Mathematical Modeling in Biology

Glenn Ledder

Department of Mathematics
University of Nebraska-Lincoln
gledder@unl.edu

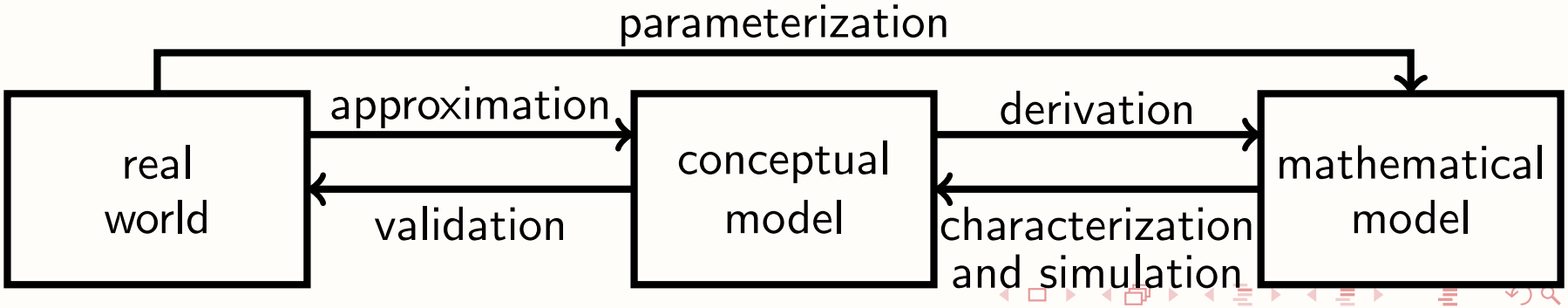
CURE Workshop, June, 2021

General Principles of Modeling

- Modeling contrasted with mathematics
- What are mathematical models?
- Models as functions of parameters

Mathematics vs Modeling

- Mathematics
  - Assumptions define the setting.
  - Conclusions follow from mathematical logic.
  - Focus is on proof.
- Modeling
  - Assumptions define a *conceptual model* of a real setting.
  - Conclusions *for the model* follow from mathematical logic.
  - Conclusions *for the setting* are only as good as the conceptual model.
  - Focus is on checking results against known outcomes.



What is a Mathematical Model?

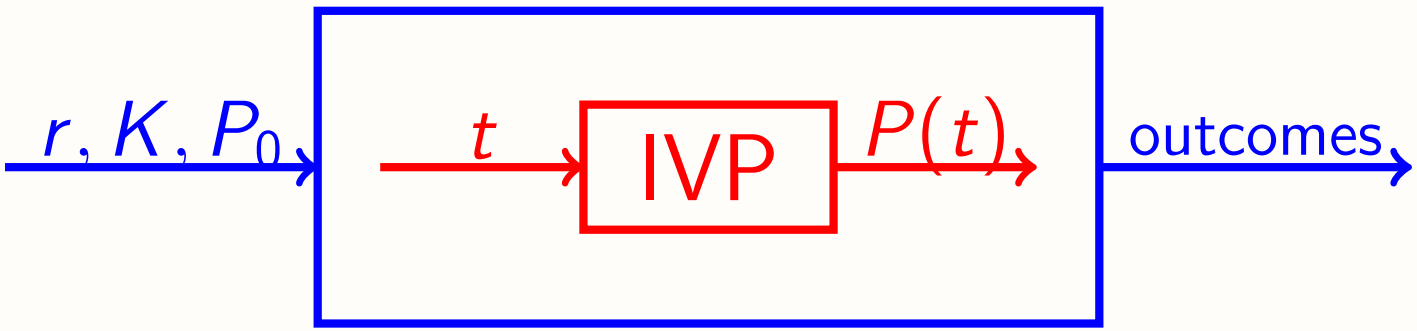
- A **mathematical model** is a self-contained collection of one or more variables together with a set of rules (usually formulas and equations) that prescribe the values of those variables.
  - Models serve as an approximate quantitative description of some actual or hypothetical real-world scenario.
  - Models are created in the hope that the behavior they predict will capture enough of the features of that scenario to be useful.
  - The value of a model depends on the setting to which it is applied and the questions it is used to address.
- A **mechanistic model** is a mathematical model based on assumptions about the scientific principles that underlie the phenomena being modeled.

