CHUNG CHENG HIGH SCHOOL (MAIN)
Sec 3 Additional Mathematics 2024
Weighted Assessment 3

Mathematical Modelling of COVID-19 Transmission in Singapore
Group Component

10

Group Members:

1. _____ ( )

2. _____ ( )

3. _____ ( )

4. _____ ( )

Group Number: _____

Class: _____

Submission Deadline: 26th July 2024

Term 3 Week 5, Friday 15 00h

The Weighted Assessment 3 consists of two components:

(a) Group Component (10 marks)

(b) Individual Component (20 marks)

Instructions for work submission

(i) Group Submission:
Each group must submit both a Microsoft Excel document and a Google Sheet. Details on the required content for these documents are provided under the various 'Group Tasks'. Submission should be made online via the Google Classroom.

(ii) Individual Submission:
Each student must submit a hardcopy of their Individual Component.

(iii) Plagiarism or any reproduction of other groups' solution will lead to loss of marks.

(iv) You may refer to the marking rubrics on page 10 of the question paper.

(v) The number of marks for each task, question, or part of a question is indicated in brackets [ ] at the end of the respective item.

1 Background

The novel coronavirus disease (COVID-19) that emerged in November 2019 has swept the world for more than 2 years, and Singapore was not spared from the outbreak. The virus that causes COVID-19 is known as the severe acute respiratory syndrome coronavirus 2. This virus is genetically similar to the virus that caused the SARS outbreak in 2003.

The first COVID-19 case in Singapore was confirmed on 23 January 2020. By late March and April, COVID-19 clusters were developed, which soon contributed to an overwhelming proportion of new cases in the country. To stem the tide of infections, circuit breaker lockdown measures were implemented from April to June 2020, after which restrictions were gradually lifted.

2 The use of Mathematical Modelling in Epidemiology

Mathematical models for transmission of infectious diseases are essential tools in epidemiology. These models help to predict the spread of infectious diseases and in making informed decisions regarding intervention measures. A simple model is the Exponential Growth Model.

The Exponential Growth Model

The exponential growth model assumes the number of infections is proportional to the current number of infected individuals. This model is used in the early stages of an outbreak when resources are not limiting the spread.

$$N(t) = N_0 \mathrm{e}^{rt}$$

where:

- $N(t)$ is the number of infected individuals at time t.
- N_0 is the initial number of infected individuals.
- r is the growth rate.
- e is the base of the natural logarithm.

Another model that involves the use of exponential functions is the Logistic Growth Model.

You can learn more about the exponential model and logistic growth model through the video below.

https://www.khanacademy.org/science/ap-biology/ecology-ap/population-ecology-ap/v/exponential-and-logistic-growth-in-populations

3 Objective of Task

In your team, you will model the transmission of COVID-19 in Singapore using an appropriate mathematical function. You will apply this function to solve related questions which require the concepts that you have learnt in Additional Mathematics, notably Exponential and Logarithmic functions, as well as Linear Law.

This task aims to enhance your understanding of the practical application of mathematics in real-life contexts using real-life data. Additionally, you will acquire knowledge and skills in using Excel for mathematical modelling.

4 Covid Transmission in Singapore between January 2020 and May 2020.

The number of COVID-19 cases recorded in Singapore for the year 2020 between January and May is given in Table 4.1 below. The number of **new cases** per week is recorded at the start of each week.

Date	No. of cases
2020-01-26	4
2020-01-27	0
2020-01-28	0
2020-01-29	0
2020-01-30	0
2020-01-31	0
2020-02-01	0
2020-02-02	14
2020-02-03	0
2020-02-04	0
2020-02-05	0
2020-02-06	0
2020-02-07	0
2020-02-08	0
2020-02-09	22
2020-02-10	0
2020-02-11	0
2020-02-12	0
2020-02-13	0
2020-02-14	0
2020-02-15	0
2020-02-16	32
2020-02-17	0
2020-02-18	0
2020-02-19	0
2020-02-20	0
2020-02-21	0
2020-02-22	0
2020-02-23	17
2020-02-24	0
2020-02-25	0
2020-02-26	0
2020-02-27	0
2020-02-28	0
2020-02-29	0
2020-03-01	13
2020-03-02	0

Date	No. of cases
2020-03-03	0
2020-03-04	0
2020-03-05	0
2020-03-06	0
2020-03-07	0
2020-03-08	36
2020-03-09	0
2020-03-10	0
2020-03-11	0
2020-03-12	0
2020-03-13	0
2020-03-14	0
2020-03-15	74
2020-03-16	0
2020-03-17	0
2020-03-18	0
2020-03-19	0
2020-03-20	0
2020-03-21	0
2020-03-22	220
2020-03-23	0
2020-03-24	0
2020-03-25	0
2020-03-26	0
2020-03-27	0
2020-03-28	0
2020-03-29	370
2020-03-30	0
2020-03-31	0
2020-04-01	0
2020-04-02	0
2020-04-03	0
2020-04-04	0
2020-04-05	387
2020-04-06	0
2020-04-07	0
2020-04-08	0

