

# 2025 NCEA Assessment Report

<b>Subject:</b>	Statistics
<b>Level:</b>	3
<b>Achievement standard(s):</b>	91584, 91585, 91586

## General commentary

Across all three papers, the format was consistent with previous years. Candidates were expected to answer a variety of linked question parts that provided opportunities for Achievement, Merit, and Excellence performance. Markers regularly note that a significant number of candidates leave parts of questions blank and that students should attempt as many parts as they can, for example restarting a question in part (b) if part (a) was challenging, and read the questions thoroughly to ensure they address what's being asked.

Marker reports also emphasise that correct numerical answers alone are not sufficient for higher grades; responses must be supported by appropriate justification (including calculations) and clear reasoning that is linked back to the context of the question or statistical report. Sufficient working should be shown to communicate thinking, and diagrams or sketches are helpful where appropriate. Additionally, careful use of calculator skills (including familiarity with standard form) and avoidance of premature rounding remain important to avoid inaccuracies.

Overall, candidates are encouraged to write succinctly, include working and explanations, and engage with the context and language of the question to demonstrate deeper statistical understanding, rather than just computational skills.

## Report on individual achievement standard(s)

### Achievement standard 91584: Evaluate Statistically Based Reports

#### Assessment

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.

#### Commentary

Candidates needed to read the report and the questions carefully, and then consider what is asked of them, 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 that they can 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 should avoid saying that there is a need to “eliminate” bias. Where bias is concerned, 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 **[−0.7%, 10.7%]**, then candidates needed to interpret this correctly by saying something like: *“We are confident that the proportion of house plant owners in New Zealand, Canada, the USA, the UK, and Australia who spend less than one hour per week taking care of their house plants is between 0.7% less and 10.7% more than the proportion who spend between one and three hours per week.”*

Candidates who then wrote a claim separately from the interpretation but related to the context were well-rewarded. For example, *“Because this confidence interval includes both negative and positive values, there is not sufficient evidence to support the claim that, in the population of **house plant owners in New Zealand, Canada, the USA, the UK, and Australia**, a higher proportion spend less than one hour per week caring for their house plants than those who spend between one and three hours.”*

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. This includes features of experimental and observation study design.

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:

- demonstrated basic statistical understanding and applied familiar methods, but often lacked depth, clarity, or contextual interpretation
- calculated margins of error and confidence intervals correctly (or approximately correctly), but did not fully interpret them in context, often omitting the population, variable of interest, or a clear claim / inference
- showed an understanding that the margin of error accounts for natural sampling variation, but were unable to clearly explain what this meant in the context of the study
- identified key features of reports using general references to statistical evidence, without including sufficient detail or explanation
- recognised that random selection helps produce a representative sample, but explanations were often surface-level
- correctly identified explanatory and response variables and could distinguish between experimental and observational studies, but did not always link this understanding to valid inference
- identified potential issues such as confounding variables, incentives, or response bias, but explanations of how these issues may have impacted the results were often brief or vague

- used terms such as bias, fair, or placebo correctly in isolation, but without clear explanation of how these concepts operated in the specific study
- referred to patterns, comparisons, or recognising differences, rather than correctly identifying order effects caused by participant changes over time (such as fatigue, boredom, or attention drift)
- could identify when it was inappropriate to extend results, but did not always clearly explain why the claim could not be generalised to the wider population.

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

- calculated confidence intervals correctly and were able to either interpret them in context or justify a claim, but not consistently both, often missing full formal structure
- explained the usefulness of the margin of error in context and could calculate it correctly but sometimes struggled to fully articulate its implications for inference
- linked confidence interval calculations to contextual inference statements, usually naming the population and variable, but with errors or omissions in either the interpretation or the final claim
- identified explanatory and response variables and described them in the context of the study or experiment.
- demonstrated understanding of random allocation, explaining how it helps balance confounding variables in context, though explanations were sometimes more descriptive than evaluative
- identified order effects (e.g. fatigue, rushing, loss of attention) and explained why randomising order reduces these effects, though links to overall validity were sometimes implicit rather than explicit
- identified and explained advantages and disadvantages of survey methods in context
- recognised non-sampling errors (e.g. incentives) and could explain how these might impact results, but explanations were sometimes limited to either what or why, rather than both
- identified mismatches between the sampling frame and the target population and explained why this limits extending results beyond the study context
- used appropriate statistical language, but occasionally relied on imprecise terms, such as “eliminates bias” instead of “reduces / minimises bias,” which prevented Excellence
- demonstrated sound statistical understanding and correct application of methods, with explanations that were generally clear and relevant to the context, but lacked the precision, synthesis, or formal wording required for Excellence.

