Mathematics 4MB3/6MB3 Mathematical Biology
2016 ASSIGNMENT 4

Group Name: Model Students

Group Members: Nicole Dumont, Melody Fong, Carolina Weishaar

This assignment was due on Monday 13 March 2017 at 11:30am.

1 Time Series analysis of Recurrent Epidemics

(a) You should have received the following data files by e-mail:

`meas_uk__lon_1944-94_wk.csv`
`meas_uk__lpl_1944-94_wk.csv`

These plain text comma-separated-value files list weekly cases of measles (in London and Liverpool, England, from 1944 to 1994). Depending on which research direction you select, you might receive other files in the same `ymdc` (year,month,day,count) format, where the count column might contain cases or deaths, for example. Write the following R functions:

(i) `read.ymdc()`. Read a file in `ymdc` format and return a data frame containing these data and including a `date` column that has R's `Date` class. The first (and potentially only) argument to this function should be the `filename` of the data file to be read.

Solution.
```
read.ymdc <- function(filename){
  data <- read.csv(filename, skip=6)
  ymdcdata <- NULL

  ymdcdata$date <- as.Date(paste(as.character(data$year), as.character(data$mont
  ymdcdata$cases <- data$cases
  ymdcdata <- as.data.frame(ymdcdata)
  return(ymdcdata)
}
```

□

(ii) `time.plot()`. Given a data frame produced using `read.ymdc()`, display the associated time plot. The first argument of the function should be the data frame. Further optional argument(s) should allow the user to smooth the time series with a moving average. By default, this function should create a new plot but there should be an option to add to an existing plot. Implement this by having a logical

`add` argument that is false by default (`add=FALSE`). This will allow you to add a smoothed version of the time series on top of the raw data, for example. The final argument should be the ellipsis (`...`) so that details such as colour and line style can be passed to the plotting commands used in this function.

Solution.
```
time.plot <- function(ymdcdata, ma.smooth=FALSE, sides = 1, add=FALSE, .
  if (dev.cur() == 1L && !identical(add, FALSE)) {
        warning("'add' will be ignored as there is no existing plot")
        add <- FALSE
  }
  #Moving average smoothing
  if(isTRUE(ma.smooth)){
    olddata <- ymdcdata
    #Replace middle values by moving average
    for (i in seq(sides+1,length(olddata$cases)-sides,1)){
      ymdcdata$cases[i] <- mean(olddata$cases[i-sides:i+sides])
    }
    #Replace ends of data with NAs
    ymdcdata$cases[1:sides] <- NA
    ymdcdata$cases[sides:length(olddata$cases)-sides+1] <- NA
  }
  if (isTRUE(add)) {
    lines(cases~date,ymdcdata, ...)
  }else plot(cases~date,ymdcdata, ...)
}
```

□

(iii) `periodogram()`. Given a data frame produced using `read.ymdc()`, display the associated *period periodogram* (power spectrum as a function of period). The first argument of the function should be the data frame. By default, the entire time series should be used, but optional argument(s) should allow the user to specify a time range of interest. Use R's `spectrum()` function to compute the power spectrum. Have `add` and `...` arguments as in `time.plot()`. Note that if `v` is a vector containing a time series of interest, you can obtain and plot its *frequency* periodogram as follows.

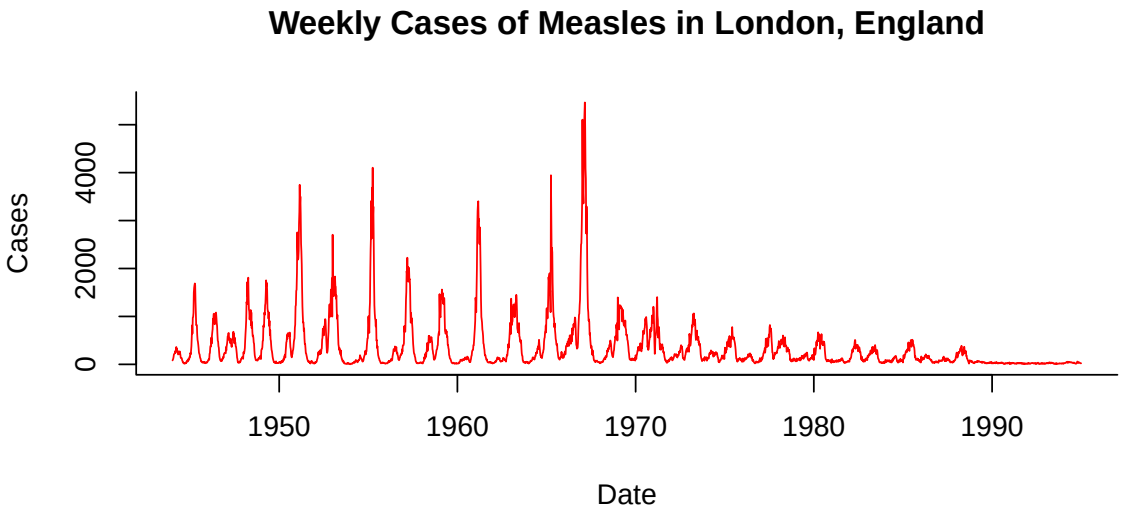
```
v <- c()
s <- spectrum(v, plot=FALSE)
plot( s$freq, s$spec, type="l")
```