Date	No. of cases
2020-04-09	0
2020-04-10	0
2020-04-11	0
2020-04-12	1110
2020-04-13	0
2020-04-14	0
2020-04-15	0
2020-04-16	0
2020-04-17	0
2020-04-18	0
2020-04-19	3693
2020-04-20	0
2020-04-21	0
2020-04-22	0
2020-04-23	0
2020-04-24	0
2020-04-25	0
2020-04-26	6701
2020-04-27	0
2020-04-28	0
2020-04-29	0
2020-04-30	0
2020-05-01	0
2020-05-02	0

Table 4.1: Covid Data Singapore, January 2020 to May 2020
Source: Data adapted from https://ourworldindata.org/coronavirus

5 Group Component [10 marks]

- There are a total of **6** tasks under the group component.
- The primary goal of the group component is to derive a Mathematical Function using Microsoft Excel, that models the transmission of COVID-19 in Singapore, specifically fitting the data points provided in table 4.1.
- Each task has a specific objective which will be described under the individual task instructions.
- Detailed steps required to complete each task will also be provided. Additionally, some tasks will be supported by instructional videos to guide you in completing them.

Task 1: Using Microsoft Excel to organise the data [2 marks]

Your team will organise the raw data so that it is easier to read, interpret and enable excel formulas to be applied more effectively for calculations and analysis.

Step 1

Open Microsoft Excel.

- Create a new worksheet and name the file using the format
  'COVID-19 Data_WA3_Class_Group number'.

Step 2

Enter the Data.

- Organise the data from Table 4.1 into two columns:
  - Column 1: Record the week number.
  - Column 2: Record the corresponding number of new COVID-19 cases at the start of each week.

Step 3

Add appropriate Headings.

- Column 1: You may label as "Week".
- Column 2: You may label as "Actual COVID19 Cases".

Task 2: Inserting formula of an exponential function to model the data [2 marks]

Your team will use an Excel formula to represent an exponential function that fits the data points in Table 4.1.

In this task, you will use the Exponential Growth Model $y = ae^{bt}$ where a and b are constants to be found using the Microsoft excel 'Sover'.

Step 1

Assume Initial Values

- Label two empty adjacent cells as "a" and "b".
- At this point, we do not know the values of constants a and b in the exponential function that fits the data points, so we will assume both values to be 1 (or any other arbitrary numbers, except 0).
- Under each respective cell "a" and "b", enter the value of 1.

Step 2

Create a new column that calculates the number of new COVID-19 cases based on a mathematical function.

- Create a new column adjacent to "Actual COVID10 Cases" and it as "Model COVID19 Cases".
- In the first cell under the column "Model COVID19 Cases", insert the formula that calculates the number of COVID-19 cases based on the Exponential Growth Model

  =A*EXP(B*<week>)

  where
  - A is the cell reference for constant a (make this an absolute reference by adding the dollar sign, e.g., 'E2').
  - B is the cell reference for constant b (make this an absolute reference by adding the dollar sign, e.g., 'F2').
  - <week> is the cell reference for the week number in the corresponding row (e.g., 'C2' if the week number is in column C).
  - EXP is the function to calculate the exponential of a given number.

- Drag the fill handle down to apply this formula to compute the number of cases for all rows (all the respective weeks).

- Formula Example: If the cell E2 contains the value of constant a, cell F2 contains the value of the constant b and 'C2' is the week number in column C, then the formula in the first cell under "Model COVID19 Cases" would be

  =E2*EXP(F2*C2)

Task 3: Finding the Root Mean Square Error (RMSE) [2 marks]

The Root Mean Square Error (RMSE) is commonly used to measure the difference between the predicted values (using the Mathematical Model) and the actual values in a dataset.

The smaller the RMSE, the more accurate the Mathematical Model.

How to calculate RMSE?

(a) Find the difference between each predicted value and the corresponding value in a data set.

(b) Find the square of each difference found in (a).

(c) Find the average of the squared differences found (b).

(d) Find the square root of the average of the squared differences found in (c).

Step 1

Create a new column (adjacent to "Model COVID-19 Cases") and label the column as "Squared Difference".

Step 2

Under the column "Squared Difference", insert a formula that calculates the 'Square Difference" which is defined as

$$(\text{Actual COVID19 Cases} - \text{Model COVID19 Cases})^{2}.$$

Step 3

Drag the fill handle to apply the formula to all rows (to compute the "Squared Difference" for all respective weeks).