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

- calculated comparison confidence intervals (including differences in proportions) and used formal statistical language to both interpret the interval and justify a statistical claim, explicitly identifying the population and variable of interest
- demonstrated clear understanding of margin of error, including when it was appropriate or inappropriate to use it to construct an approximate 95% confidence interval, and clearly explained the implications for inference
- demonstrated a strong ability to move from sample evidence to population inference, accurately writing inference statements and linked claims in full, formal wording
- explained both the “why” and the “how” of design choices (e.g. random allocation, random order) and clearly linked these to their impact on the response variable
- demonstrated clear understanding of order effects, accurately identifying participant changes over time (such as fatigue, learning, or loss of attention), and explaining how randomisation ensured observed differences could be attributed to the treatment rather than these effects
- identified and discussed non-sampling errors (e.g. incentives), explaining how these could impact results, and why they limit the strength of conclusions

- demonstrated understanding of sampling methods, including regional stratification and mismatches between the sampling frame and target population, and clearly explained the implications for extending results appropriately
- used precise statistical language consistently, avoiding common misconceptions (such as confusing patterns with order effects or claiming bias was “eliminated”)
- demonstrated strong structural awareness of common question types, allowing them to address all required components (calculation, interpretation, inference, and claim) coherently and efficiently
- demonstrated strong statistical insight, consistently applying methods correctly and explaining their reasoning clearly and precisely in context.

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

- did not answer the question asked, or wrote very little, sometimes describing the methodology from the resource material without addressing the statistical requirement of the question
- did not respond in the context of the study or experiment, instead giving general or opinion-based statements about bias, fairness, or reliability
- did not correctly calculate a margin of error or confidence interval or used an inappropriate margin of error for the situation
- used an incorrect confidence interval method, such as applying the wrong confidence interval formula, or using an overlap method rather than the required approach
- calculated parts of a method (e.g. MOE), but did not understand what it represented or how to use it to form a confidence interval
- did not correctly interpret a confidence interval, often omitting the population, variable, or any meaningful inference
- used statistical terminology incorrectly or imprecisely, including phrases such as “eliminate bias” without explanation
- focused on sample size alone as a reason to discredit studies, rather than discussing representativeness or scope of inference
- did not explain the reasoning behind study design choices (e.g. why a study was experimental, or why random selection or allocation was used)
- did not identify or correctly describe explanatory and response variables
- did not recognise when results could not be extended to a wider population, or were unable to explain why such an extension was inappropriate
- provided responses that were too brief, unclear, or fragmented to demonstrate statistical understanding.

---

## Achievement standard 91585: Apply probability concepts in solving problems

### Assessment

In this assessment candidates apply probability concepts to solve problems, including calculating probabilities in contexts involving independent, combined, conditional events, and interpreting probability from tables, trees, and other representations. Candidates need to read questions carefully to identify the necessary response, then write their responses in context and support these with necessary calculations.

## Commentary

Candidates who attempted all three questions tended to achieve higher grades than those focusing on attempting only one or two questions.

Candidates need to read questions carefully to identify the necessary response, then write their responses in context, and support these with necessary calculations.

Candidates need to ensure that they can read numbers in scientific notation from their calculators, and be careful not to mix up decimal and percentage notation.

It is useful for candidates to be able to quickly recall tests for mutually exclusive and independent events. There was confusion between the tests for mutually exclusive and independent events from many candidates. These candidates need to extend their explanation of mutually exclusive events to include the relevant probability for the intersection, rather than just using a number.

There is a noted increase in the number of students who are correctly able to interpret likelihood ratios, specifically the distinction between “times as likely” and “times more likely”, however, there is still room for improvement. Evaluating a claim does not necessarily need to have a likelihood ratio to do so. Candidates who were able to read the question and identify the justification required tended to do better than those who automatically used a likelihood ratio, especially when it was interpreted incorrectly.

Candidates' ability to identify true probability, model, and experimental estimates was weak and an area that requires improvement, along with the justification as to which estimate is the most reliable and why.

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.