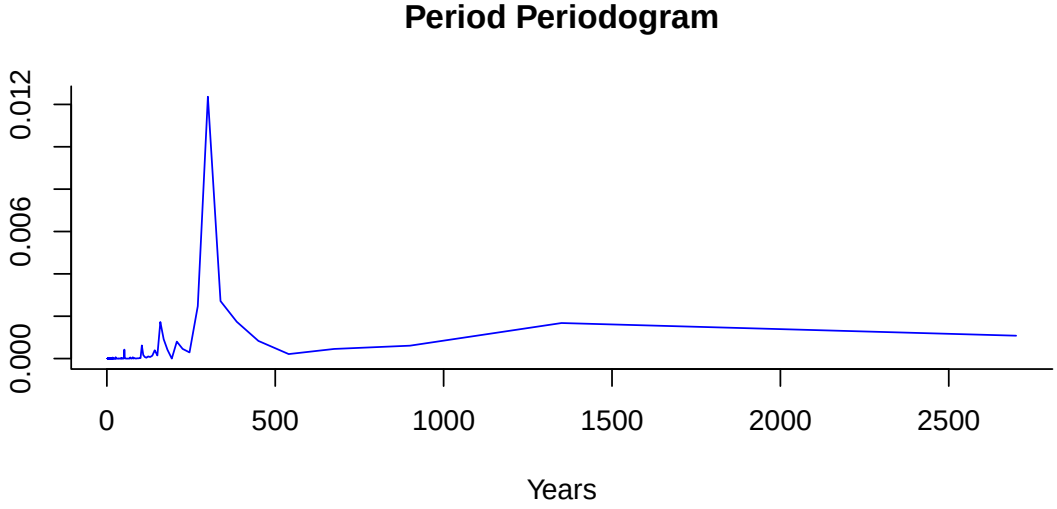
Solution.
```
periodogram <- function(ymdcdata, timerange=1:length(ymdcdata$date), add=FAL
  if (dev.cur() == 1L && !identical(add, FALSE)) {
        warning("'add' will be ignored as there is no existing plot")
        add <- FALSE
  }

  #Set data outside timerange to NAs
  ymdcdata[setdiff(1:length(ymdcdata$date),timerange),] <- NA


  ycdata <- ymdcdata
  ycdata$date <- as.numeric(ycdata$date-ycdata$date[1])/365

  s <- spectrum(ycdata, plot=FALSE)

  if (isTRUE(add)) {lon.data
    lines(1/(s$freq), s$spec, ...)
  }else plot(1/(s$freq), s$spec[,1], ...)

}
```

