Modeling for Public Health Review

Dr. Rachel Sippy
University of St Andrews

Objectives

- Review materials from the modeling workshop
- Basics for R software
- Overview of upcoming sessions

Post Questions in the Chat!

(we will have breaks to answer these during the workshop)

Workshop Schedule

Time	Topics
2:00—2:10 pm	Outline & Introduction
2:10—2:50 pm	Model Review
2:50—3:00 pm	Upcoming Workshop
3:00—3:10 pm	Break
3:10—4:00 pm	Introduction to R

Introductions!

- Dr Rachel Sippy
- Lecturer in Statistics at University of St Andrews

Types of Models

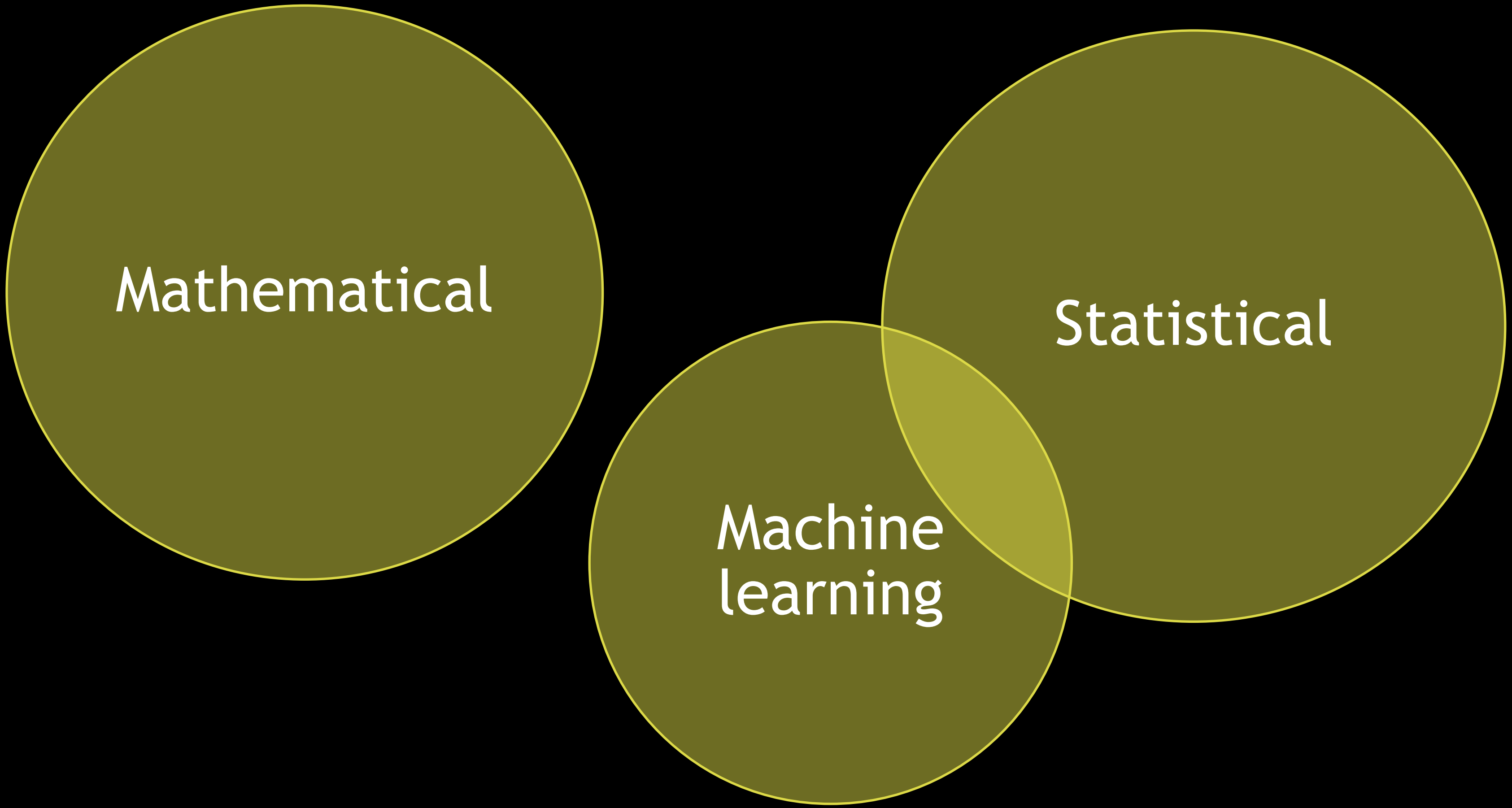
And Modeling Approaches!

Three Major Model Types

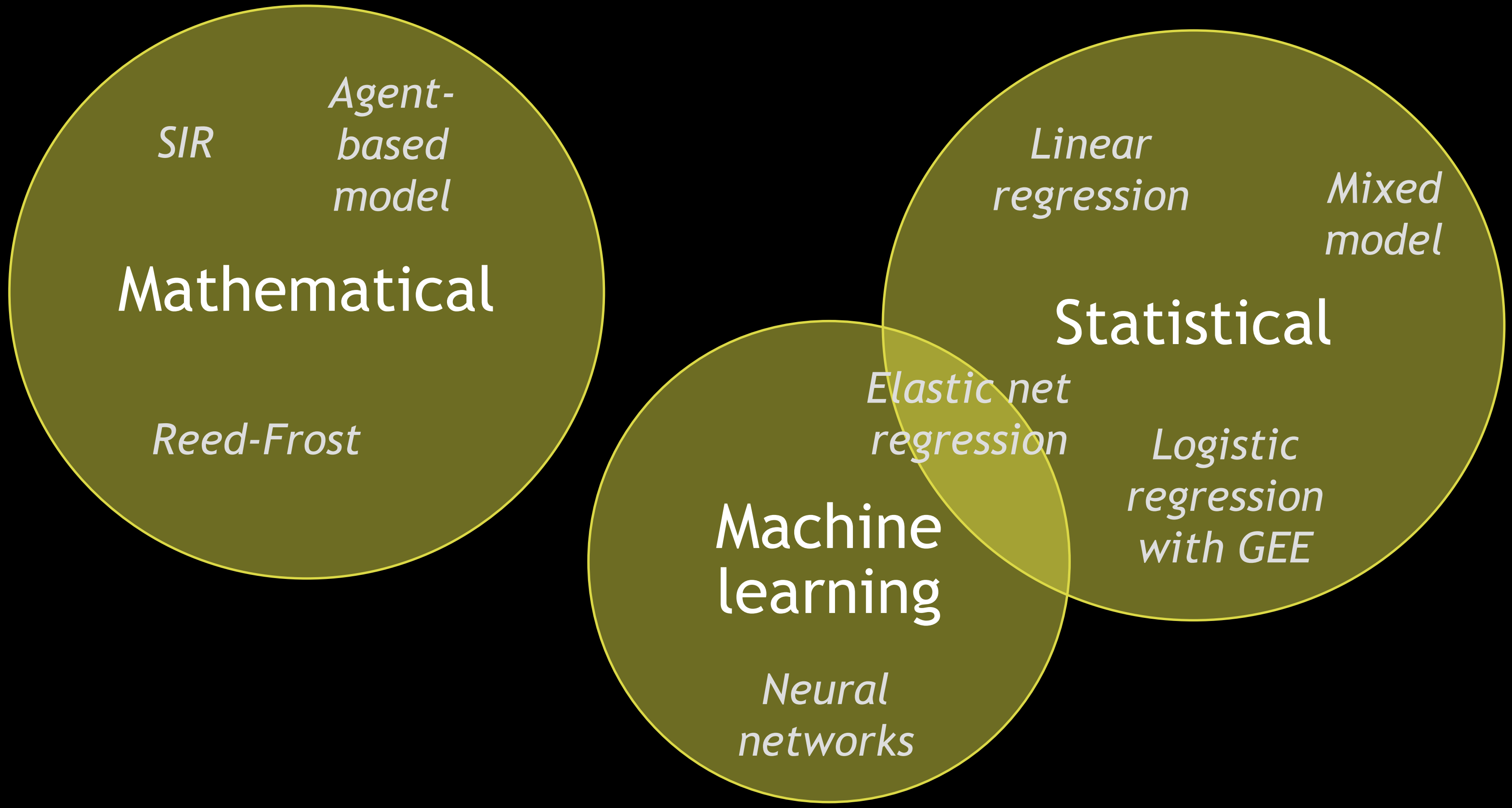
(used in epidemiology)
(general modeling approaches)

