 Year 12 Mathematics Advanced

Assessment task

MA-S2 Descriptive statistics and bivariate data analysis and MA-S3 Random Variables

Driving question

How well can mathematics predict outcomes?

Outcomes

* **MA12-8** solves problems using appropriate statistical processes
* **MA12-9** chooses and uses appropriate technology effectively in a range of contexts, models and applies critical thinking to recognise appropriate times for such use
* **MA12-10** constructs arguments to prove and justify results and provides reasoning to support conclusions which are appropriate to the context

All outcomes referred to in this unit come from [Mathematics Advanced](http://educationstandards.nsw.edu.au/wps/portal/nesa/11-12/stage-6-learning-areas/stage-6-mathematics/mathematics-advanced-2017) Syllabus  
© NSW Education Standards Authority (NESA) for and on behalf of the Crown in right of the State of New South Wales, 2017

Learning across the curriculum

* Critical and creative thinking
* Information and communication technology capability
* Literacy
* Numeracy
* Work and enterprise

Context

The normal distribution was initially devised as a result of analysing the probabilities of gambling games. Since then it has been discovered to have applications in many natural and financial situations. During this task students will investigate one such application, the modelling of random errors and apply their understanding to determine the confidence associated with a prediction generated through least-squares regression analysis.

Task

Part A

Is the wisdom of crowds superior to the wisdom of any given individual?

Over 100 years ago at an agricultural fair, a competition was run asking participants to guess the weight of an ox displayed at the fair. Over 800 people participated and no-one guessed the exact weight; however, Francis Galton, a highly reputable statistician of the time, was intrigued by the competition. He analysed the guesses and was surprised to find that the average of all the guesses was almost exactly equal to the weight of the ox. The ox weighed 1198 pounds and the average guess was 1197 pounds. This led to a theory called the Wisdom of Crowds.

Before attempting this task, research the theory of the Wisdom of Crowds. Here are some websites to consider:

[The parable of the ox](https://www.johnkay.com/2012/07/25/the-parable-of-the-ox/)

[The wisdoms of the crowd](https://jemmaths.wordpress.com/2016/10/22/the-wisdom-of-the-crowd/)

As part of their task students:

1. design and deliver an experiment in which participants freely and privately provide a guess to the outcome of a competition. Examples of experiments can be found on the internet.
2. summarise and display their results.
3. describes the results by discussing the shape, location and spread of the distribution of the results.
4. answer the driving question “Is the wisdom of crowds superior to the wisdom of any given individual?”

Part B

How do manufacturing companies ensure products match quality assurance statements?

One recognised application of the normal distribution is to model random errors. For businesses, random errors occur during manufacturing processes, i.e. when items are produced or packed there is always a margin of error. For example, during the filling of a 375mL can of soft drink an error may occur and the can may contain 376mL or 374mL. However, for a product to satisfy quality assurance standards at least 97.5% of the cans must be filled at 375mL or above. The same applies to all manufactured products. At least 97.5% of each product must satisfy the quantity or capacity stated on the labelling, otherwise the company is in breach of quality assurance standards. To determine whether products meet these standards, products are sampled during the manufacturing process.

In this part of the task, you are to investigate a brand of matchsticks. The following link shows an example of the manufacturing process for matchsticks: [YouTube - Auto match filling and box combine machine with camera](https://www.youtube.com/watch?v=XtsWz8b7IS8)

Each matchbox should contain 100 matches. Determine if the brand of matchboxes meets the quality assurance standard by investigating whether 97.5% of matchboxes contain at least 100 matches.

1. Sample the brand of matchsticks by determining the number of matchsticks in each matchbox and record the results.
2. Summarise and display the results.
3. With reference to the empirical rule for normally distributed random variables, justify whether the matchsticks are in breach of quality assurance standards. If they are not, provide suggestions so that they do meet the standards.

Part C

How well can mathematics predict outcomes?

During this part of the assessment students may use the [data.worldbank.org](https://data.worldbank.org) website to access data. Students need to investigate the population of a country over time. They will need to model the population using linear regression techniques, interpret the model in context and use the model to generate a prediction.

Using the World Bank website listed above, search “population, total” to display the world’s population over time. Each country’s population is listed below the population-time graph and can be displayed by clicking on the country name. The data for each country’s population can be accessed using the [World Bank databank](https://databank.worldbank.org/data/reports.aspx?source=2&series=SP.POP.TOTL&country=). In the variables section on the left, drop down the time option and then view the last 50 years by locating the view recent years option and select 50. Select Excel from the download options to download as a spreadsheet.

1. Use the information above to identify a country where there is a strong linear relationship between the population and time. The correlation coefficient must be greater than 0.95.
2. Identify a country where there is insufficient evidence of a linear relationship between the population and time. Explain your choice of country and explain why you think it does not follow a linear relationship.
3. Using the data from question 1, develop a correlation model for population against time.
4. Interpret the gradient from the model in the context of the question.
5. Use the model to predict the population of the country identified in question 1 for the year 2050.

It is unlikely that the correlation model fits perfectly through all of the data points. Therefore there is variation between our data points and the values generated by the model. When there is little variation, we can be confident in the accuracy of our prediction. Conversely, when there is a high level of variation, we are less confident in the accuracy of the prediction. The variation of each data point is calculated as the difference of the population from the data and the population calculated from the model, for the same value of time. This is called the residual. The residuals can be considered as random errors and are therefore normally distributed.

