. Many questions have two main parts, parts (a) and (b), and it is possible for candidates to restart a question in the second part if they have struggled in part (a). Those candidates who wrote full responses to questions, supported by calculations, and clearly showed how they reached their answers, were more likely to achieve Merit or Excellence grades than those who only wrote down solutions.

Candidates need to read questions carefully to search for keywords such as “investigate”, “compare” and “justify”. Responses 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 and should include reference to the context of the question as appropriate.

Candidates should not expect that questions in the examination will be similar to those in past examinations, as questions can be about any context and in any form. Candidates should expect that graph drawing may be required. Rough sketches should be avoided; instead, accurate graphs with scales on both axes and key parameters identified should be drawn. Scales on graphs need to be correct and even.

A continuing issue was candidates rounding answers prematurely, which resulted in inaccurate / incorrect answers and consequent difficulties in following question parts. Candidates should ensure that they do not round their work to less than 4 decimal places, until their final answer. Candidates should ensure that they are familiar with working with small numbers, particularly those in scientific notation, as probabilities may be very small or quite similar to each other. Care should be taken to revise basic number skills (working with decimals, percentages and fractions) if required.

Report on standards

Standard number 91584 Evaluate statistically based reports

Examination content and assessment specifications

The examination included three questions, each in 4-6 parts, of which candidates were required to respond to all three questions. Candidates were provided with a resource booklet and a question-and-answer booklet. The resource booklet contained three reports, one for each question. The reports were set in real-life contexts.

The questions covered the requirements of the 2022 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 inference.

Standard-specific observations

Candidates are required to assess the quality of reports using statistical methods indicated by the question, whether it is to do with the design of the study, or to identify potential issues with aspects reported in the study. Candidates need to read the report and the questions carefully, and then consider what is being asked of them, perhaps by highlighting or underlining key words, that allows them to focus better on their responses. In many cases, candidates provided a great answer that did not answer the specific question at hand, but rather another question elsewhere.

Candidates should avoid using generic, learned answers, for example, “Old people generally do not need the internet”, without considering the report's context or providing the necessary explanation to relate their observations to the statistical reports. Candidates should not say that there is a need to ‘eliminate bias’. Where bias is being discussed, the term to use is “reduce” or words to that effect.

It is important for candidates to understand the effect of sample size on the margin of error and to clearly justify this in terms of the context.

After calculating comparison confidence intervals, candidates need to interpret the confidence interval in context and appropriately discuss the underlying population. For example, if the confidence interval was [41.7%, 54.3%], then candidates needed to interpret this correctly by saying something like “I am pretty sure the proportion of UK parents who agree or strongly agree that their children are more interested in cooking at home during lockdown is somewhere between 41.7% and 54.3% more than the proportion of UK parents who disagree with this statement”. Then the candidate should write the claim SEPARATELY from the interpretation. The claim must also be in context. For example, "As this confidence interval is entirely positive, there is evidence to support that a higher proportion of UK parents agree or strongly agree that their children are more interested in cooking at home during lockdown than disagree with this statement.”

Candidates need to be able to identify, describe, and discuss both experimental and observational studies, and apply that knowledge. Further, it is a good idea to mention what type of study the report is, even if it is not clear from the question that it needs to be identified. Candidates must also realise that a causal claim can be inferred from an experimental study, but not from an observational study. Students, when discussing the features or results of an experimental or observational study, should be sure to include empirical evidence from within the report to back up any statements made. With respect to experimental studies, students need to be able to discuss why it is important to perform random allocation, and back this up with contextual evidence.

Grade related bullet points

Candidates who were awarded **Achievement** commonly:

- commented on key features by referring to statistical evidence provided in the reports without including specific details

- calculated the margin of error (MOE) and described why the MOE is required
- identified the explanatory and response variables in an experimental study in context
- correctly located information in a table of figures
- had an understanding of the need for random allocation in an experimental study
- calculated a confidence interval correctly.

Candidates whose work was assessed as **Not Achieved** commonly:

- did not refer to the statistical nature of the reports
- did not identify the correct sample size to use with the rule of thumb formula to calculate the MOE or describe why it is used
- did not identify the explanatory variables in an experiment with context
- 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
- did not identify random allocation as a key idea in experimental design
- incorrectly focused on the size of samples to discredit studies.

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
- demonstrated a sound understanding of random allocation and supported this reasoning with enough context
- identified an issue with a survey, and related it to representativeness with a specific population
- recognised when to use the rule of thumb MOE and described why it is used
- understood what a national sample meant with contextual reasoning.

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

- described an issue with study design and directly linked it to the strength of a claim in context
- 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.
- acknowledged that for randomised experiments, causal statements can be made, however, generalisations or links between variables can only extend to the participants involved in the experiment
- identified confounding/further variables and clearly commented on their implications.

Standard number 91585 Apply probability concepts in solving problems

Examination content and assessment specifications

The examination included three questions, each in 5 or 6 parts, of which candidates were required to respond to all three. Candidates were provided with a question-and-answer booklet.

The questions covered the requirements of the 2022 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.

Standard-specific observations

Candidates need to practise interpreting questions in order to determine which probability method to use to answer the question. They are encouraged to attempt all three questions, rather than focusing on only one or two. Candidates should write full responses, in context, and explain their decisions when answering questions, supporting their responses with calculations.

