Epidemic models 1

Ben Bolker

2021-02-08 ©BMB (except textbook material/other images)

motivation

- P & I data from Philadelphia 1918 flu:

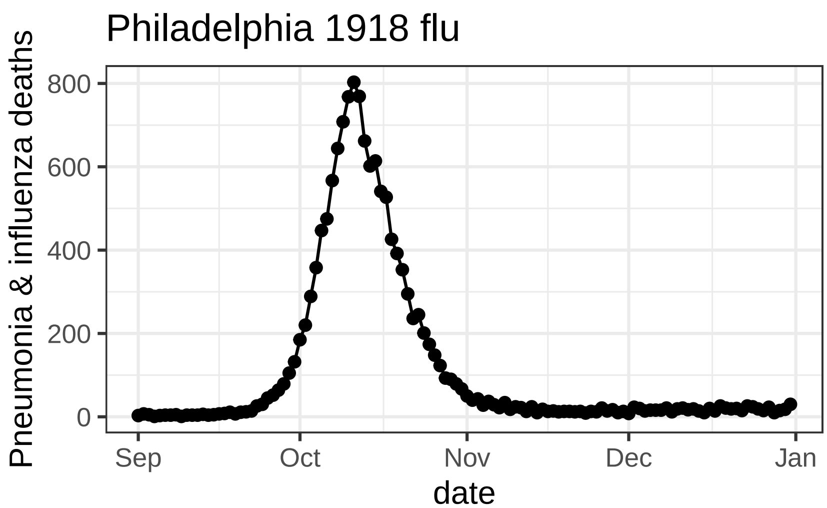


Figure 1: Phila. 1918 flu data

what do we want to figure out?

what shall we assume?

- classify individuals as S, I (**compartmental** model; **microparasite** or **intensity-independent**)
- disease is transmitted from S to I
- $S \to I$ instantaneously (zero latent period, no E)
- population is **homogeneous** (no heterogeneity in susceptibility, infectiousness, contact)
- fixed population size (birth = migration = 'natural' death = 0)
- transmission rate is time-invariant

- assumption 2 is OK (Pasteur, Koch's postulates . . .)
- all the rest are approximations

  start simple!

- parsimony

- robustness?
- applicability/estimation?

Levins (1966) (also Orzack and Sober (1993), Levins (1993), Weisberg (2007))

exponential growth

- one variable (=1D model)
- how does disease spread? $\rightarrow$ equation

---

what variables should we use?

- time (t)
- state variable: incidence, prevalence, death rate, death toll (= cumulative death?)
- deaths loosely connected to transmission

but deaths are observed!

---

when are deaths a good **proxy** for incidence?

- infection -> death time is fixed
- homogeneity? (might not matters?)
- mortality curve is shifted epidemic

(COVID context … we observe case reports, number of tests, hospitalizations, and deaths)

- **incidence**: number of infections per unit time (rate or flow)
- **prevalence**: number of currently infected people (quantity or stock)

prevalence is closer to the **mechanism**

---

model components:

- $I(t)$ (state variable: prevalence)
- $I(0)$ (initial conditions)
- β (parameter) = avg contacts **per susceptible per infective per unit time**

$$I(t+\Delta t) \approx I(t) + \beta I(t)\Delta t$$

Take $\lim \Delta t \to 0$ (and solve):

$$\frac{dI}{dt} = \beta I \to I(t) = I(0)exp(\beta t)$$

model criticism

- Ignored discrete nature of individuals
- Ignored time-varying β (e.g. **diurnal** fluctuations)
- Ignored finite infectious periods (recovery/death)

Next: What if we make infectious periods finite? (i.e., including recovery (**clearance**) or death

$$dI/dt = \beta I - \gamma I$$

mean infectious period

$$\begin{aligned} I(t) &= I(0)\exp(-\gamma t) \\ \text{proportion uninfected} &= \exp(-\gamma t) \\ \text{proportion infected} &= 1 - \exp(-\gamma t) (= \text{CDF} := C(t)) \\ \text{PDF} : &= C'(t) = \gamma \exp(-\gamma t) \\ \text{substitute } x = \gamma t \quad &\to \quad dx = \gamma\, dt \\ \text{mean} = E[t] &= \int t \exp(-\gamma t)\, dt = \int x \exp(-x)\, dx/\gamma = 1/\gamma \end{aligned}$$

dimensional analysis

rates and characteristic times/scales

- is I a proportion or a density or a number ... ?
- what are the units of β, γ ?

nondimensionalization

- standardize any values that can be eliminated **without loss of (mathematical) generality**
- what can we do here?
- $\gamma = 1$
- I ? (depends on how we have defined it initially) $\to I/N$

compare with data???

