

MONT 104N – Modeling the Environment
Chapter 5 Project
October 29 and 31, 2018

Background Information

In the text, we analyzed data from a large 2014-2015 Ebola virus disease outbreak in West Africa that you may recall. The Chapter Project for this chapter will be to look at recent data from another outbreak of this disease in the Kivu province of the Democratic Republic of Congo that is currently underway. The data we will use there comes from World Health Organization data reported on the Wikipedia page

https://en.wikipedia.org/wiki/2018_Kivu_Democratic_Republic_of_the_Congo_Ebola_virus_outbreak

On that page, you want to look for the table labelled *Timeline of reported cases and deaths*. For the project, you want to start out by getting all of this data into Excel or another spreadsheet program. You will need to convert the dates to a numerical form (days after the initial report of Ebola cases on August 1, 2018).

Questions

Most of the work you will do will involve analyzing subsets of the data along the lines of what we did in the example in the text, to try to develop answers to the following questions.

- (A) Some of the “Suspected” cases are later determined *not* to be Ebola cases and others are never verified, so let’s ignore that column since it is too speculative. Instead of the values in the Total column, let’s use the sum of the Confirmed and Probable numbers as the number of “Likely Cases.” Note that these are *cumulative totals*. This means, for instance, that the $25 + 27 = 52$ likely cases as of August 8, 2018 include the $16 + 27 = 43$ cases as of August 5, 2018, and so forth. Similarly, the 39 deaths as of August 8 include the 34 as of August 5, and so forth. Fit linear and exponential models to the Likely Cases and the Deaths as functions of the number of days after August 1. (You may want to consult Examples 5.10 and 5.11 in the text, which deal with the earlier 2014-2016 outbreak, for pointers on how to approach this.)
- (B) If you look carefully at the data from 2014-2016 you will notice that roughly 60 days after the start, after a slower start, that epidemic really started to “take off.” and reach growth rates of about 2% per day in both Cases and Deaths(!). Does something like that seem to be happening in 2018? When? (You’ll need to experiment by fitting exponential models to ranges of days starting at August 1, and going up to different ending dates, then starting at different dates and going to the end of the data to investigate this).
- (C) The CFR column records the *case fatality rate*—a percent of the total cases in some population that ended in the death of the patient. Which population was used to produce these figures?

(Hint: it's *not* the percentage of Deaths out of the figure in the Total column.) Is there a definite upward or downward trend in the CFR figures over the course of the epidemic so far? What if you leave out the first three days where you have CFR figures (maybe they're unrepresentative because they are at the very start of the epidemic)? Notes:

1. You will want to omit the days for which no CFR figure is reported.
 2. Next semester we will discuss a way to decide whether a slope of a regression line is *significantly different from zero*. Small slopes with very small R^2 values are generally not significantly different from zero.
- (D) Many of the recent news reports concerning this outbreak have emphasized one very difficult aspect of the situation. Namely, the Kivu province where this is happening is also in the middle of an armed civil conflict between rebel groups and the Congolese central government. There are large numbers of refugees displaced from their homes and an ongoing humanitarian crisis in the area because of this. This is making it difficult or even impossible for public health workers to work in the area and carry out the usual sorts of isolation and testing of contacts that are used to bring outbreaks of infectious diseases under control in less challenging circumstances. Note: I believe the figure reported in the Contacts column represents the number of contacts of known cases that have yet to be A number of recent reports have included statements to the effect that public health officials are concerned that the outbreak is “out of control.”¹ Based on your analysis, would you say that this concern is justified?

Assignment

Submit a spreadsheet showing all of your calculations with the Ebola data set. You may answer the questions above in a separate text document or in your spreadsheet. As always, document your sources if you consult things other than the Wikipedia page given above.

¹For instance, see <https://www.statnews.com/2018/10/02/officials-fear-ebola-epidemic-may-be-spinning-beyond-their-control>.