Models as Functions of Parameters

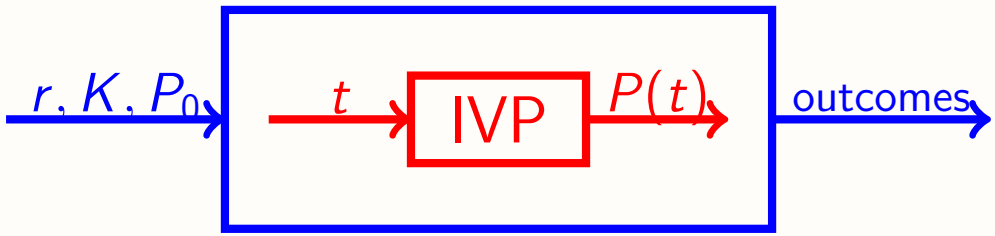
How do we view the logistic growth model

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{K}\right), \quad P(0) = P_0 > 0, \qquad r, K > 0?$$

- **Narrow** view:
  Initial value problem for $P(t)$, with parameters r, K, and P_0.
- **Broad** view:
  Function that maps parameters r, K, and P_0 to outcomes.



Models as Functions of Parameters



- **Narrow** view: Math problem with fixed parameters.
  - The narrow view is used to determine the outcomes.
    - Narrow view questions are trivial: "Given $K = 10$, $R = 1$, and $P_0 = 1$, when does the population reach $P = 5$?"
- **Broad** view: Math problem with outcomes as functions of parameters.
  - The important questions are in the broad view.
    - Do solutions with any initial condition always approach K?
    - At what point is the population growth the fastest?

An Example – the SEIR Epidemic Model

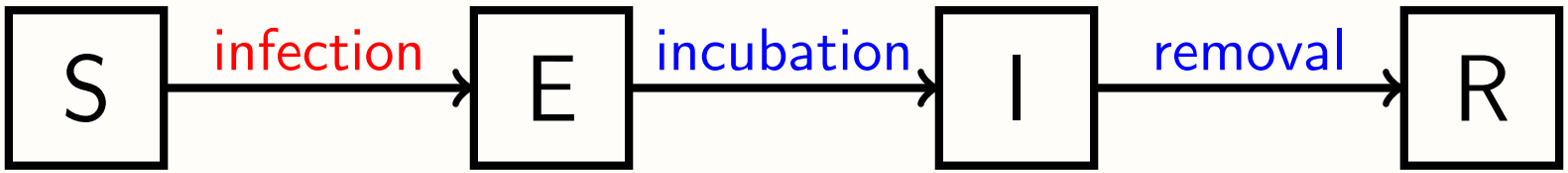
1. Assumptions and mathematical derivation
   - 1.1 Class structure
   - 1.2 Processes
   - 1.3 Differential equations
   - 1.4 Basic reproductive number
2. Designing an investigation
   - 2.1 Asking questions
   - 2.2 Choosing study parameters and outcomes
3. Overview of Methods
4. Addressing questions and reporting results
   - 4.1 Parameter studies
   - 4.2 Using simulations
   - 4.3 Simulations can suggest conjectures

1.1. Class structure

- Individuals in a population are divided into classes. These vary among different epidemiological models. For SEIR:
  - **S**: *Susceptible* – can be infected
  - **E**: *Exposed* – infected but not infectious
  - **I**: *Infectious* – can transmit the disease to susceptibles
  - **R**: *Removed* – no longer infectious
- Sometimes the names are misleading.
  - '*Exposed*' should be '*Latent*' (already infected, not merely exposed)
  - *Removed* includes people who are still sick and may include people who are deceased
- Models are designated by the class structure: SIR, SIS, SEIR, SEAIR, SEAIRHD etc

1.2 Processes

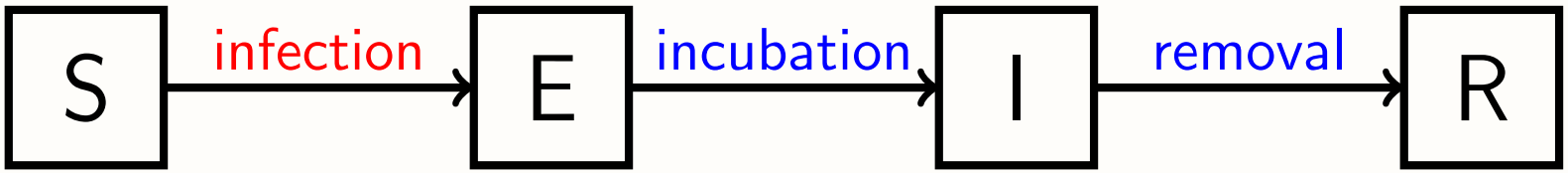
- Processes move individuals from one class to another.
  - Example: Basic SEIR model



- Rate of change of S is $-$ **infection**
- Rate of change of E is **infection** $-$ **incubation**
- Rate of change of I is **incubation** $-$ **removal**
- Rate of change of R is **removal**