- Statistical models
  - Regression (many!), time series
- Mathematical models
  - Compartmental, mechanistic, agent-based
- Machine learning
  - Uses algorithms and statistical models

Relationships

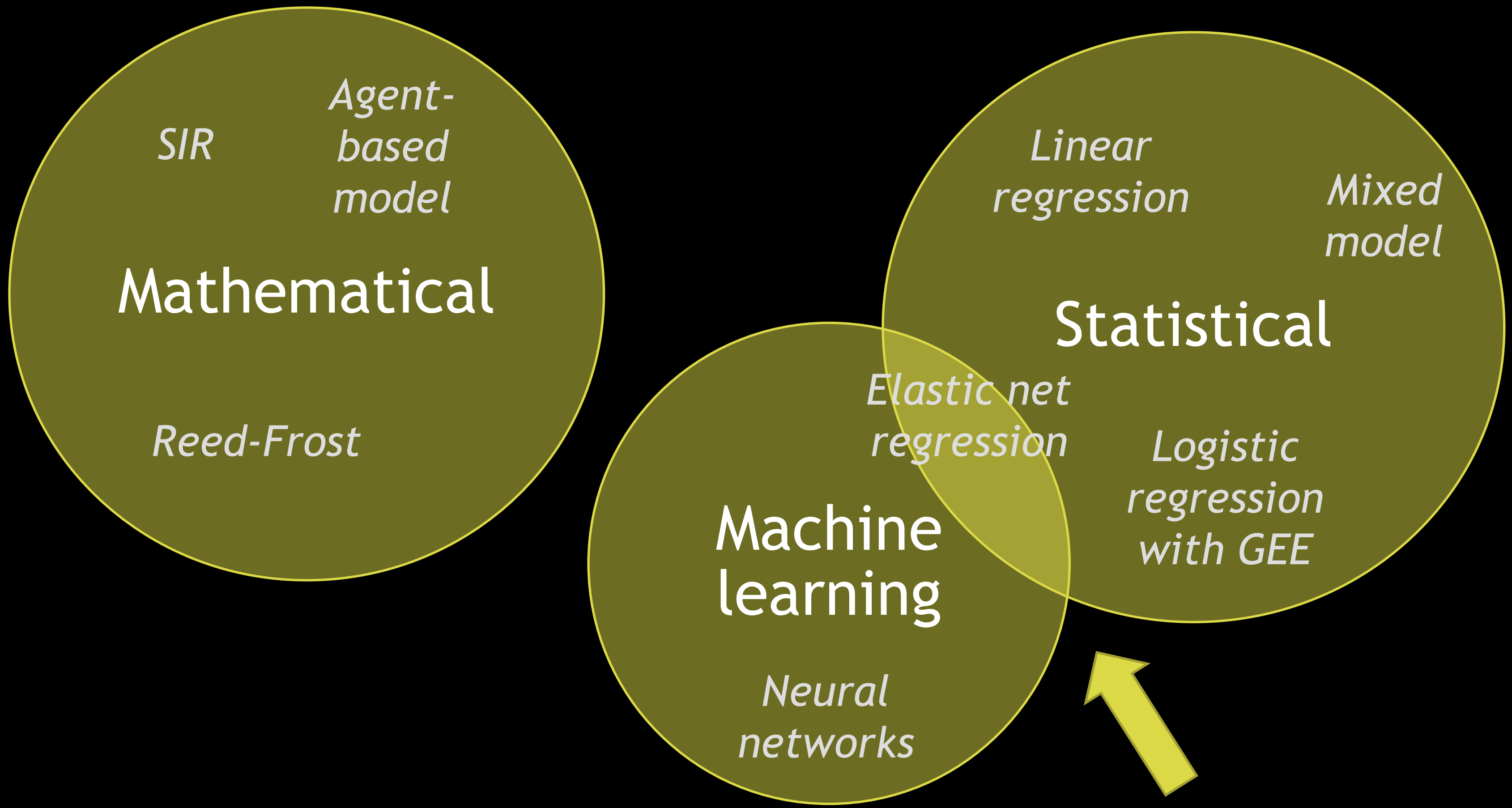


Relationships



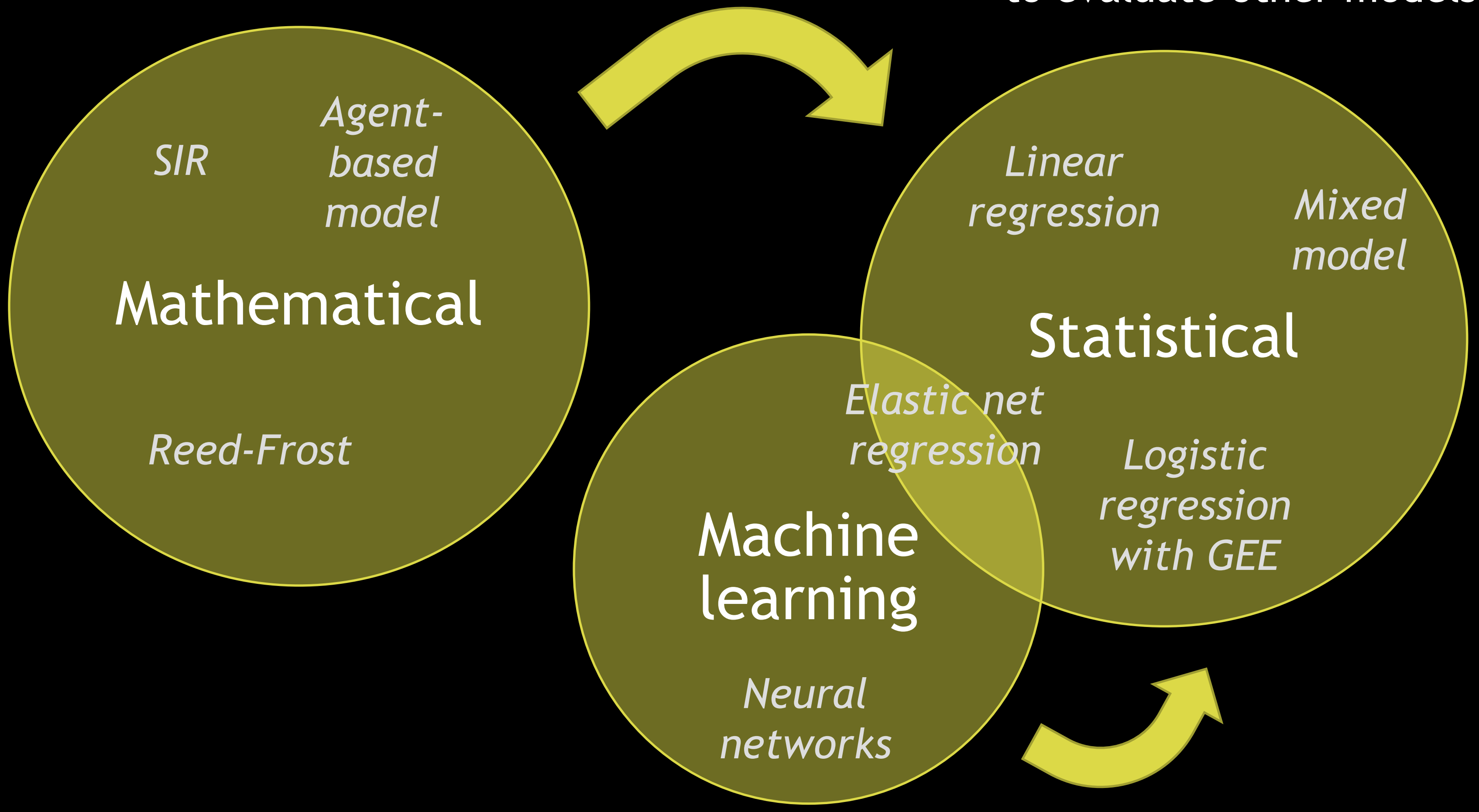
Relationships - Overlap

Same model, different context



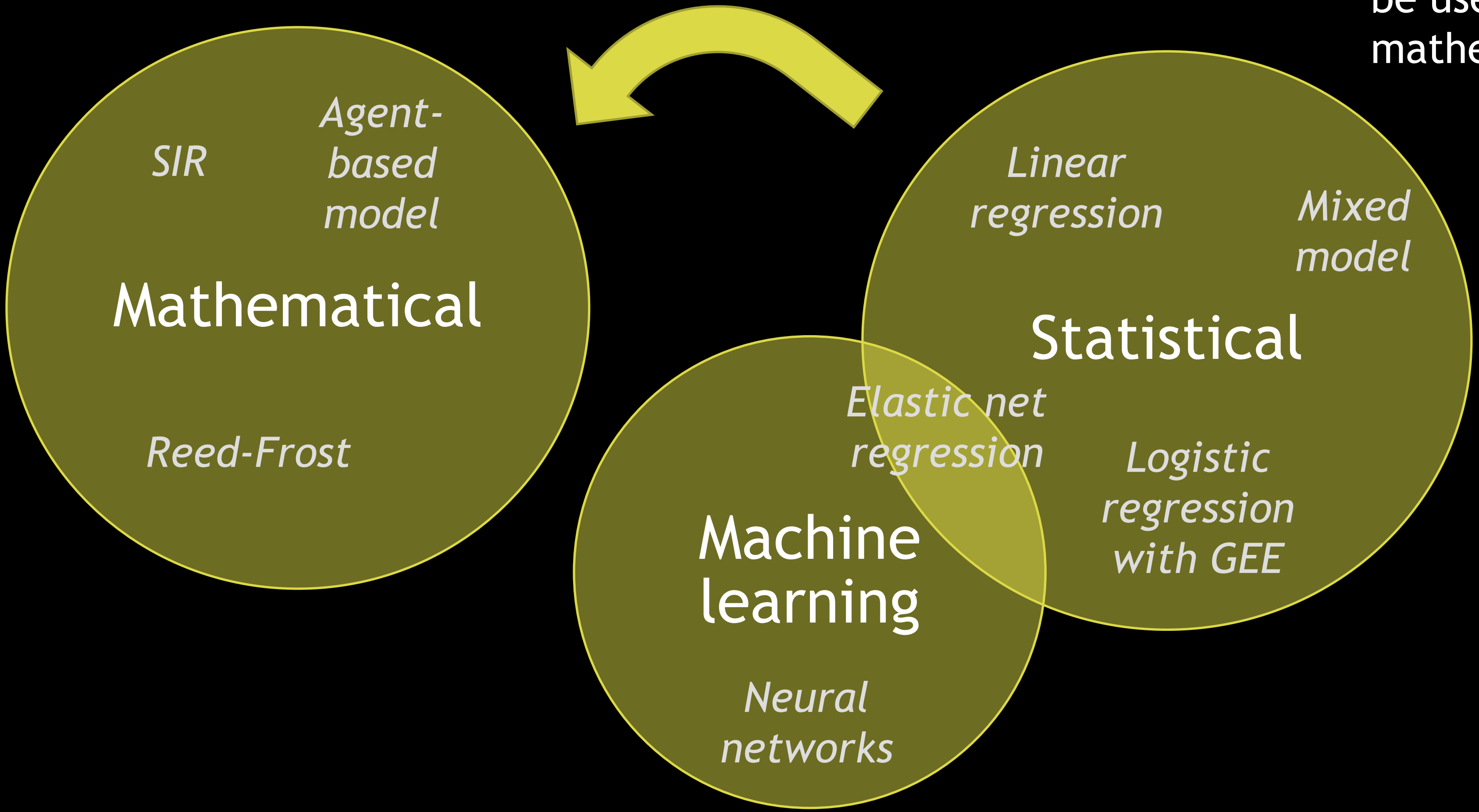
Relationships - Evaluation

Statistical methods often used to evaluate other models



Relationships – Parameter Estimates

Statistical models produce estimates to be used in mathematical models



Statistical Modeling

- Uses methods based in theory (frequentist, Bayesian, likelihood)
- Way to understand data and sources of its variability
- Concepts:
  - Our data are a sample from some larger population
  - Our data come from a distribution (parametric) or not (non-parametric)
- Generalization of statistical test