Original scale:

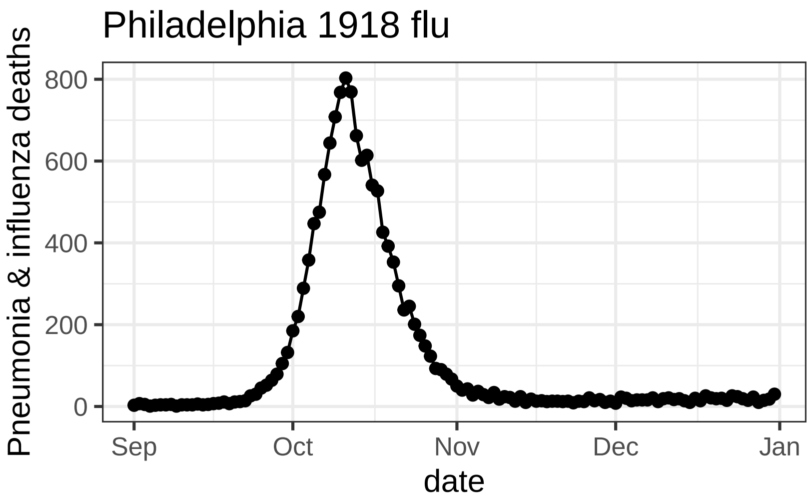


Figure 2: Philadelphia P&I

Log scale:

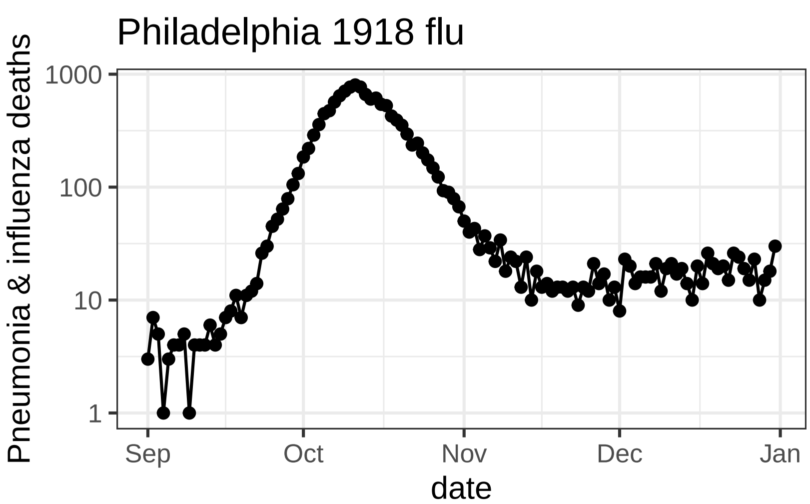


Figure 3: Philadelphia P&I, log scale

- Fit a straight line through the straight part of the curve
- slope is βN
- "intercept" is $\log(I(0))$ (zero is defined in a tricky way)

