Case Study 0: 1995 DRC Ebola outbreak

Danny Scarponi

Contents

Chapter 1

Introduction

This is the first of a series of tutorials that illustrate the process for using the Bayes Linear History Matching with Emulation (HME) package emulatorr.

Chapter 2

Description of epidemiological problem

The case study is based on an outbreak of Ebola Hemorrhagic Fever that took place in Kikwit, Democratic Republic of the Congo, in 1995. We give here a brief description of the circumstances in which the epidemic originated and of the measures put in place to characterize and contain it. More details about the outbreak can be found in the article The Reemergence of Ebola Hemorrhagic Fever, Democratic Republic of the Congo, 1995, by Khan et al.

In May 1995, the Centers for Disease Control and Prevention (CDC) was notified of an outbreak of viral hemorrhagic fever in Kikwit, DRC. Specimens sent to Belgium were forwarded to CDC for diagnostic testing, and the presence of acute or recent EBO virus infection was confirmed for all 14 persons tested. Within 48 h of diagnostic confirmation, a multinational contingent led by the World Health Organization (WHO), composed of physicians, epidemiologists, sanitarians, health educators, and logistic support personnel were dispatched to assist DRC authorities in controlling and characterizing the outbreak.

Index Patient and Initial Cluster of Cases The first identified case-patient was GM, a 42-year-old male charcoal worker and farmer who became ill on 6 January and died of a febrile hemorrhagic disease at Kikwit General Hospital on 13 January 1995. He directly infected at least 3 members of his family, all of whom died, and an additional 10 secondary cases (all fatal) occurred among members of his extended family over the next 9 weeks in an area encompassing Kikwit and 3 surrounding villages.

In ealy April1995, a small nosocomial cluster of EBO cases among the nursing staff at Kikwit II Maternity Hospital was first recognized but misdiagnosed as epidemic dysentery. Toward the end of April, a similar nosocomial cluster was identified at Kikwit General Hospital among the operating room staff who participated in a surgical procedure on a laboratory technician employed at Kikwit II Maternity Hospital.

Surveillance and case-finding

Time course of the Epidemic At least four generations of cases were traced through the hospital

and into the community from the initial nosocomial cases in Kikwit General Hospital (figure 1).

Chapter 3

Description of model and input parameters

We consider a stochastic (continuous/discrete time?) SEIR model. The human host population is divided into the categories of susceptible (S), infected, but not yet infectious (E), infectious (I) and removed (R). Assuming a homogeneous closed population of size N (i.e. with negligible births, deaths or migration) and homogeneous mixing within the population, the rate of transition of an individual from $S \to E$ can be modelled as being proportional to the proportion of infective individuals in the population. The rates of transition from $E \to I$ and $I \to R$ are constant.

If S_t, E_t, I_t and R_t are the numbers of susceptible, infected but not yet infectious, infectious and removed individuals in the population at time t, and dt is a vanishingly small time period such that only one event can occur in $[t, t + dt]$, then we can write the probabilities of events within the population as

$$P(S \to E) = \beta N^{-1} S_t I_t dt + o(dt) \tag{3.1}$$

$$P(E \to I) = \delta E_t dt + o(dt) \tag{3.2}$$

$$P(I \to R) = \gamma I_t dt + o(dt) \tag{3.3}$$

The notation $S \to E$ corresponds to the movement of a single individual from the state S to the state E (likewise for the other possible transitions). Here β is the transition parameter, δ^{-1} is the mean incubation period and γ^{-1} is the mean infectious period. While we assume δ and γ to be constant, we will let β be time-dependent. This is because we need to account for the intervention strategies put in place to control the spread of Ebola. We assume that the transmission rate is constant up to the point of intervention t_{int}, before decaying exponentially to zero:

$$\beta_t = \begin{cases} \beta & t < t_{\text{int}} \\ \beta e^{-q(t - t_{\text{int}})} & t \geq t_{\text{int}}. \end{cases} \tag{3.4}$$

Note that the γ is not affected by the interventions, since the disease is not curable.

The epidemic model specified by (3.1)-(3.4) has parameter vector $\Theta = \{\beta, q, \gamma, \delta\}$, which we would like to estimate from the available data.