Statistical Modeling

- We start with the data
- Our data and question of interest will determine the model we use
- For our model results, we want to have an unbiased estimate of the relationship between the inputs and the outcome
- A high quality statistical model will have limited biases and an appropriate model
  - Results should account for uncertainty
  - Hopefully uncertainty is limited

Statistical Modeling Questions

- What is the effect of temperature on mosquito abundance?
- Does vaccination affect disease risk?
- What will the life expectation be for people with certain characteristics?

Mathematical Modeling

- Formal description of a process we are interested in
- Model structure/form is critical first decision
  - Often depends on type of disease being studied
- Parameters/inputs for the model often come from previous knowledge or other literature
- Models commonly formulated with differential equations
- A high-quality mathematical model should examine multiple scenarios
  - Validated if possible
  - Results should account for uncertainty
  - Hopefully uncertainty is limited

Mathematical Modeling

- We start with our model: based on system processes
- Often, we use equations to create a scenario (set of conditions)
  - Example: population of city, level of interaction, certain disease, susceptibility of population
- If we can create a realistic model, then we can make changes to model and see what happens
  - Example: what happens when we vaccinate some people?
- This is similar to running many experiments and observing what happens
- Results from these observations can be compared to real data
  - Example: do model results match what happened in real life?

Mathematical Modeling Questions

- If we vaccinate the population, how many cases will occur?
- Why does the disease have a biennial pattern?
- Given the current population and disease conditions, how many deaths will occur?

Machine Learning

- Subset of artificial intelligence developed in computer science
- Using computers to develop predictions from data
- Need to make predictions in scenarios where statistical models were inappropriate
  - Image recognition
  - Handwriting recognition
  - Financial markets

Machine Learning

- Methods may be very complex and difficult to understand, usually not interpretable
- Many models run repeatedly
- Focus on prediction results

Machine Learning Questions

- Will people with certain characteristics get the disease?
- What is the temperature in this location?
- What are the types of houses in this village?

Statistical vs. Mathematical Models

- Statistical: we are finding the relationship between the variables that is the most likely explanation for the data we have (or the relationship with least uncertainty)
  - We are estimating the size and direction of the relationships
  - These are our parameters
- Mathematical: we assume relationships/model structure, make changes and calculations, then compare to real data (validation)
  - We are estimating or predicting outcome size
  - Our inputs are our parameters

What Can We Do with Models?

(and what can’t we do)

What is the Goal of Modeling?

- Inference
- Prediction

Model Use: Inference

- Estimate relationships between outcome and input data
  - Are inputs associated with the outcome?
    -
  - How are inputs related to outcome?
    -
- Understand how data were generated
  - Why are we observing a particular pattern of data?
    -

Model Use: Inference

- Estimate relationships between outcome and input data
  - Are inputs associated with the outcome?
    - Are climate variables associated with COVID-19 rates?
  - How are inputs related to outcome?
    - How is population size related to measles incidence?
- Understand how data were generated
  - Why are we observing a particular pattern of data?
    - Why are gonorrhea levels staying at a low level?

Model Use: Prediction

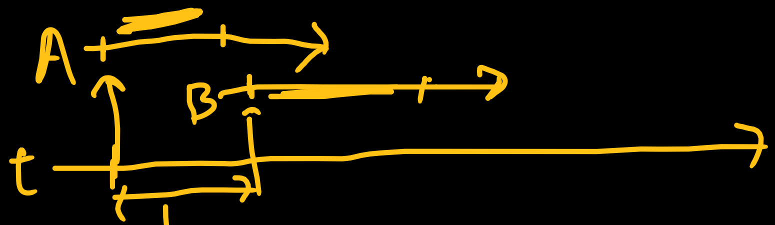
- Predicting outcomes
  - Predict general outcomes
    -
  - Predicting outcomes for an individual
    -
  - Predict impact of controls
    -
- Forecasting
  - Predicting what will happen in the future
    -

Model Use: Prediction

- Predicting outcomes
  - Predict general outcomes
    - What is the average effect of treatment on population?
  - Predicting outcomes for an individual
    - Will a person with certain characteristics get the disease?
  - Predict impact of controls
    - How will vaccination affect the number of cases?
- Forecasting
  - Predicting what will happen in the future
    - How many cases will there be in total?

Compartmental Models & Parameters

SIR + more



Epidemic Parameters

- Doubling time : time until # cases to double
- Reproductive number (R): R_0, R_t avg # people infected by 1 case
- Generation time : time between infectious states
- Herd immunity: $1 - S$ ← proportion susceptible
- Final size