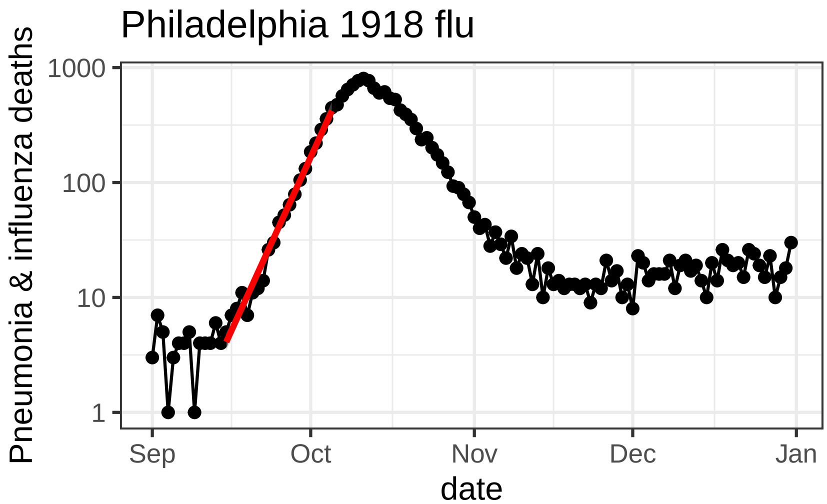
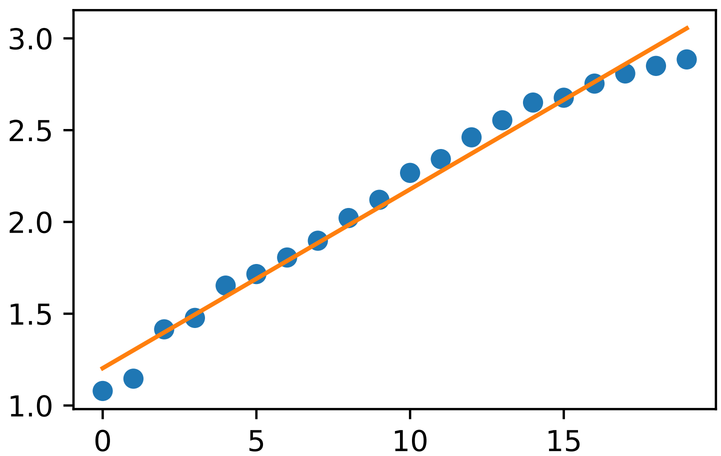


Figure 4: log-scale flu with regression

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
dd = pd.read_csv("data/pim_us_phila_city_1918_dy.csv")
## plt.plot(dd.pim)
## plt.plot(np.log10(dd.pim))
t = np.arange(20)
lw = np.log10(dd.pim)[20:40]
plt.plot(t,lw,'o')
## https://docs.scipy.org/doc/scipy/reference/generated/scipy.linalg.lstsq.html
## https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html
ta = t.reshape(-1,1)  ## make this into a column vector
reg = LinearRegression().fit(ta,lw)
plt.plot(t,reg.intercept_+reg.coef_[0]*t)
```



model assessment

- math is super-easy!
- clear, testable predictions
- parameter estimation is easy
- only consistent over a short time window
  - small t: arbitrarily close to zero
  - large t: ridiculous

Simple (SI) epidemic

- what are we missing?
- **depletion of susceptibles**
- let's take a step back and ignore death & recovery for now

$$dS/dt = -\beta SI$$
$$dI/dt = \beta SI$$

This looks 2D **but** what if we assume $S + I = N$ is constant? Then $S = N - I$

$$dI/dt = \beta(N - I)I$$

How do we solve this? **Partial fractions**

$$
\begin{aligned}
\frac{dI}{\beta(N-I)I} &= dt \\
dI\left(\frac{A}{N-I}+\frac{B}{I}\right) &= dI\cdot\frac{A+B(N-I)}{I(N-I)} \\
A = B; &\quad B = 1/N \\
\frac{1}{\beta N}(-\log(N-I)+\log(I))\Bigg|_{I(0)}^{I} &= t - t_0 \\
(-\log(N-I)+\log(I))\Bigg|_{I(0)}^{I} &= (\beta N)(t-t_0) \quad (\text{set } t_0 = 0) \\
\log\left(\frac{I}{N-I}\right) - \log\left(\frac{I(0)}{N-I(0)}\right) &= \beta N t \\
\log\left(\frac{I}{N-I}\right) &= \beta N t + -\log\left(\frac{I(0)}{N-I(0)}\right) \\
\frac{I}{N-I} &= \exp(\beta N t)\frac{I(0)}{N-I(0)} \equiv Q \\
I &= Q(N-I) \\
I(t)(1+Q) &= QN \\
I(t) &= \frac{QN}{1+Q} = \frac{N}{1+\frac{1}{Q}} \\
&= \frac{N}{1+\left(\frac{N-I(0)}{I(0)}\right)\exp(-\beta N t)}
\end{aligned}
$$

?? $\equiv I(0)\exp(\beta N t)/(1+(I0/N)(\exp(\beta N t)-1))$??

Qualitative analysis

- $I \ll N$? exponential growth
- **per capita growth rate** ($(dI/dt)/I = d(\log(I))/dt$) decreases monotonically with increasing I
- asymptotic behaviour? equilibria? periodic orbits?
- periodic orbits impossible in 1D (uniqueness of flows)

equilibrium analysis

- $I = 0$, **disease free equilibrium** (DFE)
- $I = N$, **endemic equilibrium** (EE)

  Stability? (Assume $\beta > 0$)

- **local asymptotic stability**
- **global asymptotic stability** (Lyapunov functions)

model criticism/conclusions

(Comparison to metapop, logistic growth model)

SIR model

Basic SIR model

- put the pieces together

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

- really 2D (because $S + I + R = N$)
- rescale to $N = 1$ (S, I, R as proportions)

Numerical solution (R version):