## Grade awarding

Candidates who were awarded **Achievement** commonly:

- could solve one-step problems or partially solve complex ones
- chose which probability tool was best suited to solve a problem for each part, and then used it to find the correct probability
- demonstrated limited knowledge of how to prove events are independent or mutually exclusive
- used evidence from graphs to answer questions
- processed statistical information to find a proportion
- provided probabilities for model estimates, experimental estimates and true probability.

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

- selected and used an appropriate test for independent events
- selected and used an appropriate test for mutually exclusive events using probabilities
- justified a claim using calculations and / or statistical statements
- completed a Venn diagram and found the required probability
- used conditional probabilities to solve problems
- used a probability tree to calculate probabilities
- carried out a range of probability calculations accurately and with correct notation.

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

- applied probability concepts using extended abstract thinking in solving problems

- used probabilities and probability concepts to validate or refute a claim with clear reasoning and correct mathematical statements
- understood the concepts of true probability, experimental estimates, and model estimates and were able to compare them to make a reasoned decision as to which was better to use for future predictions
- calculated and explained relative risk in context
- explained independence fully by giving a clear contextual description of the relationship between the two events
- related a context to how results could make impact on a statistical claim
- demonstrated a good level of statistical literacy and were able to answer the question being asked.

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

- did not 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 their answers
- did not read scientific notation off their calculators correctly
- used a percentage sign after a decimal answer eg. 0.25% instead of 0.25 or 25%.
- could not correctly complete a Venn diagram
- could not apply formulae for mutually exclusive events, independence, and set theory (unions and intersections)
- could not complete a discrete probability distribution table with a single missing value to correctly sum to 1
- provided answers without any supporting numerical evidence or failed to use probability theory to justify their conclusions
- did not demonstrate understanding the context of the question.

## **Achievement standard 91586: Apply probability distributions in solving problems**

### Assessment

In this assessment candidates apply probability distributions in solving problems, including selecting appropriate continuous or discrete distributions (normal, uniform, triangular, binomial, or Poisson), calculating probabilities, and justifying choices of models and parameters in context.

### Commentary

Candidates need to be able to identify which of the uniform, normal, triangular, binomial, and Poisson distributions is 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.

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

When identifying and discussing the conditions of a probability distribution model, candidates must ensure that they unpack the condition in terms of the context of the question. General statements of conditions are insufficient.

It is important to note that, for a Poisson distribution, the conditions of independence and a constant rate are often mistakenly described in terms of probabilities. The requirements are that events occur independently of one another and that the average rate of occurrence per unit of time or space remains constant. While Poisson processes often involve rare events and discrete counts, these features are not formal conditions of the distribution; rather, they are useful indicators that a Poisson model may be appropriate.

Candidates with strong calculation and probability 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. Often parts of questions are linked together, or a question 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 must be Not Achieved.

When solving a problem, candidates need to ensure that they identify the distribution they are using and its parameters. Often a similar answer to the correct answer can be obtained using the wrong distribution. In this case, a failure to identify the distribution and parameters will prevent the award of a grade.

Premature rounding or incorrect rounding continues to be a problem. Candidates should ensure that they do not round their work to less than 4 decimal places, until their final answer.

The ability to calculate relative frequencies from graphs, particularly histograms, and use this information to discuss the appropriateness of a probability distribution model 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 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
- stated the independence assumption in context.
- calculated a proportion from a histogram and used it to evaluate a claim.

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

- completed multi-step problems across a range of distributions
- identified reasons why a particular probability distribution model was appropriate to the context of a problem
- calculated a conditional probability
- discussed the suitability of the parameters of a given probability distribution model to the given context of a problem
- communicated their thinking using appropriate statements and calculations
- calculated a standard deviation for a normal distribution given a probability
- found a rate given a probability (Inverse Poisson) and linked it to the context of the question.

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 a standard deviation that fulfilled a set of constraints using the normal distribution
- identified an assumption made when using a probability distribution, evaluated its validity, and explained how altering it would affect the accuracy of a probability calculation
- used statistical reasoning to justify their choice of a probability distribution and its parameters.

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

- did not identify which distribution and parameters they were using when calculating a probability
  - could not calculate a given probability for a normal, binomial, Poisson, or triangular distribution
  - could not calculate the mean from a table showing the probability distribution of a random variable
  - made calculation errors or rounded prematurely
  - were not able to handle calculator answers given in scientific notation
  - failed to show their working or link their response to the context of the problem
  - could not calculate an observed proportion from a frequency graph and use it to support a claim
  - could not explain the independence assumption in context
  - did not understand that probabilities are between 0 and 1.
-