The Mathematics of Infectious Diseases

Leila Amiri

Department of Community Health Sciences
University of Manitoba
Canada

November 20, 2018

Contents

The likelihood function for a counting process

The connection between the hazard and survival functions:

$$\begin{aligned}\lambda^*(t) &= \frac{f(t)}{S(t)} \\ &= \frac{f(t)}{1-F(t)} \\ &= \frac{\partial}{\partial t}\Big(1-F(t)\Big)\end{aligned}$$

Therefore,

$$S(t) = \exp\Big(-\int_0^t \lambda^*(s)ds\Big)$$

So,

$$f(t) = \lambda^*(t)\exp\Big(-\int_0^t \lambda^*(s)ds\Big)$$

The likelihood function for a counting process

- Consider a counting process N leading to the intensity $\lambda^*(t)$.
- Denote the ith jump time of $N^*(t)$ by τ_i^* and let τ_i^* be infinity if $N(t)$ does not make i jumps.

The likelihood function for a counting process observed up to time t is given as

$$L(\theta, t) = \left\{ \prod_{\tau_i^* \leq t} \lambda^*(\tau_i^*) \right\} \exp\left(- \int_0^t \lambda^*(s) ds \right)$$

The likelihood function for a counting process

- Multivariate counting process $\mathbf{N}^* = (N_1^*, \ldots, N_k^*)$
- intensity process $\lambda^* = (\lambda_1^*, \ldots, \lambda_k^*)$
- q-dimensional parameter $\theta = (\theta_1, \ldots, \theta_q)$

The likelihood function is

$$L(\theta, t) = \prod_h \prod_{s \leq t} \{\lambda_h^*(s)\}^{\Delta N_h^*(s)} \exp(-\int_0^t \lambda_\cdot^*(s)\, ds).$$

where $\lambda_\cdot^*(s) = \sum_h \lambda_h^*(s)$. The log-likelihood function up to time t for a multivariate counting process can be written as

$$\begin{aligned} l(\theta, t) &= \log(L(\theta, t)) \\ &= \sum_h \left[\int_0^t \log(\lambda_h^*(s))\, dN_h^*(s) - \int_0^t \lambda_h^*(s)\, ds \right] \end{aligned}$$

SIR (susceptible-infective-recovered) model

A system of three ordinary differential equations describes SIR model:

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{I_t S_t}{N} \\ \frac{dI}{dt} &= \beta \frac{I_t S_t}{N} - \gamma I_t \\ \frac{dR}{dt} &= \gamma I_t \end{aligned}$$

SIR model

Event	Rate
$(S(t), I(t)) \longrightarrow (S(t)-1, I(t)+1)$	$\beta \frac{I_t S_t}{N}$
$(S(t), I(t)) \longrightarrow (S(t), I(t)-1)$	γI_t

- $$\lambda^*_{Inf}(t) = \beta \frac{I_t S_t}{N} \qquad \text{and} \qquad \lambda^*_{Rec}(t) = \gamma I_t$$
- The log-likelihood is

$$l(\beta, \gamma) = \int_0^\tau [\log\{\beta\, \bar{S}(u)\, I(u)\}\, \mathrm{d}N(u) - \beta\, \bar{S}(u)\, I(u)\, \mathrm{d}u + \log\{\gamma\, I(u)\}\, \mathrm{d}R(u) - \gamma\, I(u)\, \mathrm{d}u]$$

where $\bar{S}(u) = S(u)/N$ and $N(u) = S(0) - S(u)$ is the number of individuals infected in $(0, u]$.

SIR model

The maximum likelihood estimation of the parameters are

$$\hat{\beta} = \frac{n_I}{\int_0^t S_u I_u du}$$

$$\widehat{\gamma^{-1}} = \frac{\int_0^t I_u du}{n_R}$$

where n_I and n_R are the total numbers of infected and recovered in $(0, \tau]$, respectively

SEIR (susceptible-exposed-infective-recovered) model

- If S_t, E_t, I_t and R_t are random variables representing the number of susceptible, latent, infective and removed individuals at time t, respectively
- dt is a vanishingly small time period such that only one event can occur in $[t, t + dt)$

The probabilities of events within the population are

$$P(\mathcal{S} \to \mathcal{E}) = \beta N^{-1} S_t I_t dt + o(dt),$$
$$P(\mathcal{E} \to \mathcal{I}) = \delta E_t dt + o(dt),$$
$$P(\mathcal{I} \to \mathcal{R}) = \gamma I_t dt + o(dt).$$

- The notation $S \longrightarrow E$ corresponds to the movement of a single individual from state S to state E

SEIR model

- Closed population of N individuals
- A series of exact continuous (ordered) event times, $t(0) = 0 < t_{(1)} < \ldots < t_{(k)} < t$ can be observed