```
## define gradient function
SIRgrad <- function(t, y, parms) {
    g <- with(as.list(c(y,parms)), {
        c(-beta*S*I, beta*S*I-gamma*I, gamma*I)
    })
    return(list(g))
}
library(deSolve)
## initial conditions and parameters
y0 <- c(S=0.99, I=0.01, R=0)
p0 <- c(beta=4, gamma=1)
tvec <- seq(0,8,length=101)
## solve (LSODA by default)
sir_R <- ode(y=y0, times=tvec, parms=p0, func=SIRgrad)

## plot
par(las=1,bty="l") ## cosmetic
matplot(tvec, sir_R[,-1],
        type="l", lwd=2, ## solid lines, thicker
        xlab="time",ylab="proportion")
legend("right",names(y0), col=1:3, lty=1:3, lwd=2)
```

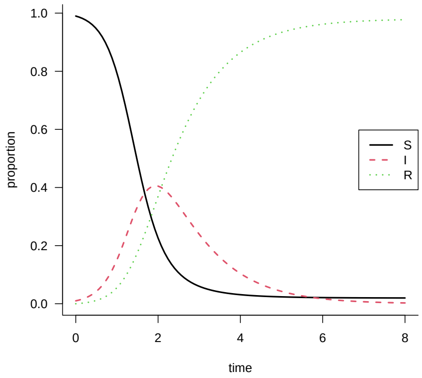


Figure 5: SIR model (R)

Phase plane plot

```
par(las=1,bty="l") ## cosmetic
plot(I~S,type="l",data=as.data.frame(sir_R))
with(as.data.frame(sir_R), points(S,I, cex=0.75,pch=16))
```

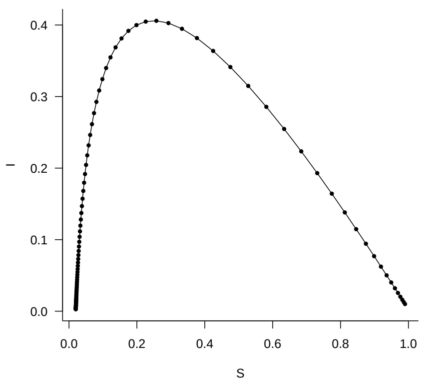


Figure 6: SIR phase plane (R)

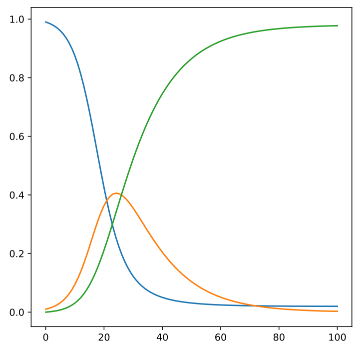
Solve using Python

```
import numpy as np
import scipy.integrate
def SIR_grad(x,t,params):
   """basic gradient definitions for SIR model"""
   beta,gamma = params   ## unpack parameters
   S,I,R = x             ## unpack state variables
   return(np.array([-beta*S*I, beta*S*I-gamma*I, gamma*I]))

t_vec = np.linspace(0,8,101)
params = (4,1) ## extra parameters (beta, gamma)
y0 = (0.99, 0.01, 0)
SIR_sol1 = scipy.integrate.odeint(SIR_grad,
                                  y0=y0,
```

```
                                       t=t_vec,
                                       args=(params,))

