Modeling the Spread of Infectious Disease

Section Workshop 4- Project 2

Names of project team members:

Assign Roles for Project Teams

Roles will be assigned based on your birthdays. Whoever has the earliest birthday in the year is assigned to role 1, next birthday is role 2 and so on. If there are only 3 members in your group, combine roles 3 and 4.

1. Discussion coordinator - ensures all are heard
2. Notes coordinator - ensures documentation in a shareable form
3. Reporter - ready to speak for the group in with your Section Workshop leader and/or other groups
4. Submission coordinator - makes sure everyone understands what is left to do (if anything!) before the submission deadline, makes plans to get together, makes plans for uploading assignments to Gradescope (this person does not necessarily need to upload but is responsible for the group coming to a consensus on a plan).

Team project

Learning Goals

- In this project you will use the tools of calculus you have learned so far to analyze a data set and draw conclusions. You will analyze data and derive a logistic differential equation that models the data. You will use qualitative analysis to determine whether your model makes sense and does a reasonable job fitting the data.
- You will use both Microsoft Excel and Desmos to do this project. Include any graphs or other tables appropriately formatted and labeled (according to the course writing rubric) directly in your REPORT.

  The content of this workshop is related to section 8.1 of the textbook, Introduction to ODE

Preview

In April 2009, a new strain of H1N1 influenza virus, referred to as pandemic influenza, was first detected in humans in the United States, followed by an outbreak in the state of Veracruz, Mexico. Soon afterwards, this new virus kept spreading worldwide resulting in a global outbreak. In China, the second Circular of the Ministry of Health pointed out that as of December 31, 2009, the country's 31 provinces had reported 120,000 confirmed cases of H1N1.

The origin of the word epidemic is Greek, where epi means “upon” and demic means “the people”. When the disease affects several geographic regions at the same, it is called a pandemic (from Greek pan meaning “all”). Epidemiology is the scientific study of contagious diseases, and it deals with the essential features of the disease, in particular, the math behind modeling the rate of infection of a population.

In this assignment, you will model and study the infection rates for the H1N1 2009 flu pandemic in Europe from weeks 36– 68 of the outbreak.

Group Report

Submit a written REPORT explaining your work and addressing the questions outlined in Parts I, II and III of this project. Your report should explain the context of the data you are examining from the H1N1 flu spread, the goals of your analysis to create a model that describes the spread, your conclusions based on your analysis and what you can learn from it. Think of this as an executive summary that you would pass to another team of researchers studying the H1N1 spread in China instead of Europe.

Your work in this report will be graded based on the analysis of the data as well as the writing in the report. You will be graded on:

- Mathematical Accuracy: Your model should be a reasonably good fit for the data and you should explain the steps in your analysis that led you to the model. You need to provide supporting graphs/plots to explain your work.
- Discussion of findings: Describes the data that you want to model. Discusses the reasons why a logistic model is appropriate. Discusses the relationship between the plots for $N(t)$ and $\frac{dN}{dt}$ including inconsistencies and what could account for them. An explanation of the model and how it was derived. A discussion of the graphical representation of the model as a slope field and whether it agrees with the logistic regression produced with Desmos.
- Coherence: Sections of the report should build on each other without repetition or gaps. Transitions should be present between sections. All material presented should be relevant to the main purpose of the report, which is to share your analysis and findings with another research team.
- Include any citations in APA format.

Your REPORT is due 1 week after your section meeting. There should be one submission per group. The written portion of the report (not including graphs and charts) should not exceed 2 pages.

Analyzing the Data

Part I – Examining the data

1. Open the spreadsheet H1N1.xls (Table 0). This data tracks the cumulative number of deaths in Europe from the H1N1 virus during weeks 36–68.

2. Create Table 1 and create a scatterplot of the shifted data.
    2.1. Make the first column in Table 1, titled "New Week", start at 0 and populate the column by subtracting 36 from each Week value in Table 0.
    2.2. Make the second column in Table 1 "Cumulative Deaths" by copying over the second column of Table 0.
3. Plot any three different trendline options in Excel to plot against the scatterplot produced from Table 1 (e.g. linear, exponential, logarithmic, or polynomial.) For this data, briefly describe why the three choices you chose are not good fits to the data based on the behavior of the function. Include the three plots and your explanation in your REPORT.
4. Briefly describe a logistic curve and why it would be a good fit for this data. Include this description in your REPORT. Notice that Excel does not have a logistic regression option (it can't produce a logistic function that "best-fits" the data.) We will use Calculus to help find a logistic regression curve.

Part II – What does N' say about N?

5. Using the data in Table 1, create a new table, Table 2 in Excel.
    5.1. Make a first column in Table 2 titled "New Week" by copying over column 1 from Table 1.
    5.2. Let $N(t)$ be the function that gives the cumulative number of deaths in week t . Add a second column to Table 2 where you approximate the "Death rate" of the disease 9the rate of change of $N(t)$ using the approximation $\frac{dN}{dt} \approx \Delta N$. Round your values to one decimal place. Leave the first cell blank as there is no prior data to find $\frac{\Delta N}{\Delta t}$.
6. Using Table 2, create a scatterplot of the data ("Death rate" vs. "New Week"). Label your axes and title the graph.
7. Explain the properties of the rate of change of $N(t)$ based on the scatterplot. Relate it to the spread of disease. Include a scatterplot in your REPORT and address the following points:
    - What is the relationship between the plot you produced from Table 1 and the plot you produced from Table 2? Use the language of derivatives and highlight any and all key points.
    - Are there inconsistencies between the analysis of the real data and what you would expect to see? What could account for this? Explain your reasoning.