Specifically, candidates need to focus on learning to interpret the likelihood ratio (times as likely vs times more likely). They need to be able to analyse situations that may or may not involve conditional probabilities and apply tests for independence when required. It is useful for students to be able to quickly recall tests for complementary, mutually exclusive, and independent events.

The ability of students to describe how to use a simulation and statistically reason using 'true, model and experimental' probability in context is generally pleasing, but this continues to be an area where improvement is still required.

Grade related bullet points

Candidates who were awarded **Achievement** commonly:

- assumed independence of events to calculate the probability of a combined event
- incorporated contextual aspects of the chance situation being modelled in their answers
- used two-way tables appropriately to model situations
- used probability trees appropriately to model situations
- used straightforward probability methods to solve problems
- included clear and correct probability statements as part of their working.

Candidates whose work was assessed as **Not Achieved** commonly:

- did not order decimals correctly
- could not correctly complete a two-way table or Venn diagram
- did not interpret the question correctly – were confused by the probability information

- did not have written responses that answered the question
- did not calculate conditional probabilities
- did not calculate numbers of individuals, or missed doing so
- did not give reasons for non-independence
- did not agree or disagree with a claim
- did not calculate correct proportions
- did not explain true and model probability
- did not identify and explain sampling variability.

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

- applied probability concepts using relational thinking in solving problems
- understood when to use sampling without replacement
- answered claims correctly by stating whether they supported the claim or not and backed up their statement with calculations
- recognised an independence question written in context
- selected an appropriate test for independence
- recognised conditional probability written in context
- calculated a proportion and answered the question in context
- provided valid reasons for estimates and in context
- interpreted the claim ‘at least twice as likely’
- understood the concept of a total probability of 1.

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

- applied probability concepts using extended abstract thinking in solving problems
- integrated contextual and statistical knowledge and gave responses in context
- understood concepts that would / could be assumed when taking a sample
- reasoned why / why not survey results can be applied to other situations
- demonstrated statistical reasoning using ‘true, model and experimental’ probability in context
- explained non-independent events, explained the effect of two events and said if claims are justified or not
- identified the effects of low samples.

Standard number 91586 Apply probability distributions in solving problems

Examination content and assessment specifications

The examination included three questions, each in 5 or 6 parts, of which candidates were required to respond to all three. Candidates were provided with a question-and-answer booklet.

The questions covered the requirements of the 2022 assessment specifications which were to clearly identify the probability distribution applied in solving each problem, stating assumptions made and to calculate probabilities from probability distributions presented as formulae, tables or graphs of data, simulation results or written information.

Candidates needed to be familiar with, and justify the use of, the normal, Poisson, binomial, uniform, and triangular distributions and they were expected to understand the calculation of the mean and standard deviation of a random variable.

Standard-specific observations

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.

A common issue was candidates rounding answers prematurely which resulted in inaccurate / incorrect answers. Candidates should ensure that they do not round their work to less than 4 decimal places, until their final answer.

Understanding and identifying the random variable or event being discussed in a question is key to success in this assessment. Many candidates struggled with this.

Some candidates were unclear about the conditions of each probability distribution model and, when discussing a particular probability distribution, frequently confused its conditions with those of other probability distributions. Merit and Excellence candidates need to explain how the conditions of a given probability distribution model are appropriate or inappropriate to the given context of that random variable.

Candidates often wrote answers that had insufficient detail or did not link their answers to the context. Merit and Excellence questions frequently have instructions such as “investigate”, “compare” and “justify”. Responses to such questions should include statistical calculations, and sufficient working should be shown to indicate the candidate’s thinking and reasoning. Discussions about any assumptions made when applying the distribution model and the validity of these assumptions in terms of the given context should always be given even when they are not specifically asked for. Final responses need to be linked back to the question or statement that has been given.

Grade related bullet points

Candidates who were awarded **Achievement** commonly:

- identified which distribution to apply and calculated a probability for that distribution
- understood terms such as ‘at least’ and ‘more than

- identified the correct parameters needed to solve a probability distribution problem
- calculated the mean and standard deviation for the distribution of a discrete random variable in table form
- provided an assumption in context relating to one of the distributions
- calculated an observed proportion from a frequency graph.

Candidates whose work was assessed as **Not Achieved** commonly:

- did not identify which distribution to apply
- did not calculate a given probability for a binomial, Poisson, or triangular distribution
- did not calculate probabilities from a table showing the probability distribution of a random variable
- provided responses containing errors or rounded prematurely
- did not understand the difference between a condition of a distribution model and an assumption
- did not describe the conditions of a distribution in context
- equated $\frac{1}{3}$ with 0.3 or $\frac{33}{100}$ or 0.33
- did not calculate probabilities for problems involving “less than”.

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

- completed multi-step problems across a range of distributions
- calculated and compared means and standard deviations for two sets of data
- calculated a straightforward inverse Poisson problem
- compared observed and model distributions probabilities to draw a conclusion
- understood assumptions made when using probability distribution models and were able to discuss whether these assumptions were likely to be valid in the context of a problem
- communicated their thinking using appropriate statements and calculations.

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 by considering features of the probability distribution, statistical evidence and / or the context of the situation
- used conditional probability correctly in the context of the triangular distribution
- justified the suitability of a probability distribution model for a given observed distribution by comparing the features of the model with that of the graph of the observed distribution and its context.