## https://community.rstudio.com/t/how-to-display-the-plot-in-the-python-chunk/22039/3
import matplotlib.pyplot as plt
fig, ax = plt.subplots()
ax.plot(SIR_sol1);
plt.show()
```



dimensional analysis

- initial growth rate (time^{-1}) $\beta - \gamma$
- mean infectious period $1/\gamma$ (time)
- basic reproduction number $\mathcal{R}_0 = \beta/\gamma$

initial growth rate

$$\begin{aligned}\frac{dI}{dt} &= \beta S - \gamma I \\ &= (\beta S - \gamma) I \\ &\approx (\beta - \gamma) I \quad \text{near } DFE\end{aligned}$$

or calculate **Jacobian** ($\partial X_i / \partial X_j$):

$$\begin{pmatrix} -\beta I & -\beta S & 0 \\ \beta I & \beta S - \gamma & 0 \\ 0 & \gamma & 0 \end{pmatrix}$$

Evaluate at DFE ($\{1, 0, 0\}$):

$$\begin{pmatrix} 0 & -\beta & 0 \\ 0 & \beta - \gamma & 0 \\ 0 & \gamma & 0 \end{pmatrix}$$

Eigenvalues of this are pretty boring! But useful approach.

Per capita rates

In general we can express *per capita* gradients in X as gradients of $\log(X)$:

$$\begin{aligned}\frac{dX}{dt} &= Xf(X,Y,Z,\dots) \\ \frac{\frac{dX}{dt}}{X} &= f(X,Y,Z,\dots) \\ \frac{d\log(X)}{dt} &= f(X,Y,Z,\dots)\end{aligned}$$

Another way to see that $\beta - \gamma$ is the slope on the log scale.

Stability of DFE

- $\beta > \gamma$ ($r > 0$)
- $\beta/\gamma > 1$ ($\mathcal{R}_0 > 1$)

Local asymptotic stability **or**

- $\frac{dI}{dt} = \beta SI - \gamma I$
- non-dimensionalize: $\gamma = 1$, $\beta = \mathcal{R}_0$
- $\frac{dI}{dt} = (\mathcal{R}_0 S - 1)I$
- $\frac{d\log I}{dt} = \mathcal{R}_0 S - 1$

Since $S \leq 1$, $\mathcal{R}_0 < 1 \to$ deriv of $\log I$ is always negative (don't really need the last step)

Automated analysis

```
library(phaseR)

## -------------------------------------------------------------------------------
## phaseR: Phase plane analysis of one- and two-dimensional autonomous ODE systems
## -------------------------------------------------------------------------------
##
## v.2.1: For an overview of the package's functionality enter: ?phaseR
##
## For news on the latest updates enter: news(package = "phaseR")

par(las=1,bty="l",xaxs="i",yaxs="i") ## cosmetic
SIRgrad_2d <- function(t, y, parms) {
    g <- with(as.list(c(y,parms)), {
        c(-beta*S*I, beta*S*I-gamma*I)
    })
    return(list(g))
```

```
}
## plot(0:1,0:1,type="n",xlab="S",ylab="I")
f1 <- flowField(SIRgrad_2d,
          xlim=c(0,1),
          ylim=c(0,1),
          parameters=p0,
          state.names=c("S","I"),
          add=FALSE)
n1 <- nullclines(SIRgrad,
                 xlim=c(0,1),
                 ylim=c(0,1),
                 parameters=p0,
                 state.names=c("S","I"))
t1 <- trajectory(SIRgrad_2d,parameters=p0,
           state.names=c("S","I"),
           ## n=10,
           y0=y0[1:2],
           tlim=c(0,5))
```

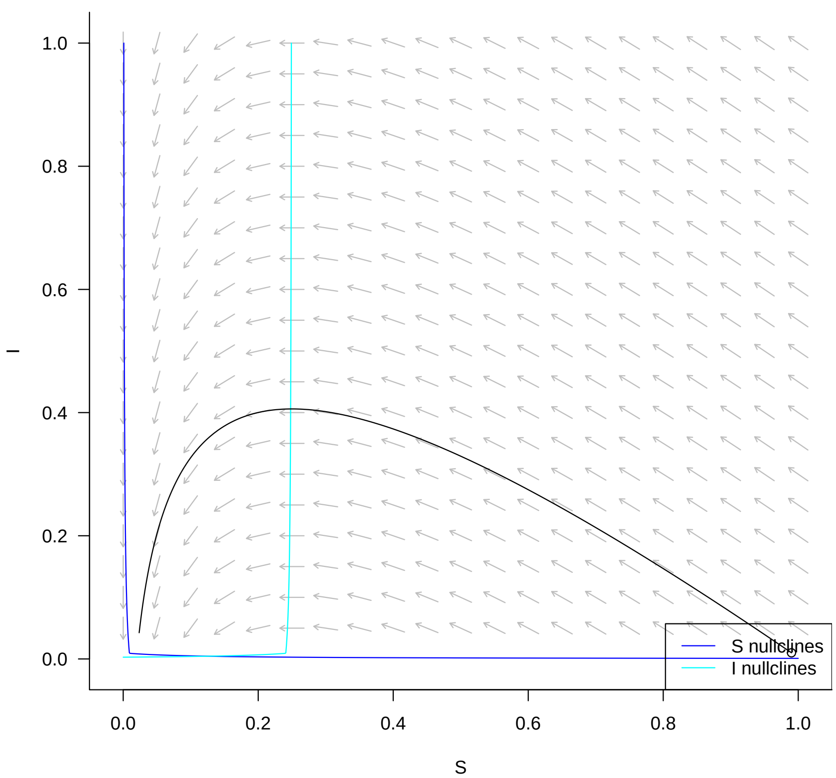


Figure 7: phase plane analysis in R