□

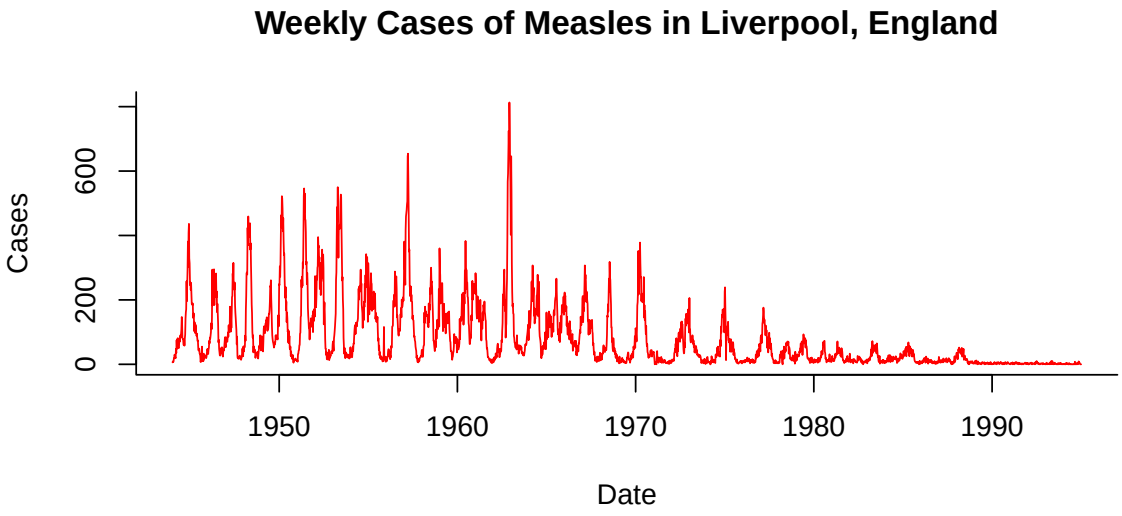
(b) Using your functions, make a multi-panel plot that clearly shows the temporal pattern of the time series and how its frequency structure changes over time. Think carefully about how to make this multi-panel figure as clear as possible for yourselves and your readers. Describe your figure, explaining what aspects of your figure you feel are puzzling or interesting and may be possible to understand using mechanistic mathematical modelling. (Repeat this for each of the epidemic time series you are given.)

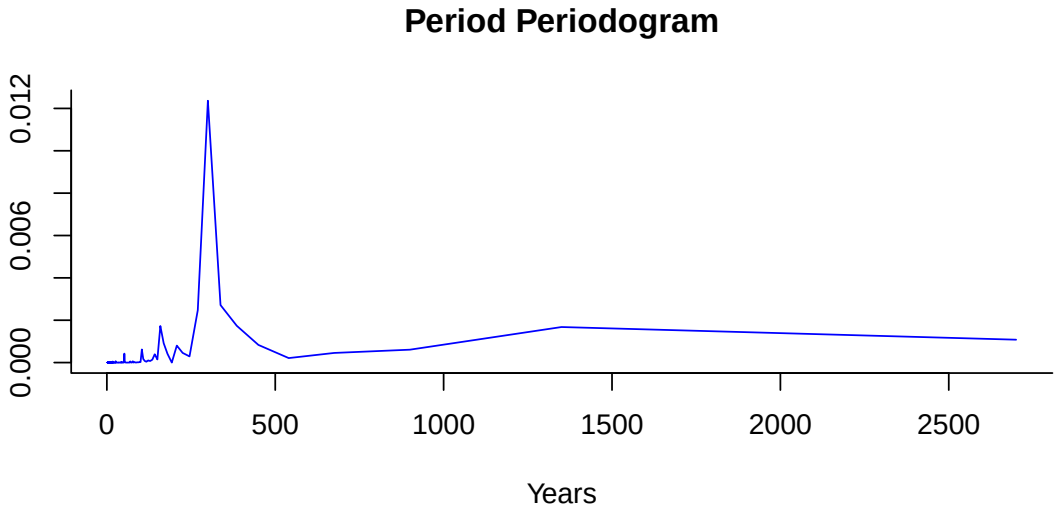
Solution.
```
lon.data <- read.ymdc("meas_uk__lon_1944-94_wk.csv")
par(mfrow=c(2,1))
time.plot(lon.data, ma.smooth=FALSE, add=FALSE, type="l", xlab="Date", ylab="Cases",
periodogram(lon.data, add=FALSE, type="l", xlab="Years", ylab="", main="Period Perioc
```






```
lpl.data <- read.ymdc("meas_uk__lpl_1944-94_wk.csv")
par(mfrow=c(2,1))
time.plot(lpl.data, ma.smooth=FALSE, add=FALSE, type="l", xlab="Date", ylab="Cases",
periodogram(lpl.data, add=FALSE, type="l", xlab="Years", ylab="", main="Period Period
```





□

2 Stochastic Epidemic Simulations

Consider the SI model,

$$\frac{dI}{dt} = \beta(N - I)I\,, \qquad I(0) = I_0, \tag{1}$$

where β is the transmission rate, N is the population size and $I(t)$ is the number of infected individuals at time t.