The likelihood function would be given by

$$
\begin{aligned}
f(\boldsymbol{t} \mid \beta, \delta, \gamma) = & \\
& \left\{ \prod_{i=1}^{k} \left[\left(\frac{\beta}{N} S_{t_{(i-1)}} I_{t_{(i-1)}} \right)^{\eta_{1(i)}} \left(\delta E_{t_{(i-1)}} \right)^{\eta_{2(i)}} \left(\gamma I_{t_{(i-1)}} \right)^{1-\eta_{1(i)}-\eta_{2(i)}} \right] \right. \\
& \left. \times \exp \left[- \left(\frac{\beta}{N} S_{t_{(i-1)}} I_{t_{(i-1)}} + \delta E_{t_{(i-1)}} + \gamma I_{t_{(i-1)}} \right) \left(t_{(i)} - t_{(i-1)} \right) \right] \right\} \\
& \times \exp \left[- \left(\frac{\beta}{N} S_{t_{(k)}} I_{t_{(k)}} + \delta E_{t_{(i-1)}} + \gamma I_{t_{(k)}} \right) \left(t - t_{(k)} \right) \right] .
\end{aligned}
$$

SEIR model

- t is the current point of the epidemic
- $\eta_{(i)} = (\eta_{1(i)}, \eta_{2(i)})$ are a set of binary indicator variables
  - $\eta_{1(i)} = 1$ if event $t_{(i)}$ is an infection (e.g. a move from $S \longrightarrow E$), and 0 otherwise
  - $\eta_{2(i)} = 1$ if event $t_{(i)}$ is a move from $E \longrightarrow I$, and 0 otherwise.
  - $S_{t_{(i)}}$, $E_{t_{(i)}}$ and $I_{t_{(i)}}$ are counts of susceptible, latent and infective individuals at time $t_{(i)}$, respectively

Individual level model

- This model determines how transmission occurs from infectious to susceptible individuals within the population
- $P(i,t)$: the probability that a previously uninfected (susceptible) individual i is infected within the time interval $(t, t+1]$.

$$P(i,t) = 1 - Pr(i \text{ is not infected in} (t, t+1] | i \text{ is not infected in} (-\infty, t-1])$$

- Deardon et al. (2010)

$$P(i,t) = 1 - \exp\left[\left(-\Omega_S(i) \sum_{j \in I(t)} \Omega_T(j) k(i,j)\right)\right] + \varepsilon(i,t)$$

Individual level model

- $\Omega_S(i)$ represents risk factors associated with susceptible i (i.e., susceptibility)
- $\Omega_T(j)$: represents risk factors associated with infectious individual j (i.e., transmissibility)

They could be linear functions of covariates and associated parameters representing the effect of covariates (such as age or genomic information) on the risk of an individual contracting, or passing on, the disease.

Individual level model

- $k(i,j)$ is an infection kernel representing risk of infection over some distance measure (function of Euclidean or road distance)
  Common examples:
  - Exponential: $k(i,j) \propto \exp(-\beta d_{ij})$
  - Gaussian: $k(i,j) \propto \exp(-\beta d_{ij}^2)$
  - Power law: $k(i,j) \propto d_{ij}^{-\beta}$

  where d_{ij} is the distance between susceptible individual i and infective j
- $\varepsilon(i,t)$ represents infections that are not well explained by the $\Omega_S(i)$, $\Omega_T(j)$ and $k(i,j)$

Individual level model

Given a record of infection events at discrete time points $t = 0, \ldots, T$ during the epidemic, the likelihood is

$$f(\boldsymbol{S}, \boldsymbol{I}, \boldsymbol{R}|\boldsymbol{\theta}) = \prod_{t=0}^{T} f_t(\boldsymbol{S}, \boldsymbol{I}, \boldsymbol{R}|\boldsymbol{\theta}),$$

where

$$f_t(\boldsymbol{S}, \boldsymbol{I}, \boldsymbol{R}|\boldsymbol{\theta}) = \left[\prod_{i \in \boldsymbol{I}(t+1) \setminus \boldsymbol{I}(t)} P(i,t)\right]\left[\prod_{i \in \boldsymbol{S}(t+1)} \{1 - P(i,t)\}\right]$$

Individual level model

- θ is the vector of unknown parameters
- $S = \{S(t)\}_{t=1}^{T}$, $I = \{I(t)\}_{t=1}^{T}$ and $R = \{R(t)\}_{t=1}^{T}$
- $f_t(S, I, R|\theta)$ is the probability that all new infections observed in time interval $[t, t+1)$, as denoted by $I(t+1) \setminus I(t)$, are infected under the model, and all susceptible individuals observed not to be infected within this interval, as denoted by $S(t+1)$, are not infected under the model.

