SISMID Module 6: Stochastic Epidemic Models with Inference – Exercise Session 2

Dongni Zhang, Stockholm University

July 13, 2021

Introductions to Exercise 2.1

Estimation of R_0 (a)

- Assume a homogeneous mixing population and all individuals are initially susceptible.
- No prevention measures.
- In case of a large outbreak, we observe that a fraction $\tilde{\tau}$ get infected.

The estimate of R_0 is given by the observed value:

$$\hat{R}_0 = -\ln(1 - \tilde{\tau})/\tilde{\tau}.$$

Estimation of R_0 (b)

Now if we know that a fraction r was initially immune, and there were a fraction $\tau_{overall}$ infected during the outbreak.

- The fraction infected among those initially susceptibles $\tilde{\tau} = \tau_{overall}/(1-r)$.
- The estimate of R_0 is now given by

$$\hat{R}_0 = -\ln(1-\tilde{\tau})/(1-r)\tilde{\tau}.$$

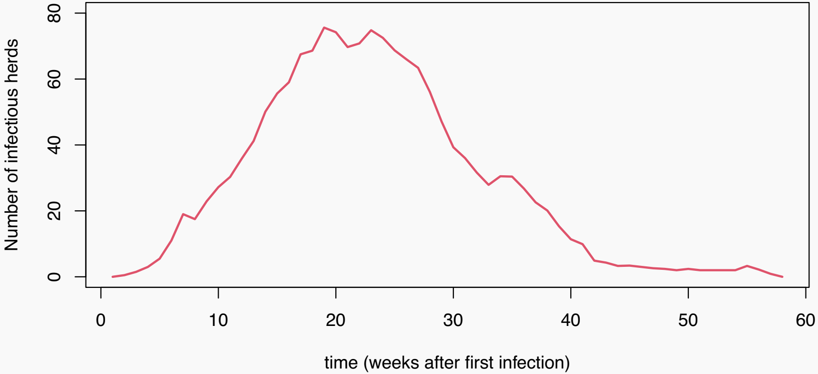
Introductions to Exercise 2.2

Estimating parameters: SIR model

Background: Classical Swine Fever Virus(CSFV) in the Netherlands

- A highly contagious disease of pigs and wild boar.
- A huge outbreak in the Netherlands took place between February 1997 and May 1998.
- There were 429 infected herds detected and stamped out.
- Netherlands has approximately $N = 21\ 500$ pig herds.

Plot of the weekly number of infectious herds



Estimating parameters: Gaussian observations

- We have n observations $y_i = I(t_i)$ at time points $t_1, \cdots, t_n$ with mean $\mathbf{E}[y_i; \theta]$, which is determined by the SIR differential system.
- Least squares estimates $\theta = (\beta, \gamma)$ minimizing the function

$$l(\theta) = \sum_{i=1}^{n} (y_i - \mathbf{E}[y_i; \theta])^2,$$

corresponds to Maximum Likelihood Estimate for Gaussian observations with

$$I(t_i) \sim N(\mathbf{E}[y_i; \theta]; \sigma^2),$$

with the variance of the observation noise σ^2.

Estimating parameters: MLE for CSFV Data(1)

Define the log-likelihood function

```
 ll.gauss <- function(theta){
#determine the solution of SIR ODE
...  <- lsoda(...)
return(sum(dnorm(data, mean =..., sd = 1, log =
TRUE))
}.
```

Estimating parameters: MLE for CSFV Data(2)

Maximize the log-likelihood and compute MLE

```
mle <- optim(
#initial values for theta to be optimized over
...,
#log-likelihood function
fn = ll.gauss,
#maximize the function
control = list(fnscale = -1) ).
```

Introductions to Exercise 2.3

Estimating parameters: SEIR model(1)

In this exercise, we are supposed to fit the SEIR model with time changing $\beta(t)$ from Exercise 1.3 to the data of reported cases in Stockholm during Feb-Apr 2020.

$$\beta(t) = \begin{cases} \beta_0 \text{ ,if } t \leq t_1 - w, \\ \beta_0 + \frac{\beta_1 - \beta_0}{2w}(t - (t_1 - w)) \text{ ,if } t_1 - w < t \leq t_1 + w, \\ \beta_1 \text{ ,if } t > t_1 + w, \end{cases}$$

The parameters here to optimize for are

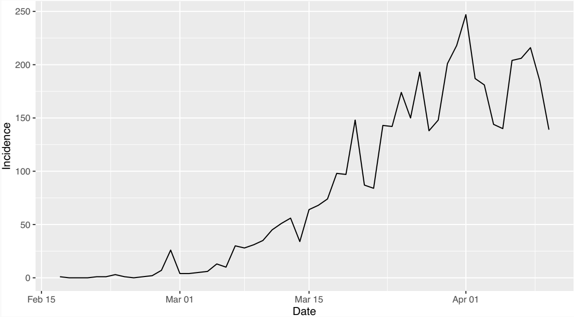
$$\theta = (\beta_0, \beta_1, t_1, w, \gamma).$$

Estimating parameters: SEIR model(2)

Assumptions:

- $N = 2374550, \rho = 1/5.$
- let $I(t)$ match the number of reports on calendar day t, with $I(0) = 1$ and $t = 0$ is equal to 2020-02-17.

Plot of the time series:



Least Square Approach for fitting(1)

Define the Least Square function

```
 ll.sq <- function(theta,I0){
#determine the solution of SEIR ODE
sol <- lsoda(y=, times=, func=,
parms=exp(theta))
sum((...-...)∧2 )
}.
```

Least Square Approach for fitting(2)

Compute the estimates

```
 theta_hat <- optim(
#starting values ..., fn = ll.sq,
#minimize the function
method="Nelder-Mead", IO=1)
```

Note: While using `optim`, try out more starting values (t_1 is not too small, γ not too large, $\beta_1 < \beta_0$...) to get a reasonably well-fitted curve.