total # infected

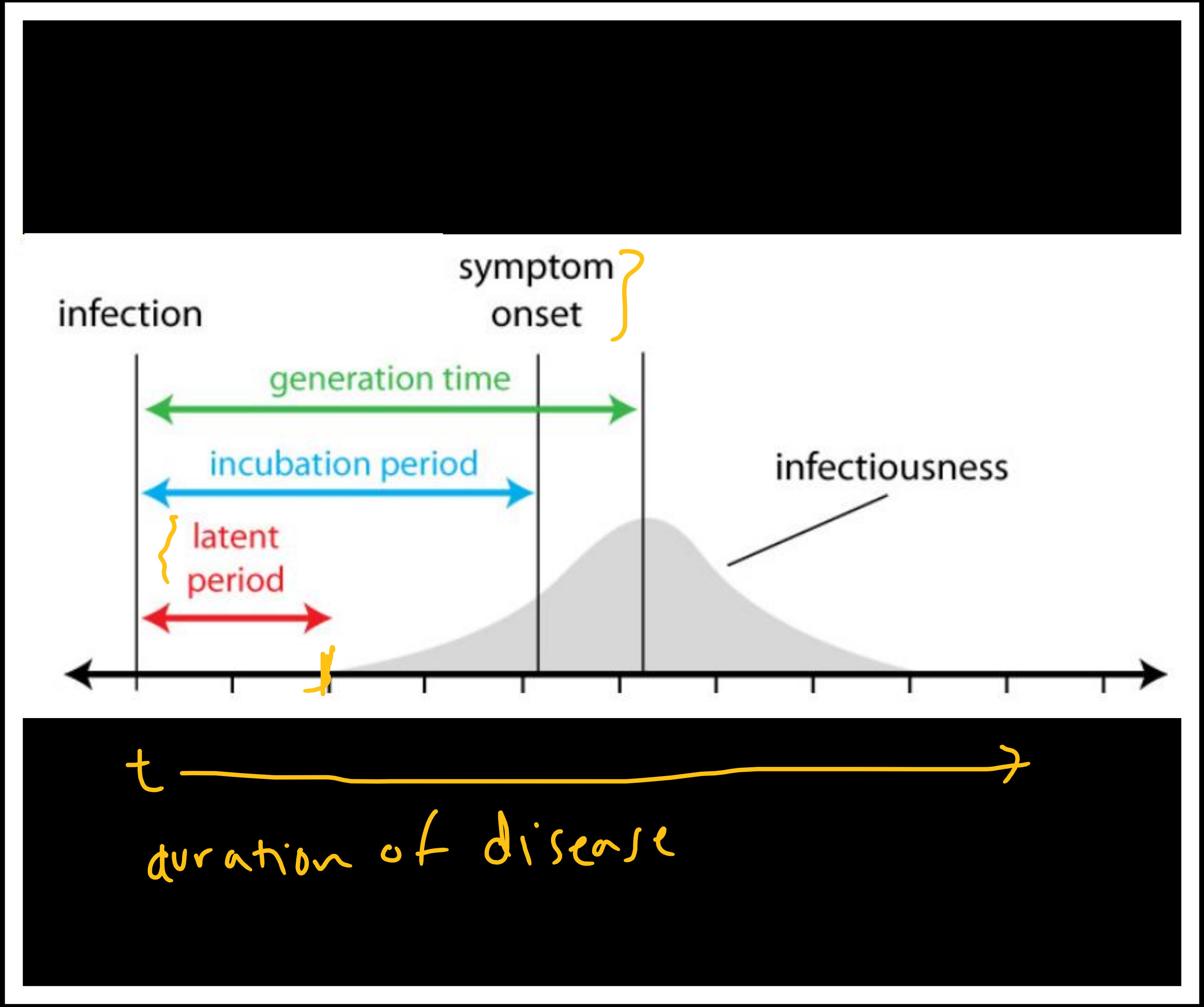
R_0 = everyone susceptible

R_t = susceptible (partial)

$$HIT = 1 - 1/R_0$$

$$DT = \frac{\ln 2}{\ln\left(\frac{R}{GT}\right)}$$

$$R = 2^{GT/DT}$$



Generation Time & Epidemics

- Incubation period
  - time between infection and development of symptoms
- Latent period
  - time between infection and becoming infectious
- Infectious period
  - period when a case is infectious
- Generation time is most important
  - depends on biological factors and number of contacts

Data for Parameters

- How do we know what values to enter for our model parameters?
- How do we know what to use for starting values?
- Models (depending on the structure) need a lot of inputs, such as:
    - birth rates
    - death rates
    - transmission rate
    - latency rate
    - recovery rate
    - number in each compartment
    - Where might we find these numbers?

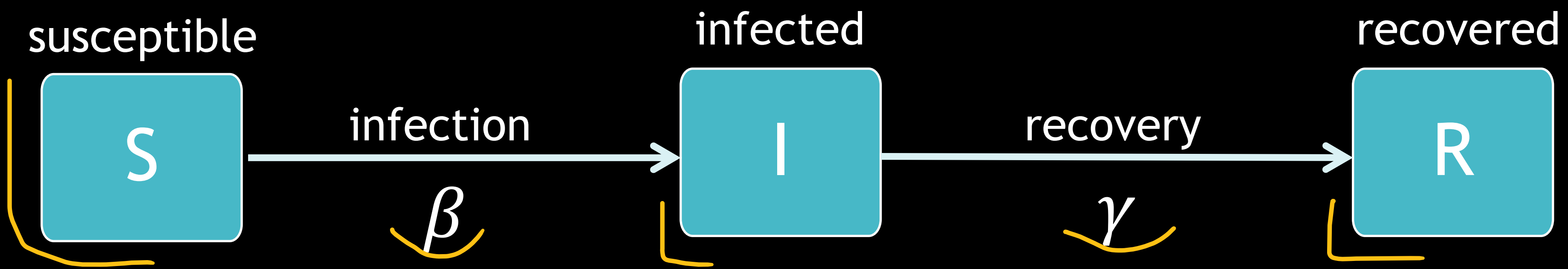
natural history parameters

(important term)

surveillance
population data

Compartmental models

SIR Model: Kermack & McKendrick



- β is a transmission coefficient
  - it is the probability of a susceptible becoming infected if they contact an infected person

$$\begin{cases} \dfrac{dS(t)}{dt} = -\beta S(t) I(t) \\[2ex] \dfrac{dI(t)}{dt} = \beta S(t) I(t) - \gamma I(t) \\[2ex] \dfrac{dR(t)}{dt} = \gamma I(t) \end{cases}$$

- γ is the recovery rate
  - 1/ γ is the duration of infectiousness

SIR $\boxed{S} \to \boxed{I} \to \boxed{R}$

SEIR $\boxed{S} \to \boxed{E} \xrightarrow[\sigma]{} \boxed{I} \to \boxed{R}$

SIS $\boxed{S} \to \boxed{I} \to \boxed{S}$

type of disease

treatment ?

$$\begin{tikzcd} \boxed{S} \arrow[r] & \boxed{I} \arrow[r, "\gamma"] \arrow[dr, "\gamma_T"'] & \boxed{R} \\ & & \boxed{T} \end{tikzcd}$$

vaccination ?