Example: Spatial Models with Covariates

- If the individuals being modeled are humans or animals, some of which are vaccinated, we may wish to estimate, or account for, a vaccination effect in the susceptibility function.

$$\Omega_S(i) = \alpha + \alpha_0 X_i$$

where X_i is a binary indicator variable such that

$$X_i = \begin{cases} 1, & \text{if } i \text{ has not been vaccinated,} \\ 0, & \text{if } i \text{ has been vaccinated.} \end{cases}$$

$\alpha > 0$ and $\alpha_0 > 0$ are constant infectivity and vaccination effect parameters, respectively

Example: Spatial Models with Covariates

- Let $k(i,j) = d_{ij}^{-\beta}$ where β is a spatial parameter and d_{ij} is the distance between susceptible individual i and infective j

$$P(i,t) = 1 - \exp\left[-(\alpha + \alpha_0 X_i) \sum_{j \in I(t)} d_{ij}^{-\beta}\right]$$

Simple case

$$P(i,t) = 1 - \exp\left\{-\alpha \sum_{j \in I(t)} d_{ij}^{-\beta}\right\} \quad \alpha, \beta > 0$$

$$\frac{\partial l(\theta)}{\partial \theta_i} = \frac{\partial}{\partial \theta_i}\left[\sum_{t=1}^{t_{max}} \sum_{i \in S(t+1)} \log(1 - P(i,t))\right] + \frac{\partial}{\partial \theta_i}\left[\sum_{t=1}^{t_{max}} \sum_{i \in I(t+1) \setminus I(t)} \log(P(i,t))\right]$$

$$= \left[\sum_{t=1}^{t_{max}} \sum_{i \in S(t+1)} \frac{\partial}{\partial \theta_i}\left(-\alpha \sum_{i=1}^{n} d_{ij}^{-\beta}\right)\right] + \left[\sum_{t=1}^{t_{max}} \sum_{i \in I(t+1) \setminus I(t)} \frac{\partial}{\partial \theta_i} \log\left(1 - \exp\left(-\alpha \sum_{i=1}^{n} d_{ij}^{-\beta}\right)\right)\right]$$

$$S(\alpha) = \frac{\partial l(\theta)}{\partial \alpha} = \left[\sum_{t=1}^{t_{max}} \sum_{i \in S(t+1)} \left(\sum_{j \in I(t)} d_{ij}^{-\beta}\right)\right] + \left[\sum_{t=1}^{t_{max}} \sum_{i \in I(t+1) \setminus I(t)} \frac{\sum_{j \in I(t)} d_{ij}^{-\beta}}{\exp\left(\alpha \sum_{j \in I(t)} d_{ij}^{-\beta}\right) - 1}\right]$$

$$S(\beta) = \frac{\partial l(\theta)}{\partial \beta} = \left[\sum_{t=1}^{t_{max}} \sum_{i \in S(t+1)} \alpha \sum_{j \in I(t)} d_{ij}^{-\beta} \log(d_{ij})\right] + \left[\sum_{t=1}^{t_{max}} \sum_{i \in I(t+1) \setminus I(t)} \frac{\alpha \sum_{j \in I(t)} d_{ij}^{-\beta} \log(d_{ij})}{1 - \exp\left(\alpha \sum_{j \in I(t)} d_{ij}^{-\beta}\right)}\right]$$

Simple case

The maximization of the log-likelihood can be carried out by setting the gradient equal to 0 and solving for the optimal parameter values. Although the analytic derivatives (and second derivatives) are tractable, Newton-Raphson method can be used.

Epidemic Curves: R code

- Create an SIR function

```
SEIR =function(time, state, parameters) {
with(as.list(c(state, parameters)), {
dS= -beta * S * I
dE = (beta * S * I) - (delta * E)
dI = (delta * E) - (gamma * I)
dR = gamma * I
return(list(c(dS, dE, dI, dR)))
})
}
```

- Set parameters

  Proportion in each compartment: Susceptible 0.999999, Exposed 0.000001, Infected 0.000001, Recovered 0

```
init = c(S = 1-1e-6, E = 1e-6, I = 1e-6, R = 0.0)
```

Epidemic Curves: R code

- beta: infection parameter; gamma: recovery parameter
  parameters= c(beta = 1.4247, delta=.5, gamma = 0.14286)
- Time frame times =seq(0, 70, by = 1)
- library(deSolve)
  Solve using ode (General Solver for Ordinary Differential Equations)
  out = ode(y = init, times = times, func = SEIR, parms = parameters)
- change to data frame out = as.data.frame(out)

Epidemic Curves: R code

- matplot(x = times, y = out, type = “l”, xlab = “Time”, ylab = “Susceptible and Recovered”, main = “SEIR Model”, lwd = 1, lty = 1, bty = “l”, col = 2:5)
- legend(40, 0.7, c(“Susceptible”, “Exposed”, “Infected”, “Recovered”), pch = 1, col = 2:5, bty = “n”)

Epidemic Curves