```
phasePlaneAnalysis(SIRgrad_2d,xlim=c(0,1),
                   parameters=p0,
```

```
            state.names=c("S","I"),
            ylim=c(0,1))
```

Solution

- can't get analytical solution for $S(t)$, $I(t)$
- **but**: we can solve for $I(S)$:

$$\begin{aligned}
\frac{dI}{dS} &= \frac{dI/dt}{dS/dt} = -1 + \frac{1}{\mathcal{R}_0 S} \\
\int_{I(0)}^{I} (t)\, dI &= \int_{S(0)}^{S(t)} \left(-1 + \frac{1}{\mathcal{R}_0 S}\right) dS \\
I - I(0) &= -(S - S(0)) + \frac{1}{\mathcal{R}_0} \log(S/S(0)) \\
I + S - (I(0) + S(0)) &= \frac{1}{\mathcal{R}_0} \log(S/S(0))
\end{aligned}$$

Final size calculations

- $t \to \infty$:

$$(I_\infty + S_\infty) - (I(0) + S(0)) = \frac{1}{\mathcal{R}_0} \log S_\infty / S(0)$$

- newly invading pathogen: $S \approx 1$, $I(0) \ll 1$ (≈ 0), $I_\infty \to 0$
- in the limit $I(0) \to 0$:

$$S_\infty) - 1 = \frac{1}{\mathcal{R}_0} \log S_\infty$$

- "final size" $Z = 1 - S_\infty$
- $-Z = \frac{1}{\mathcal{R}_0} \log(1 - Z)$

Lambert W functions

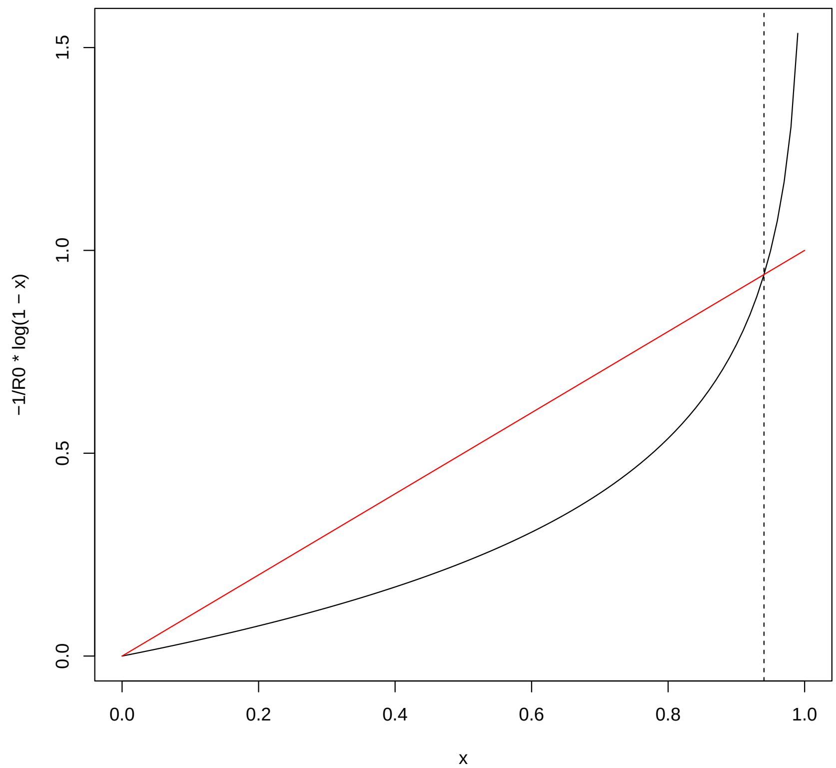
- How do we solve this?
- Newton's method (or whatever)
- *Lambert W* (Corless et al. 1996): solves $W \exp(W) = Z$