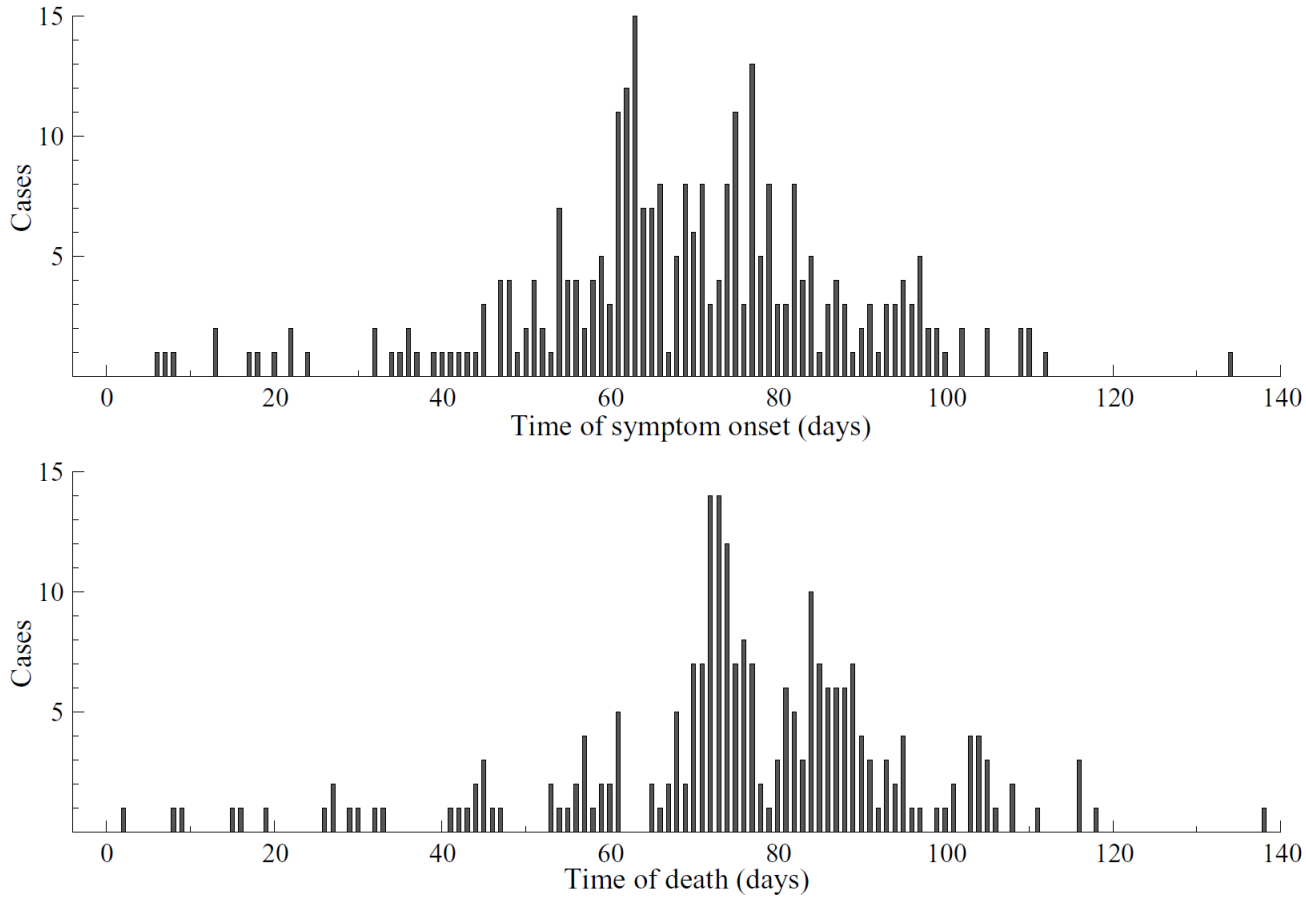
Chapter 4

Description of fitting outputs

4.1 Description of data on 1995 Ebola DRC outbreak

The data consist of two time series (see Figure below) recorded from March 1 to July 16, namely, daily counts of Ebola cases by date of symptom onset, accounting for a total of 291 cases, and daily counts of deaths from Ebola, accounting for a total of 236 deaths. It is also documented that the first case became ill on January 6, 1995, the last case died on July 16, and a total of 316 cases were identified resulting in a rate of 81% fatality.

The epidemic lasted for about 200 days with control measures being introduced about 130 days after the start of the epidemic. The exact starting time and evolution of the epidemic prior to March 1 is unobserved. Furthermore, from the total number of 316 identified cases, it can be deduced that the dates of symptom onset for 25 cases and the dates of removal from the infectious class for 80 cases are not reported in the given time series.



4.2 Missing data handling

Chapter 5

Description of calibration methodology

5.1 Construction of emulators

5.1.1 General structure of an emulator (role of various parameters)

5.1.2 Sampling points (training/validation points)

5.1.3 Training first emulators

5.2 Emulator diagnostics

5.2.1 Implausibility measure (choice of cut-off)

5.2.2 Various forms of diagnostics

5.2.2.1 Simulator/emulator outputs comparison

5.2.2.2 Standard error of emulator outputs

5.2.2.3 Emulator/simulator implausibility

5.3 Points generation

5.4 Further waves

Appendix A

Answers

Solution 4

```
2 + 2
```

```
## [1] 4
```

Return to task on P??

Answer 5

Why 42 of course!

Return to task on P??

Solution 7

Base R

```
## base R solution
mean(iris$Petal.Length[
    iris$Species == "setosa"])
```

```
## [1] 1.462
```

tidyverse

```
## tidyverse solution
iris %>%
    filter(Species == "setosa") %>%
    select(Petal.Length) %>%
    summarise(mean = mean(Petal.Length))
```

```
##    mean
## 1 1.462
```

Return to task on P??