(a) Write an R function `SI.Gillespie()` that uses the Gillespie algorithm to produce a realization of a stochastic process whose mean field dynamics are given by equation (1)

in the limit $N \to \infty$. Your function should have arguments `beta`, `N`, `I0` and `tmax` (the time at which to end the simulation). You may find it helpful (conceptually) to write equation (1) in two-variable form:

$$\frac{dS}{dt} = -\beta SI\,, \qquad S(0) = N - I_0, \tag{2a}$$

$$\frac{dI}{dt} = \beta SI\,, \qquad I(0) = I_0. \tag{2b}$$

Note that there is only one type of event that can occur, so the second part of the Gillespie algorithm (what type of event occurred) is trivial for this model.

Note: To make stochastic simulations exactly reproducible use `set.seed()`.

Solution.
```
SI.Gillespie <- function(beta, N, I0, tmax){
  timeseries <- c()
  timeseries <- rbind(timeseries, c(0,I0))
  index <- 1
  while(timeseries[index,1] < tmax){
    if (timeseries[index,2] < N){
      rate <- beta*(timeseries[index,2])*(N-timeseries[index,2])
      timestep <- rexp(n=1, rate= rate)
      timeseries <- rbind(timeseries, c(timeseries[index,1]+timestep, timeseries[inde
    }else{
      timeseries <- rbind(timeseries, c(timeseries[index,1]+timestep*3, timeseries[in
    }
    index <- index+1
  }
  return(timeseries)
}
```

□

(b) Make a multi-panel plot comparing the deterministic and stochastic dynamics of the SI model for $\beta = 1$, $I_0 = 1$ and $N \in \{32, 10^2, 10^3, 10^4\}$ ($N = 32$ is close to $10^{1.5}$). Each panel should correspond to a different value of N and should show 30 stochastic realizations together with the deterministic solution.

Solution.
```
## Vector Field for SI model
SI.vector.field <- function(t, vars, parms=c(beta=2,gamma=1)) {
  with(as.list(c(parms, vars)), {
    dx <- -beta*x*y # dS/dt
    dy <- beta*x*y  # dI/dt
    vec.fld <- c(dx=dx, dy=dy)
    return(list(vec.fld)) # ode() requires a list
```

```
  })
}
```

```
## Draw solution
library("deSolve")
draw.soln <- function(ic=c(x=1,y=0), tmax=1,
                      times=seq(0,tmax,by=tmax/500),
                      func, parms, ... ) {
  soln <- ode(ic, times, func, parms)
  lines(times, soln[,"y"], col="blue", lwd=3, ... )
}
```

```
beta <- 1
I0 <- 1
tmax <- 0.3
N <- c(32, 10^2, 10^3, 10^4)

par(mfrow=c(length(N),1))
for (j in 1:length(N)){
  ## draw box for plot
  plot(0,0,xlim=c(0,tmax),ylim=c(0,N[j]+1),xlab="Time", ylab="Prevalence (I)", main=p
  line <- c(1:30)
  for (i in 1:length(line)) {
    set.seed(line[i])
    lines(SI.Gillespie(beta, N[j], I0, tmax),lty="dotted",
            col=line[i+1]) # use a different line colour for each solution
  }
  draw.soln(ic=c(x=N[j]-I0,y=I0), tmax=tmax,
            func= SI.vector.field,
            parms=c(beta,gamma=1))
}
```