Part III – Building a model

8. Create a new column table in Excel (Table 3).
    8.1. The first column header will be "N(t)[A-N(t)]," where A represents $\lim_{t\to\infty} N(t)$. Since we don't know what the limit will be in reality, we will use $A = 1261$ since it is the last value on the data table.
    8.2. Populate the first column of Table 3 using a formula in Excel to compute $N(t) \cdot \big(a - N(t)\big)$
    8.3. Population the second column of Table 3 with the growth rate $\frac{dN}{dt}$and the original data $N(t)$.
9. Create a scatterplot from Table 3 (plotting $\frac{dN}{dt}$versus$N(t) \cdot \big(A - N(t)\big)$.) Label your axes and title the graph.

10. What does the value $\left(A - N(t)\right)$ represent? Explain your answer from the point of view of $\frac{dN}{dt}$ and the original data.
11. Create a linear regression in Excel to describe the mathematical relationship between the variables on the scatterplot from Table 3. On the scatterplot, show the linear regression, its equation, and analyze the residual R. Set the intercept of the linear regression to zero (by checking the box near "display equation".) Include this graph in your REPORT.
12. Based on your findings in the previous step, derive a differential equation of the form $\frac{dN}{dt} = \cdots$(Hint: think about what the input and output are representing in the previous question.) Include your model and an explanation of how you got it in your REPORT.

Part IV – Does your model make sense?

13. Go to Desmos and use the slope field application to analyze your model. Make a copy of this graph for your group to edit.
14. Use the sliders to set your values of r and A from your differential equation (this might take some algebra on your part!)
15. Copy the data from Table 1 and pate it into Desmos to produce a scatterplot of the data on top of the slope fields produced based on your differential equation.
16. **CHECKPOINT:** Does this scatterplot make sense based on the general shape of solutions in the slope field? If yes, congratulations; your differential equation has done a reasonable job of modeling the data! If not, revisit the work you did to produce the differential equation leading up to question 12.
17. Unlike Excel, Desmos can do logistic regressions. In Desmos, below the table of data, enter the following formula using x_1 and y_1, which will model the data using the logistic curve equation

$$N(t) = \frac{A}{1 + Be^{rt}}$$

18. Desribe Desmo's regression function qualitatively. Is it a good fit for the data? In you rREPORT, describe and interpret its behavior as t tends to infinity. Include a screenshot of the slope field your produced with the data and regression plotted over it.

Individual Project for Final Portfolio

1. Focus on Writing: Watch the first few minutes of an interview of US Surgeon General Dr Jerome Adams (https://www.youtube.com/watch?v=HK2ypT2xweA). Around minute 2:15, he says "...we can get America to flatten the curve". What does "flatten the curve" mean in terms of Calculus concepts? Be as technical and accurate as possible, using all relevant calculus definitions, examples, and notation. Include graphs and formulas as needed. Be sure to include APA formatted citations.
   (Hint: It might be helpful to first watch the "All Around Math Guy" as he relates the analysis you did in the group project to the idea of "flattening the curve." You can watch a YouTube video that discusses the relationship here and highlights data collected from early in the COVID-19 pandemic (March 2020.))
2. Focus on Writing Suppose you are a spokesperson for the Baltimore City Department of Public Health. You are charged with welcoming new community health workers working as contact tracers and care coordinators. Calculus is not required for the position so assume this group has

never taken calculus before. The team does not need a technical understanding of calculus concepts, but they do need to understand the importance of their role in "flattening the curve". Write a one-page memo to distribute to new employees. Include an argument in support of data driven analysis, the differences in modeling with a logistic vs. exponential curve, the concept of "flattening the curve", and an interpretation of the inflection point. Specifically write for this audience. Do not use any calculus definitions or notations.

Further Reading and Resources

Papers

The two references below are papers that discuss the spread of H1N1 in 2009.

JIN Z, ZHANG J, SONG LP, SUN GQ, KAN J, ZHU H. MODELLING AND ANALYSIS OF INFLUENZA A (H1N1) ON NETWORKS. BMC PUBLIC HEALTH. 2011;11 SUPPL 1(SUPPL 1):S9. PUBLISHED 2011 FEB 25. DOI:10.1186/1471-2458-11-S1-S9

SAMANLIOGLU F, BILGE AH. AN OVERVIEW OF THE 2009 A(H1N1) PANDEMIC IN EUROPE: EFFICIENCY OF THE VACCINATION AND HEALTHCARE STRATEGIES. J HEALTHCARE ENG. 2016;2016:5965836. DOI:10.1155/2016/5965836

Desmos Support

For help with logistic models, reference this illustrative example that uses sample data: https://www.desmos.com/calculator/naf1qogfjn.

For extra help creating regressions visit the regression page on Desmos: https://learn.desmos.com/regressions.

We have created a slope field application for you to use to analyze your model: https://www.desmos.com/calculator/a5p9qv8xtm

Excel Support

Important Note: Make sure that you open the data spreadsheet in Excel or Office 365; you can access Office365 via https://portal.office.com/ and log in with your JHU credentials. Do not use Pages on Mac or GoogleSheets because the hints embedded below will not work for you!)

Adding trend lines: https://support.microsoft.com/en-us/office/add-a-trend-or-moving-average-line-to-a-chart-fa59f86c-5852-4b68-a6d4-901a745842ad

Flattening the curve

Mr. Eric Balzarini of Walnut Grove Secondary School in Langley, BC, Canada, is also known as "The all around Math guy." His YouTube video https://www.youtube.com/watch?v=CUzcuiRE15c discussion will help with the individual portion of the project.

This assignment was prepared by Emily Braley and Joe Cutrone, Johns Hopkins University Mathematics, Spring 2022.