```
import sympy as sym
z, R = sym.symbols('z R')
sym.solve(sym.Eq(z,-1/R*sym.log(1-z)),z)

## [(R + LambertW(-R*exp(-R)))/R]

finalsize <- function(R0) {
    1+1/R0*lambertW(-R0*exp(-R0))
}
```

```
R0 <- 3
curve(-1/R0*log(1-x), from=0,to=1)
curve(1*x, add=TRUE,col="red")
library(emdbook)
abline(v=finalsize(R0),lty=2)
```



Epidemic threshold

Assuming vaccination (or other perfect *prophylaxis* [protection]) at rate p

$$R_0 = 1 - 1/p$$

speed-based intervention:

$$\begin{aligned} \beta SI - (\gamma + \phi)I < 0 \\ I(\beta - \gamma - \phi) < 0 \\ \phi > (\beta - \gamma) = r \end{aligned}$$

Comparing Epidemic threshold vs. final size

```
library(emdbook)
finalsize <- function(R0) {
```

```
    1+1/R0*lambertW(-R0*exp(-R0))
}
par(las=1,bty="l")
curve(finalsize(x),from=1,to=10,xlab=expression(R[0]),
      ylab="proportion")
curve(1-1/x, add=TRUE, col=2)
legend("bottomright",
       c("final size","herd immunity threshold"),
       col=1:2, lty=1)
```

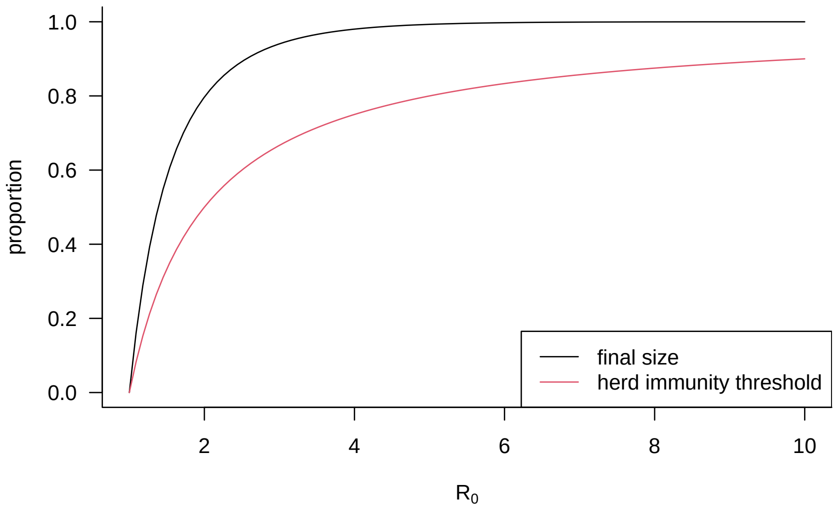


Figure 8: final size vs herd immunity

Estimating R from data

- Euler-Lotka equation

$$\begin{aligned}
I(t) &= \int_0^t I(t-\tau)K(\tau)\,d\tau \\
I(0)\exp(rt) &= \int_0^t I(0)\exp(r(t-\tau))K(\tau)\,d\tau \\
1 &= \int_0^t \exp(-r\tau)K(\tau)\,d\tau \\
1 &= \int_0^t \exp(-r\tau)\mathcal{R}_0 g(\tau)\,d\tau \\
\frac{1}{\mathcal{R}_0} &= \int_0^t \exp(-r\tau)g(\tau)\,d\tau
\end{aligned}$$

SIRS/SIR with vital dynamics

- models of **endemic** disease

- e.g. "childhood diseases" (measles, mumps, rubella, pertussis, polio, chickenpox, ...)
  - directly transmitted, acute, immunizing
- SIRS model
- SIR model *with vital dynamics*

$$
\begin{aligned}
\frac{dS}{dt} &= \mu N - \beta/NSI - \mu S \\
\frac{dI}{dt} &= \beta/NSI - \gamma I - \mu I \\
\frac{dR}{dt} &= \gamma I - \mu R
\end{aligned}
$$

- Balanced population (birth rate = μN). Can consider more complex demography but often don't need to
  - disease-induced death rates low relative to natural mortality
  - demographic time scales much longer than epidemic time scales
- Scaling β/N is much easier for dealing with applications/real data, scaling $N = 1$ is easier for doing math
- $\mathcal{R}_0$ is $\beta/(\mu+\gamma)$ ($\approx \beta/\gamma$ for most human diseases)

Most of the following is taken from Brauer, Castillo-Chavez, and Feng (2019)

$$
\begin{aligned}
S^* &= \frac{\mu+\gamma}{\beta} = 1/\mathcal{R}_0 \quad \text{(this is \textbf{very general})} \\
I^* &= \frac{\mu}{\mu+\gamma} - \frac{\mu}{\beta} = \frac{\mu}{\beta}(\mathcal{R}_0 - 1)
\end{aligned}
$$

- at equilibrium the **force of infection** is βI^*, so the **average age at infection** is $A = 1/(\beta I^*)$
- average lifespan is $L = 1/\mu$
- $L/A = \beta I^*/\mu = \mathcal{R}_0 - 1$
  - another way to estimate $\mathcal{R}_0$! (also, $S^* = 1/\mathcal{R}_0$)
  - tells us something about risk by age, effects of vaccination

Jacobian at EE:

$$
\begin{pmatrix}
-\mu\mathcal{R}_0 & -(\mu+\gamma) \\
\mu(\mathcal{R}_0 - 1) & 0
\end{pmatrix}
$$

Trace = $-\mu\mathcal{R}_0$, Det = $\mu(\mu+\gamma)(\mathcal{R}_0 - 1)$

$$
\begin{aligned}
\lambda &= (1/2)\left(-\mu\mathcal{R}_0 \pm \sqrt{\mu^2\mathcal{R}_0^2 - 4\mu(\mathcal{R}_0-1)(\mu+\gamma)}\right) \\
&\approx (1/2)\left(-\mu\mathcal{R}_0 \pm \sqrt{-4\mu(\mathcal{R}_0-1)\gamma}\right) \\
&= (1/2) - \mu\mathcal{R}_0 \pm i\sqrt{(1/L)(L/A)(1\gamma)}
\end{aligned}
$$

- So the imaginary part is $\sqrt{A\tau}$ where τ is the length of the infective period.
- e.g. for measles $A \approx 70\text{yr}$, $\tau \approx 1/26$, epidemic interval is about 2.76 years.