□

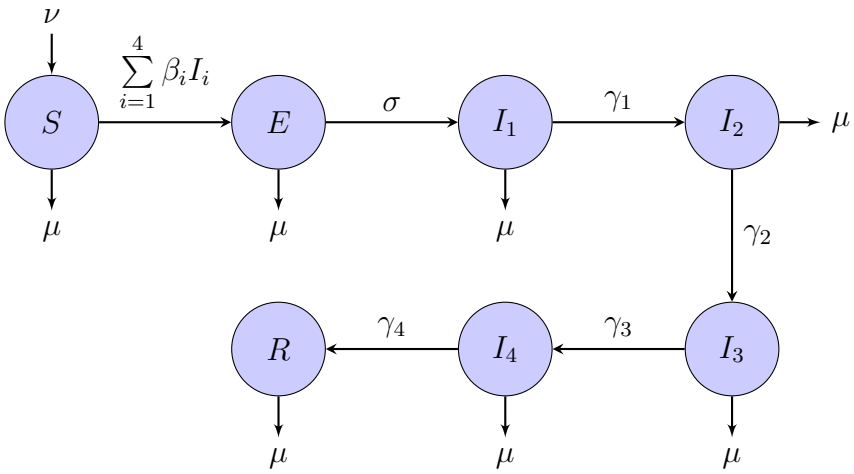
3 $\mathcal{R}_0$ for smallpox

The natural history of smallpox is shown in Figure 1. The US Centers for Disease Control and Prevention (CDC) has recently discovered that a group of bioterrorists plans to reintroduce smallpox to the United States. The CDC has reason to believe that the terrorists are also bioengineers and have successfully altered the virus so that it causes the early rash stage to be twice as long as it was when the virus was last circulating naturally in the 1970s. Moreover, the existing smallpox vaccine apparently provides no protection against the altered virus.

The CDC wants your opinion on how the alterations to the virus will affect $\mathcal{R}_0$ and the expected final size of an epidemic if the planned attack is successful.

(a) Construct a compartmental (ODE) smallpox transmission model based on the natural history specified in Figure 1, including vital dynamics but ignoring disease-induced death.

Solution. A compartmental (ODE) smallpox transmission model based on the natural history and assuming time spent in each stage of the disease is expontentially distributed is given by



$$\frac{dS}{dt} = \nu - S\sum_{i=1}^{4} \beta_i I_i - \mu S \tag{3}$$

$$\frac{dE}{dt} = S\sum_{i=1}^{4} \beta_i I_i - \sigma E - \mu E \tag{4}$$

$$\frac{dI_1}{dt} = \sigma E - \gamma_1 I_1 - \mu I_1 \tag{5}$$

$$\dots \tag{6}$$

$$\frac{dI_i}{dt} = \gamma_{i-1} I_{i-1} - \gamma_i I_i - \mu I_i, \qquad 0 < i \leq 4 \tag{7}$$

$$\dots \tag{8}$$

$$\frac{dR}{dt} = \gamma_5 I_5 - \mu R \tag{9}$$

The notation of the compartments is as follows

Notation	Definition
S	Proportion of population that is susceptible
E	Proportion of population in the incubation stage
I_1	Proportion of population in the prodrom stage
I_2	Proportion of population in the early rash stage
I_3	Proportion of population in the pustular rash stage
I_4	Proportion of population in the scarbs stage
R	Proportion of population that is recovered

The notation of the rates and waiting times is as follows

Notation	Definition	Value for original virus
ν	Birth rate	
β_i	Transmission rate due to individuals from compartment I_i	
$\frac{1}{\sigma}$	Mean time spent in compartment E	12 days
$\frac{1}{\gamma_1}$	Mean time spent in compartment I_1	3 days
$\frac{1}{\gamma_2}$	Mean time spent in compartment I_2	4 days
$\frac{1}{\gamma_3}$	Mean time spent in compartment I_3	5 days
$\frac{1}{\gamma_4}$	Mean time spent in compartment I_4	11 days
μ	Death rate	

□

(b) Use a biological argument to find a formula for $\mathcal{R}_0$.

Solution. . . . beautifully clear and concise text to be inserted here. . . □

(c) Calculate $\mathcal{R}_0$ using the next generation matrix approach. *Note:* Your solution should include $\mathcal{F}$, $\mathcal{V}$, F, V, and FV^{-1}, in the most human-friendly form you can find. However, feel free to use symbolic manipulation software such as `Maple`, *Mathematica* or sage to help with the necessary algebra and matrix computations.