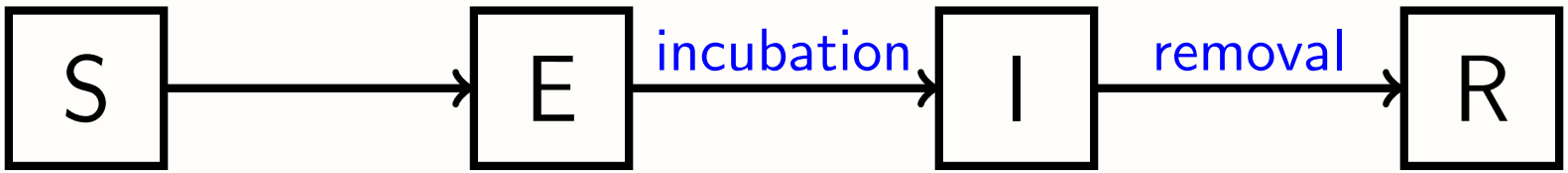
1.2 Processes

- Processes are either **transmissions** or **transitions**.



  - **Transmissions** require interaction with another class.
    - Susceptibles are infected by Infectives.
  - **Transitions** happen without any interaction.
    - Incubation of Latents and removal of Infectives happen spontaneously (but perhaps in phases).

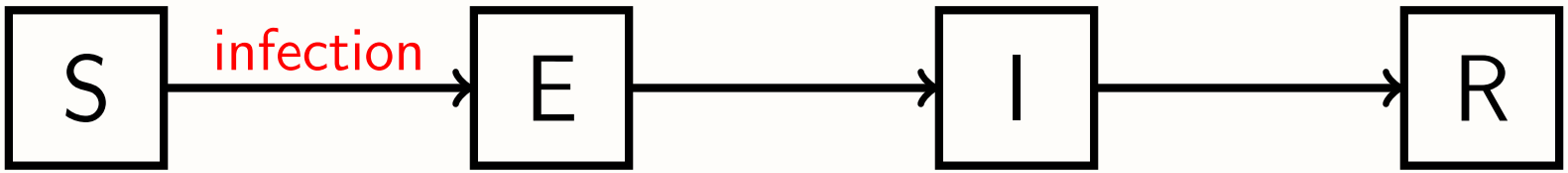
1.2 Transitions



- Transition rates are proportional to the **leaving** class (assuming one-phase transitions)
  - incubation rate $=$ constant $*$ $E = \eta E$
  - removal rate $=$ constant $*$ $I = \gamma I$

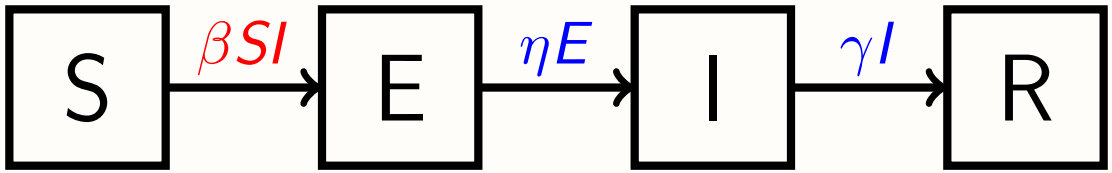
- **Rate constants are reciprocals of average time in class.**
  - Average removal time 10 days $\rightarrow \gamma = 0.1$

1.2 Transmissions



- Transmission rates are proportional to the **leaving** class size
  - infection rate $=$ force of infection $* S = \lambda S$
- The force of infection is (usually) proportional to the sum of the **transmitting** classes (just I for SEIR)
  - force of infection $=$ constant $* I = \beta I$
- The infection rate is $\beta I * S = \beta SI$

1.3 Differential equation model



$$S' = -\beta SI$$

$$E' = \beta SI - \eta E$$

$$I' = \eta E - \gamma I$$

$$R' = \gamma I$$

- Let $N = S + E + I + R$. Then $N' = 0$, so N is constant. (*Without loss of generality, we can take* $N = 1$.)
- The model is an **autonomous dynamical system** (rates of change depend only on the state of the system).

1.4 Basic reproductive number

- **Basic reproductive number** $\mathcal{R}_0$:
  the average number of secondary infections caused by one primary infective in a fully susceptible population.
  - $\mathcal{R}_0 > 1$ is needed to start an epidemic.

- The total number is the average rate times the average time.

- Calculation of average transmission rate:
  - Recall that the **transmission rate** is βSI
  - Transmission rate **per infective**: βS
  - Rate per infective in a **fully-susceptible population**: βN

1.4 Basic reproductive number

- **Basic reproductive number** $\mathcal{R}_0$:
  average transmission rate per infective in a fully susceptible population multiplied by average time in the Infectious class.
- Average transmission rate per infective in a fully susceptible population: βN
- Average time in infectious class: $1/\gamma$ (reciprocal of γ)
- Average number is average rate times average time:

$$\mathcal{R}_0 = \beta N \cdot \frac{1}{\gamma} = \frac{\beta N}{\gamma}.$$

- Other diseases (like COVID-19) can be more complicated.