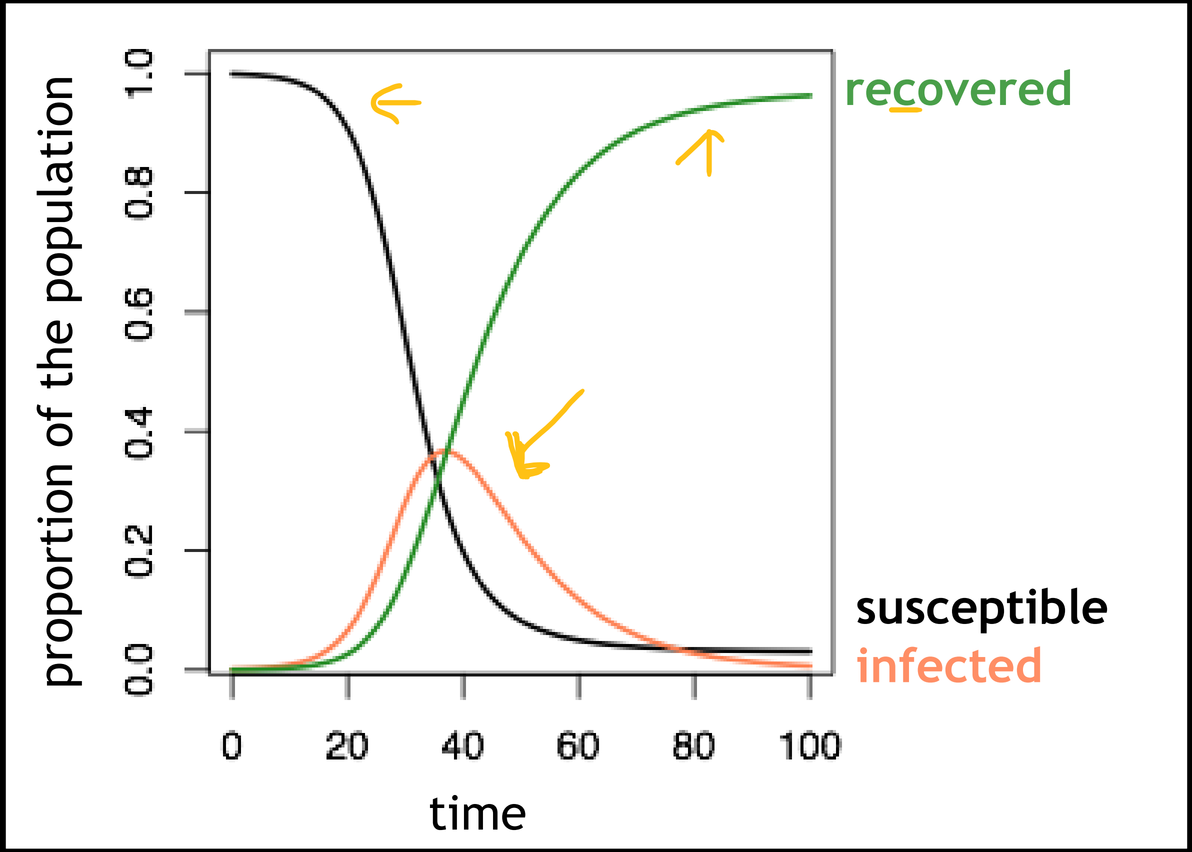
$$\begin{tikzcd} \boxed{S} \arrow[r, "\beta"'] \arrow[rrr, bend right, "v"'] & \boxed{E} \arrow[r, "\sigma"'] & \boxed{I} \arrow[r, "\gamma"'] & \boxed{R} \end{tikzcd}$$

modify to answer questions

Thought Process for Building Models

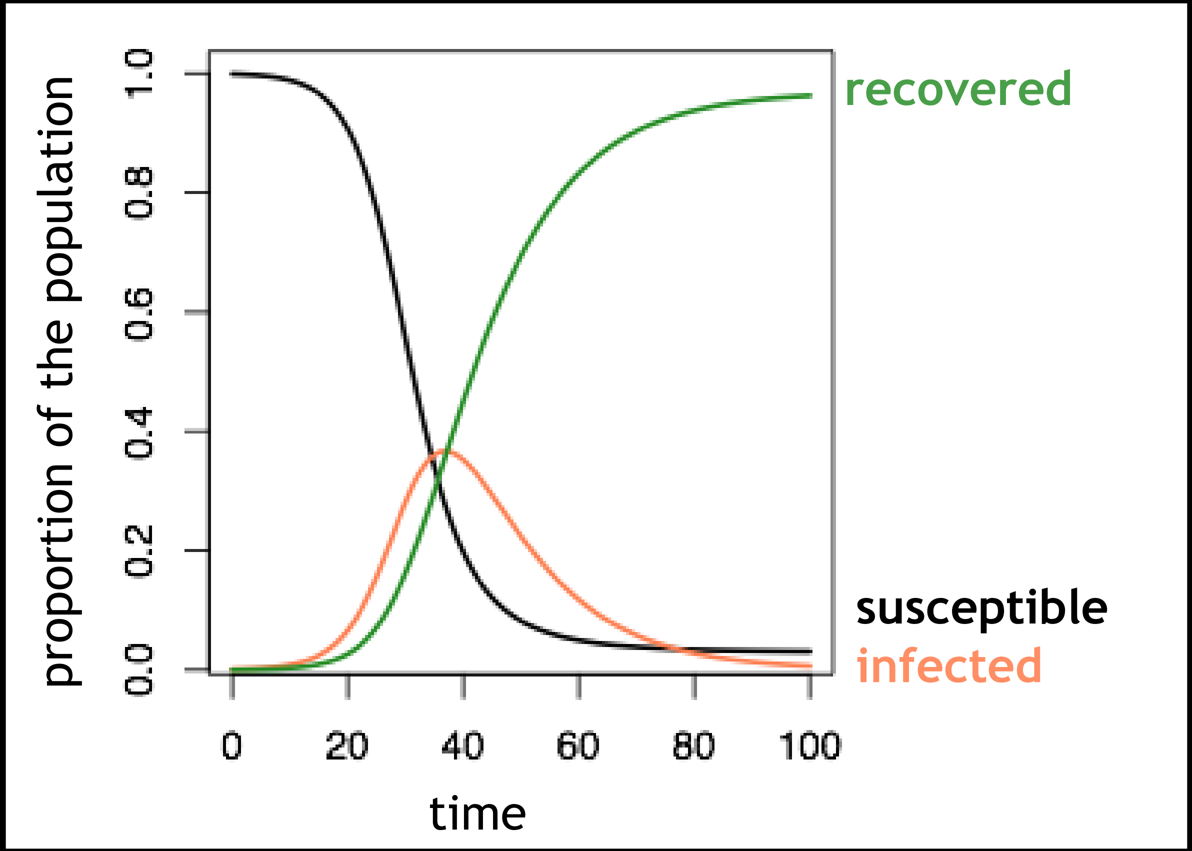
- The basic SIR model can be adapted for many situations
- For any new model, consider:
  - What is changing?
  - Do I need a new compartment? Fewer compartments?
  - Do I need to change rates?
  - Do I need additional transitions?
- Translate this to your differential equations
  - represent the flow of individuals between compartments
  - update your equations to reflect any new transitions