Solution.

$$\frac{d}{dt}\begin{bmatrix} E \\ I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} S\sum_{i=1}^{4}\beta_i I_i - \sigma E - \mu E \\ \sigma E - \gamma_1 I_1 - \mu I_1 \\ \gamma_1 I_1 - \gamma_2 I_2 - \mu I_2 \\ \gamma_2 I_2 - \gamma_3 I_3 - \mu I_3 \\ \gamma_3 I_3 - \gamma_4 I_4 - \mu I_4 \end{bmatrix} = \mathcal{F} - \mathcal{V}$$

$$\mathcal{F} = \text{The inflow of new infected into infected compartments} = \begin{bmatrix} S\sum_{i=1}^{4}\beta_i I_i \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathcal{V} = \text{The outflow of from infected compartments} = \begin{bmatrix} \sigma E + \mu E \\ -\sigma E + \gamma_1 I_1 + \mu I_1 \\ -\gamma_1 I_1 + \gamma_2 I_2 + \mu I_2 \\ -\gamma_2 I_2 + \gamma_3 I_3 + \mu I_3 \\ -\gamma_3 I_3 + \gamma_4 I_4 + \mu I_4 \end{bmatrix}$$

The Jacobians of these vectors at the disease free equilibrium, $(S^*, I_1^*, \ldots, I_4^*) = (1, 0, \ldots, 0)$, are

$$F = D\mathcal{F}_{(S^*, I_i^*)} = \begin{bmatrix} 0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$V = D\mathcal{V}_{(S^*, I_i^*)} = \begin{bmatrix} \sigma + \mu & 0 & 0 & 0 & 0 \\ -\sigma & \gamma_1 + \mu & 0 & 0 & 0 \\ 0 & -\gamma_1 & \gamma_2 + \mu & 0 & 0 \\ 0 & 0 & -\gamma_2 & \gamma_3 + \mu & 0 \\ 0 & 0 & 0 & -\gamma_3 & \gamma_4 + \mu \end{bmatrix}$$

The inverse of V is

$$V^{-1} = \begin{bmatrix} \frac{1}{\sigma+\mu} & 0 & 0 & 0 & 0 \\ \frac{\sigma}{(\sigma+\mu)(\gamma_1+\mu)} & \frac{1}{\gamma_1+\mu} & 0 & 0 & 0 \\ \frac{\sigma\gamma_1}{(\sigma+\mu)(\gamma_1+\mu)(\gamma_2+\mu)} & \frac{\gamma_1}{(\gamma_1+\mu)(\gamma_2+\mu)} & \frac{1}{\gamma_2+\mu} & 0 & 0 \\ \frac{\sigma\gamma_1\gamma_2}{(\sigma+\mu)(\gamma_1+\mu)(\gamma_2+\mu)(\gamma_3+\mu)} & \frac{\gamma_1\gamma_2}{(\gamma_1+\mu)(\gamma_2+\mu)(\gamma_3+\mu)} & \frac{\gamma_2}{(\gamma_2+\mu)(\gamma_3+\mu)} & \frac{1}{\gamma_3+\mu} & 0 \\ \frac{\sigma\gamma_1\gamma_2\gamma_3}{(\sigma+\mu)(\gamma_1+\mu)(\gamma_2+\mu)(\gamma_3+\mu)(\gamma_4+\mu)} & \frac{\gamma_1\gamma_2\gamma_3}{(\gamma_1+\mu)(\gamma_2+\mu)(\gamma_3+\mu)(\gamma_4+\mu)} & \frac{\gamma_2\gamma_3}{(\gamma_2+\mu)(\gamma_3+\mu)(\gamma_4+\mu)} & \frac{\gamma_3}{(\gamma_3+\mu)(\gamma_4+\mu)} & \frac{1}{\gamma_4+\mu} \end{bmatrix}$$

The next generation matrix (computed using Mathematica) is

$$FV^{-1} = \begin{bmatrix} & & A^T & & \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where

$$A = \begin{bmatrix} \frac{\sigma}{(\sigma+\mu)(\gamma_1+\mu)}\left(\beta_1 + \frac{\beta_2\gamma_1}{\gamma_2+\mu} + \frac{\beta_3\gamma_1\gamma_2}{(\gamma_2+\mu)(\gamma_3+\mu)} + \frac{\beta_4\gamma_1\gamma_2\gamma_3}{(\gamma_2+\mu)(\gamma_3+\mu)(\gamma_4+\mu)}\right) \\ \frac{1}{\gamma_1+\mu}\left(\beta_1 + \frac{\beta_2\gamma_1}{\gamma_2+\mu} + \frac{\beta_3\gamma_1\gamma_2}{(\gamma_2+\mu)(\gamma_3+\mu)} + \frac{\beta_4\gamma_1\gamma_2\gamma_3}{(\gamma_2+\mu)(\gamma_3+\mu)(\gamma_4+\mu)}\right) \\ \frac{1}{\gamma_2+\mu}\left(\beta_2 + \frac{\beta_3\gamma_2}{\gamma_3+\mu} + \frac{\beta_4\gamma_2\gamma_3}{(\gamma_3+\mu)(\gamma_4+\mu)}\right) \\ \frac{1}{\gamma_3+\mu}\left(\beta_3 + \frac{\beta_4\gamma_3}{\gamma_4+\mu}\right) \\ \frac{\beta_4}{\gamma_4+\mu} \end{bmatrix}$$