```
2*pi*sqrt(5/26)
```

```
## [1] 2.755359
```

- **Damping factor**: amplitude decreases by a factor $\exp(-\mu\mathcal{R}_0 T/2)$ per cycle

Stochasticity

- Reed-Frost model
  - household infection model: fixed (small) population, discrete infection generations
  - start with **index cases**
  - then ...
- Gillespie algorithm
- stochastic ODEs

References

Brauer, Fred, Carlos Castillo-Chavez, and Zhilan Feng. 2019. "Endemic Disease Models." In *Mathematical Models in Epidemiology*, edited by Fred Brauer, Carlos Castillo-Chavez, and Zhilan Feng, 63–116. Texts in Applied Mathematics. New York, NY: Springer. https://doi.org/10.1007/978-1-4939-9828-9_3.

Corless, Robert M., G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. 1996. "On the Lambert W Function." *Advances in Computational Mathematics* 5 (4): 329–59. https://doi.org/10.1007/BF02124750.

Levins, R. 1966. "The Strategy of Model Building in Population Biology." *American Scientist* 54: 421–31. https://www.jstor.org/stable/27836590.

Levins, Richard. 1993. "A Response to Orzack and Sober: Formal Analysis and the Fluidity of Science." *Quarterly Review of Biology* 68 (4): 547–55.

Orzack, Steven Hecht, and Elliott Sober. 1993. "A Critical Assessment of Levins's the Strategy of Model Building in Population Biology (1966)." *Quarterly Review of Biology* 68 (4): 533–46.

Weisberg, Michael. 2007. "Forty Years of 'the Strategy': Levins on Model Building and Idealization." *Biology & Philosophy* 21 (5): 623–45. https://doi.org/10.1007/s10539-006-9051-9.