SIR Model Output



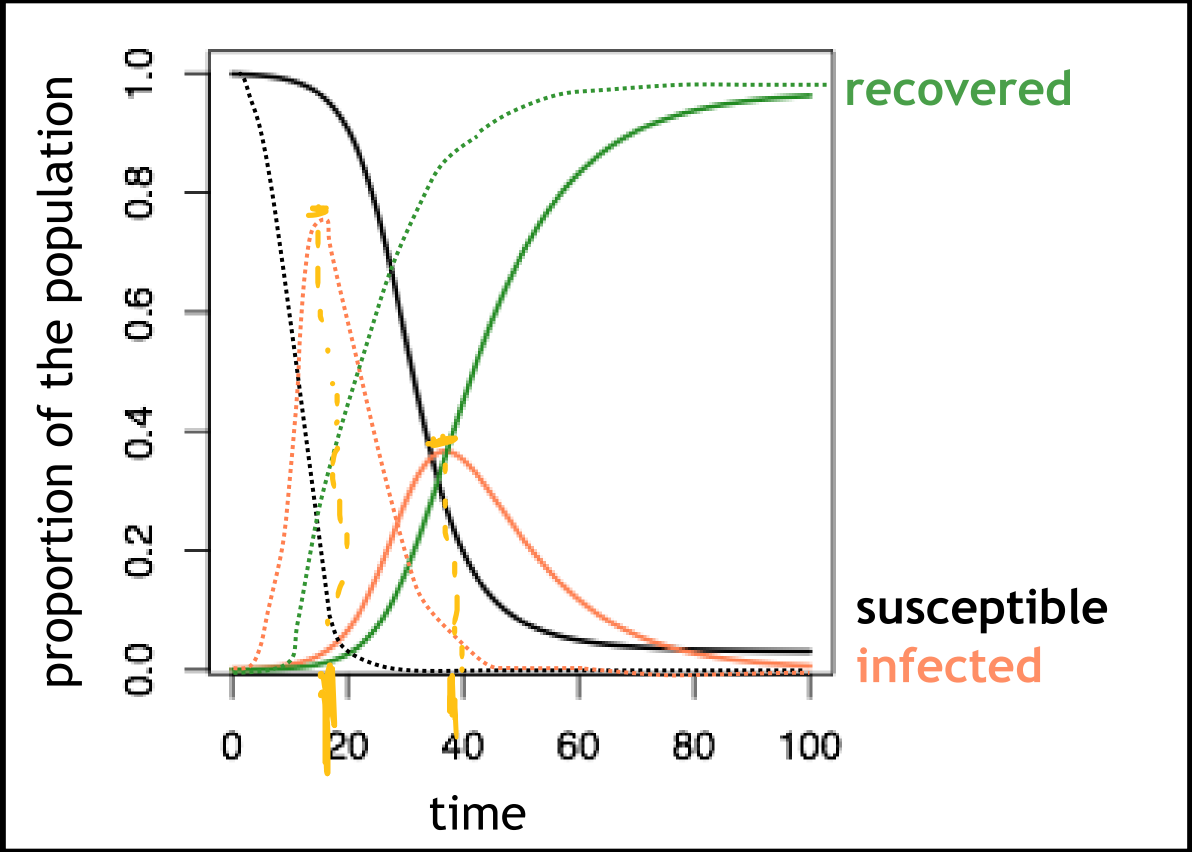
- We can examine the epidemic dynamics
- We will have values for S(t), I(t), and R(t) for each time step t
  - these can be expressed as population proportions and plotted
  - we can also find totals for each compartment at the end of the epidemic

SIR Model Output



- In our first examples last week, we input parameter values and starting values
  - parameter values for beta and gamma
  - starting values for number of people in each compartment S, I, R
- What about our other epidemic determinants?

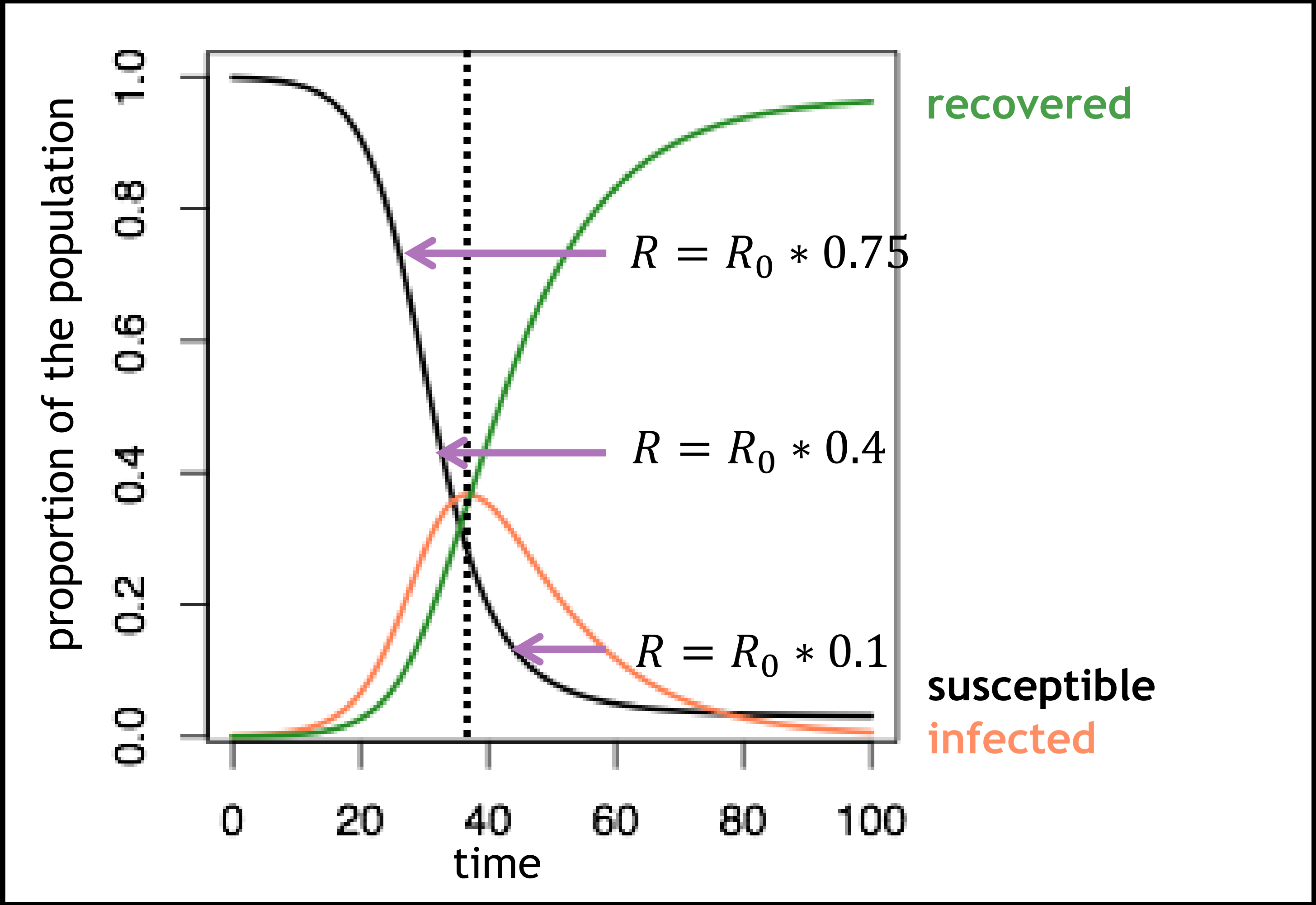
SIR Model Output



- Shape of curves will depend on epidemic determinants
  - low • solid lines → $R_0 = 3.6$
  - high • dotted lines → $R_0 = 13.2$
- What differences can we note about these two epidemics with two different basic reproductive numbers?