Step 4

In an empty cell, insert a formula that calculates the average of the squared differences in all the respective weeks.

Step 5

In an empty cell, insert a formula that calculates the square root of the average squared difference. This value is the Root Mean Square Error.

Task 4: Finding the values of the parameters a and b using Microsoft Solver [2 marks]

Microsoft Solver in Excel allows you to perform optimisation tasks such as finding the maximum or minimum value of a formula in the spreadsheet.

In this task, you will find use the Microsoft Solver to find the values of the constants a and b such that the difference between the COVID-19 cases calculated by the exponential function and the actual number of cases is minimised i.e. the Root Mean Square Error is the minimum.

By doing so, you will identify the exponential function that fits the real data as closely as possible.

Step 1

Open Solver and set up the parameters to find a and b.

Step 2

Set the objective as the cell with the Root Mean Square Error.

Step 3

Choose 'Min' so that the objective is to find the values of a and b that will give the minimum Root Mean Square Error.

Step 4

In the 'By Changing Variable Cells' box, enter the cells containing the values of the parameters a and b.

Step 5

Select 'GRG Nonlinear' as the solving method.

- If necessary, you can click 'Add' to include constraints for parameters a and b (such as setting them as non-negative constraints).

Step 6

Click 'Solve'.

- Once the solver finds a solution, the values of a and b will be adjusted such the value for the Root Mean Square Error is the minimum.
- **Do not** round off the values of a and b displayed by the solver.

Task 5: Import the Microsoft Excel document into Google Sheet.

Name the Google Sheet using the **same name** as the Microsoft Excel document.

Your team will submit need to submit both Google Sheet and Microsoft Excel document.

Task 6: Representing the data in Charts in the Google Sheet [2 marks]

Your team will present the following in the SAME chart:

- The data given in Table 4.1
- The graph of the exponential function that was identified in **Task 4**.

Step 1

Representing the Data using a Statistical Diagram

- In the Google Sheet (**from Task 5**), insert a chart that display two graphs:
  - Graph 1: plots the number of actual new COVID-19 cases against time.
  - Graph 2: plots the number of new COVID-19 cases calculated by the mathematical model against time.

Step 2

Ensure a Smooth Curve

- Under chart editor
  - Go to the **Step up** tab and select '**Smooth line chart'**.
  - Go to **Customise** tab, navigate to '**Chart Style**' and set the layout of the chart as '**smooth**'.

Step 3

Label the Chart and Legends

- Provide an appropriate title for the chart.
- Label the x-axis as 'Week'.
- Label the y-axis as 'Number of new COVID-19 Cases'.
- Ensures that the legend clearly differentiate between the actual COVID-19 cases and the model COVID-19 cases.

Marking Rubrics

Criteria	2 marks	1 mark	0 mark
Task 1 Using Microsoft Excel to organise the data.	• Data is accurately entered and matches the source exactly. • Both column headings are labelled appropriately.	• Data is mostly accurate with minor errors or omissions. • Column headings are omitted or not labelled appropriately.	• Data contains significant errors or omission.
Task 2 Inserting formula of an exponential function to model the data.	• Correctly inserts the exponential growth formula, with cell numbers for constants a,b and cell number for exponent correctly reflected. • Formula is inserted for all rows of data.	• Inserts the exponential growth formula but with errors such as incorrect 'placement' for constants a and b.	• Significant errors in the formula (the formula itself is not the exponential growth model)
Task 3 Finding the Root Mean Square Error (RMSE).	• A column 'Squared Difference' is created, and formula is inserted correctly and applied to all rows. • Formula to find average of the squared difference is correct. • Formula to find square root of average squared formula is correct.	• At least 2 out of 3 items (refer to column 2) are correctly carried out.	• Only 1 item is carried out or none of the items are carried out (refer to column 2)
Task 4 Finding the values of the parameters a and b using Microsoft Solver. (The teacher will check by running the solver using your Microsoft Excel file)	• The values for a and b are displayed in the Microsoft Excel sheet. • The solver successfully finds and adjusts the values for a and b, such that RMSE is minimised.	• The values for a and b are displayed in the Microsoft Excel sheet. • The values for a and b are not adjusted such that RMSE is minimised.	• The values for a and b are not displayed in the Microsoft Excel sheet. • No evidence that solver was used to find a and b.
Task 6 Representing the data in Charts in Google Sheet	• In the same chart, the graphs of actual covid-19 cases and model covid-19 cases against time are both displayed. • Both graphs are smooth curves. • Chart title, axes and legend are appropriately labelled.	• The graphs of actual covid-19 cases and model covid-19 cases are displayed in separately charts. • Graphs were not smooth curves (e.g. line segments or dots) • Incomplete graph labels.	• No graphs were displayed.