$\mathcal{R}_0$ is the spectral radius - i.e. the maximum eigenvalue - of this matrix. Using Mathematica to compute this:

$$\mathcal{R}_0 = \rho(FV^{-1}) \tag{10}$$

$$= \frac{\sigma}{(\sigma+\mu)(\gamma_1+\mu)(\gamma_2+\mu)}\left(\beta_1(\gamma_2+\mu) + \frac{\beta_2\gamma_1(\gamma_3+\mu)}{\gamma_3+\mu} + \frac{\beta_3\gamma_1\gamma_2(\gamma_4+\mu)}{(\gamma_3+\mu)(\gamma_4+\mu)} + \frac{\beta_4\gamma_1\gamma_2\gamma_3}{(\gamma_3+\mu)(\gamma_4+\mu)}\right) \tag{11}$$

$$= \frac{\sigma}{(\sigma+\mu)(\gamma_1+\mu)(\gamma_2+\mu)(\gamma_3+\mu)(\gamma_4+\mu)}(\beta_1(\gamma_2+\mu)(\gamma_3+\mu)(\gamma_4+\mu)$$
$$+ \beta_2\gamma_1(\gamma_3+\mu)(\gamma_4+\mu) + \beta_3\gamma_1\gamma_2(\gamma_4+\mu) + \beta_4\gamma_1\gamma_2\gamma_3) \tag{12}$$

□

(d) Based on your model, and $\mathcal{R}_0 \sim 5$ for unaltered smallpox, what can you say about the difference in $\mathcal{R}_0$ that can be expected for the newly engineered virus vs. the original virus?

Solution. Let $\tilde{\mathcal{R}}_0$ be the value of $\mathcal{R}_0$ for the newly engineered virus. All of the parameters are the same as for unaltered smallpox except $\frac{1}{\tilde{\gamma}_2} = \frac{2}{\gamma_2}$. □

(e) Write a paragraph that you can imagine e-mailing to the CDC, in which you do your best to answer their questions.

Solution. . . . beautifully clear and concise text to be inserted here. . . □

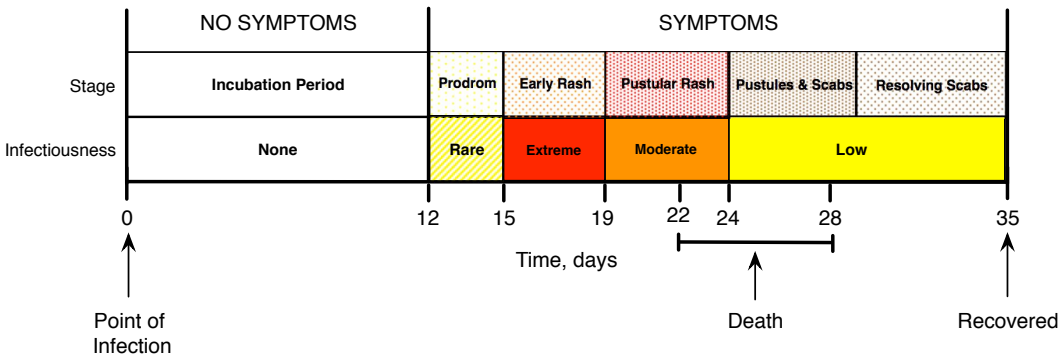


Figure 1: The natural history of smallpox infection. The prodrom stage begins with fever but the patient is very rarely contagious. Early rash is the most contagious stage, when the rash develops and transforms into bumps. During the pustular rash stage bumps become pustules, which then turn into scabs during the pustules and scabs stage and fall off during the resolving scabs stage. The infected person is contagious until the last scab falls off. (*This is Figure 3.4 from page 82 of Olga Krylova's 2011 McMaster University PhD thesis.*)

— END OF ASSIGNMENT —

Compile time for this document: March 10, 2017 @ 22:25