dynamics different!

Reproductive Number during Epidemic

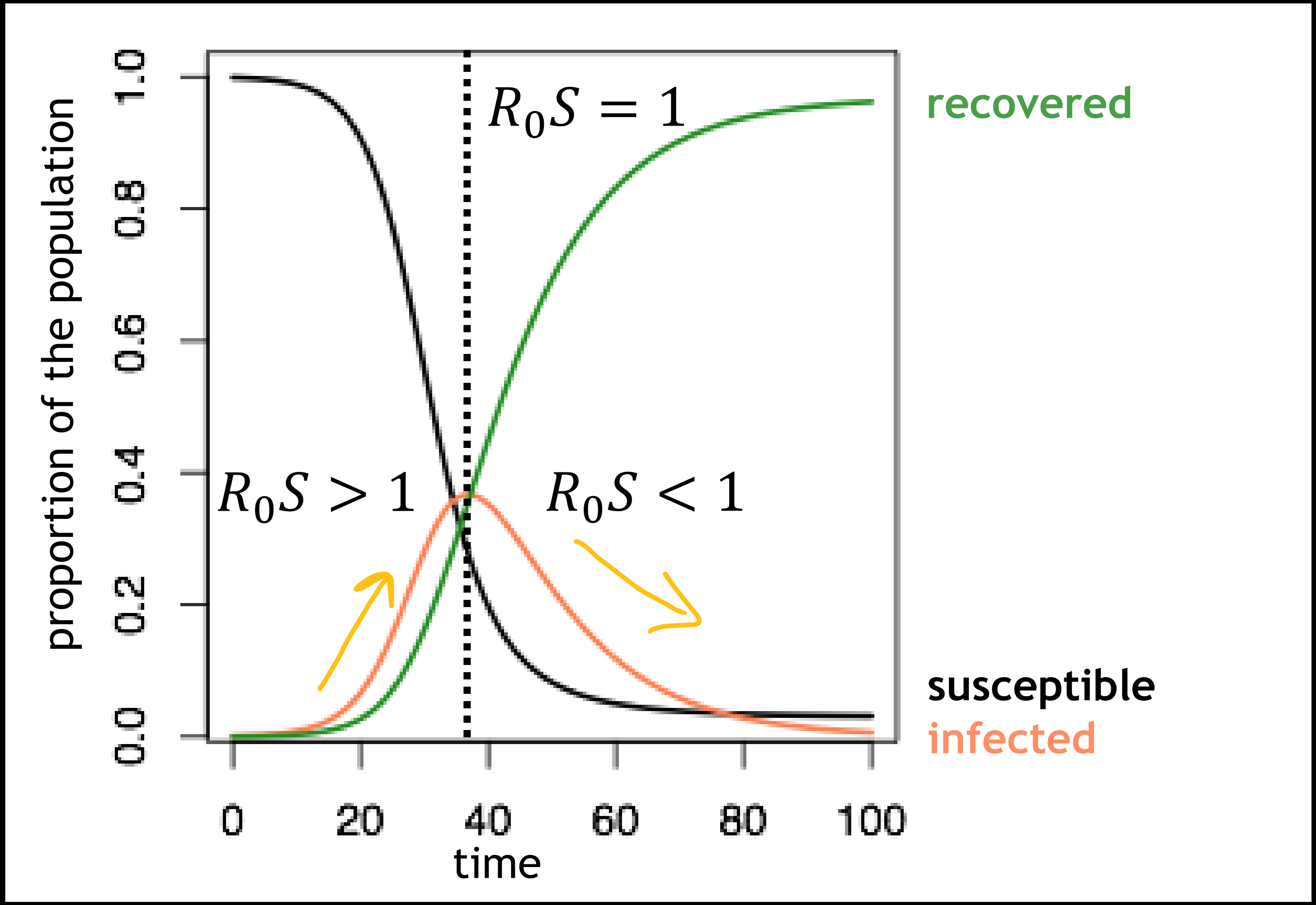


- Can translate R0 to R based on proportion of susceptibles at specific time points

$R_0 \to R_t$

epidemic stable? growing? decreasing?
R=1 R>1 R<1

Reproductive Number during Epidemic



- The basic reproductive number is driving the behavior of the epidemic
  - how we calculate R0 depends on the type of model
- During the epidemic, we can use the relationship between R0*S and 1 to see when the epidemic will increase, stabilize, or decrease

Model Development

- Important to reflect on purpose of model & wider considerations
- 9-step process from Habbema *et al.*, 1996

1. Identify the question
2. Identify existing knowledge
3. Choose model structure
4. Choose modeling method
5. Model quantification
6. Model validation
7. Prediction & optimization
8. Decision making
9. Model transfer

Upcoming Workshop

Surveillance

Surveillance Workshop Topics

- Surveillance Goals & Uses
- Case-based Surveillance
- Serological Surveillance
- Molecular Surveillance
- Genetic Surveillance
- Wildlife/Veterinary Surveillance
- Tracking Vaccination
- Surveillance Assessment
- Example Programs
- Bangladesh Case Study
- Bangladesh Assessment

R Session Topics

- Exploratory Analysis
- Data Manipulation
- Mapping
- Tidyverse
- Debugging
- Plotting & Visualization
- Forecasting

Additional Materials

- To be sent:
  - Link to videos/materials from modeling workshop
  - Syllabus (when dates are set)
  - Copies of slides/script

Workshop Schedule

Time	Topics
~~2:00—2:10 pm~~	~~Outline & Introduction~~
~~2:10—2:50 pm~~	~~Model Review~~
~~2:50—3:00 pm~~	~~Upcoming Workshop~~
3:00—3:10 pm	Break
3:10—4:00 pm	Introduction to R