It is expected students will research background information to support their progress through the following questions. Students may benefit by researching key phrases like ‘understanding residuals’, ‘residuals in Desmos’, ‘prediction interval’ and so on.

1. Use the residual values to calculate the standard deviation for the variation of the data from the model.
2. Use the standard deviation to develop a 95% prediction interval.
3. Evaluate the process of regression analysis to develop a correlation model and the setting of a prediction interval to answer the driving question, “How well can mathematics predict outcomes?”

What to submit:

As part of this task students need to submit

* Handwritten or electronic responses to each question
* References to sources of data and/or include tables of all data used. This may be in the form of a spreadsheet
* Evidence of the experiments being performed. Students may like to provide photos or a short video with an annotation
* References to all websites used during any research

Success criteria

Part A

| Fluency, understanding and communication | Problem solving, reasoning and justification |
| --- | --- |

| Criteria | Working towards developing | Developing | Developed | Well developed | Highly developed |
| --- | --- | --- | --- | --- | --- |
| 1. Design and deliver an experiment  **MA12-8** | Student designs and delivers a suitable experiment. | Student delivers the experiment to an appropriate number of participants and maintains the confidentiality of the responses. |  |  |  |
| 2. Summarise and display results  **MA12-8 MA12-9** | Student displays results using an appropriate table. | Student displays the results using an appropriate display.  Student calculates appropriate statistics. |  |  |  |
| 3. Describe the distribution of the results  **MA12-8** | Student calculates appropriate statistics to represent the results. | Student describes the distribution of results using appropriate language, while interpreting the statistics to describe location and spread. |  |  |  |
| 4. Answer the driving question  **MA12-10** |  |  |  | Student interprets the results of the experiment to draw a conclusion. | Student discusses their confidence in the conclusion and provides suggestions to develop a more rigorous conclusion. |

Part B

| Fluency, understanding and communication | Problem solving, reasoning and justification |
| --- | --- |

| Criteria | Working towards developing | Developing | Developed | Well developed | Highly developed |
| --- | --- | --- | --- | --- | --- |
| 1. Sampling  **MA12-8 MA12-9** | Student performs the experiment. |  |  |  |  |
| 2. Summarise and display results  **MA12-8 MA12-9** | Student displays results using an appropriate table. | Student displays the results using an appropriate display.  Student calculates appropriate statistics. |  |  |  |
| 3. Using empirical rule  **MA12-8 MA12-10** |  |  | Student applies some understanding of the empirical rule and is able to show some interpretation of the results. | Student provides a reasonable conclusion by interpreting and referencing the results. Most of their reasoning is appropriate but incomplete. | Student provides a reasonable conclusion by interpreting and referencing the results. Reference has been made to the empirical rule in the context of the question. |

Part C

| Fluency, understanding and communication | Problem solving, reasoning and justification |
| --- | --- |

| Criteria | Working towards developing | Developing | Developed | Well developed | Highly developed |
| --- | --- | --- | --- | --- | --- |
| 1. Identify a linear correlation  **MA12-8 MA12-9** | Student identifies a country with linear relationship for population growth and a correlation coefficient greater than 0.95 |  |  |  |  |
| 2. Identify a country with no linear correlation  **MA12-8 MA12-9 MA12-10** | Student interprets the correlation coefficient to identify a country that does not display a linear relationship for population growth. | Student links the absence of a linear relationship to an event or events in the country’s history |  |  |  |
| 3, 4 and 5. Develop, interpret and use the model  **MA12-8 MA12-9** | Student develops a linear correlation model | Student interprets the gradient in the context of the question | Student uses the model to make a prediction for the population in the year 2050 |  |  |
| 6, 7 and 8. Calculating and interpreting the variation of the model  **MA12-8 MA12-9 MA12-10** | Student calculates the residual values from the data and the model | Student uses the residual values to calculate the standard deviation for the variation of the data. | Student applies their understanding of the empirical rule to develop 95% prediction intervals. | Student interprets the prediction intervals to quantify how well mathematics can predict outcomes. | Student suggests an improvement to the prediction and justifies it with mathematics. |

Note to staff

The success criteria above has been designed for students and staff alike to use. Students should be presented the rubric as part of the assessment task package. Students and staff follow the process of the task downwards through the rubric and the depth of responses, for each element, across the rubric. Students should be encouraged to use the rubric to self-assess their progress as an assessment-as-learning strategy.

The aim of the assessment task is to develop students’ deep content knowledge. This is reflected in the descriptors, **working towards developing** through to **highly developed**. The level of skill and understanding required in each part of the task is different; some parts require **highly developed** or **well-developed** skills, other parts only capture a **developing** skill set.

None of the working mathematically elements are distinct and when demonstrating one element, you are invariably demonstrating another. As an example, communication runs concurrently through all the other working mathematically elements. Students cannot respond to this assessment without communicating in some form. However, it is envisaged that there is a general progression through the working mathematically elements, starting with fluency and leading to understanding, problem solving, reasoning and justification, with increasingly higher levels of communication accompanying each element. Careful consideration has been given to the position of the success criteria statements so they reflect the working mathematically elements demonstrated.

This assessment task has been designed to illuminate the style of questions and the types of responses needed to elicit deep content knowledge, however, staff are encouraged to use and adapt the assessment task and the success criteria to their school context. Staff may like to enhance or amend sections of the task. Staff may like to adapt the rubric to assign marks to the descriptors in order to differentiate between responses that address the same statement. All changes are the responsibility of the staff using the assessment.