2.1 Asking questions

- Models must be designed to answer specific questions.
  - If we want to know the impact of COVID-19 on health care resources, we need to modify the SEIR model to track hospitalizations and/or ICU patients.
  - If we want to know the impact of mitigation strategies on COVID-19, we need to build mitigation into the model.
    - My COVID-19 model is SEAIHRD with λ impacted by testing and a contact factor.
- Some common question types:
  - Is a specific claim supported by modeling or not?
  - What effect does parameter x have on outcome y?

2.2 Choosing study parameters

- Model parameters are not always the best study parameters, especially when they are hard to measure.
  - Transition times are more fundamental than transmission rate constants. (Take $\gamma = 1/T_i$)
  - The transmission parameter β is dependent on population size, while the basic reproductive number $\mathcal{R}_0$ is a fundamental disease property. (Take $\beta = \gamma \mathcal{R}_0 / N$.)
- Some parameters are more fundamental than others.
  - The effect of the disease duration T_i, given fixed $\mathcal{R}_0$, is simply to change the time scale for the results.
  - Changing a model to dimensionless form (see Ledder, *Mechanistic Modeling*, Section 5) eliminates scale parameters.

2.2 Choosing outcomes

- Maximum number of new infections?
- Maximum number of hospitalizations per million? (compared to an average of 2800 hospital beds per million in the US)
  - Serves as a measure of the stress on the health care system
- Percent deaths? (0.1% in the US is 325,000 people)
  - Serves as a measure of the human cost
- Final fraction of susceptibles?
  - Serves as a measure of the risk of a new outbreak
- Times for any of these events?

3. Overview of methods

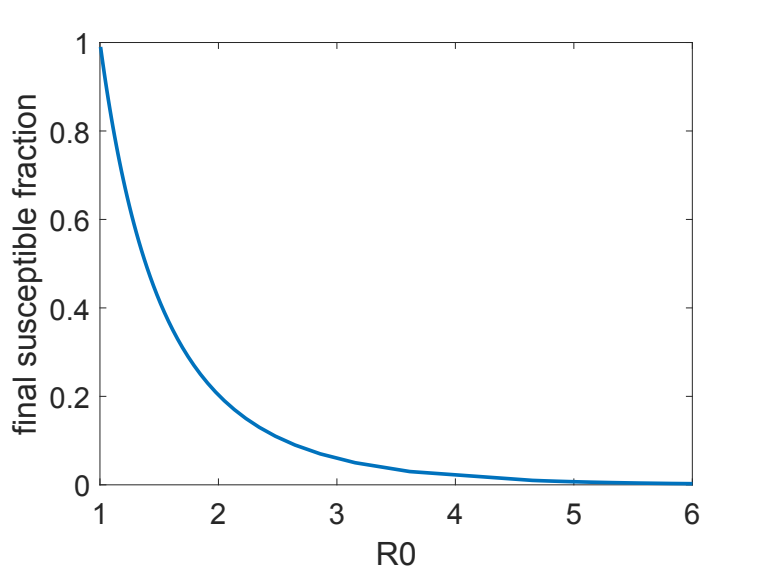
- Solution – seldom useful
  - Very few differential equation problems can be solved using methods of calculus. Even when possible, the results are often less useful than those obtained using other methods.
- Characterization – using hand computation to obtain properties of a model
  - We can often determine long-term behavior using mostly algebra.
  - Results are often general (parameters left unspecified).
- Simulation – using numerical computation to obtain results for individual scenarios
  - Results are never general (parameters must be specified).
  - Can still be used to address general questions.

4. Addressing questions and reporting results

- Answers to math questions are often numbers or formulas. **Modeling questions require verbal answers**, supplemented with visual aids.
  - 'The graph goes up and then comes down' is merely a *description*. An *explanation* connects to the real world scenario and **offers a reason** for the observed results.
- Graphs must be informative and not misleading.
  - No negative values for populations or parameters.
  - Axes must be labeled.
  - Sometimes multiple curves on the same axes are more informative than multiple graphs.
  - Measured data should be plotted as points; simulation results should be plotted as dot-to-dot 'curves'.

4.1 Parameter studies

- Sensitivity analysis can determine how important a parameter is, but it does not determine the parameter's effect.
- Parameter studies determine the quantitative effect of a parameter on one or more outcomes.
  - SEIR example: Assume no initial immunity or mitigation. How does the fraction of people who don't get the disease depend on the basic reproductive number? (Homework problems 2–3)



4.2 Using simulations

Modeling with simulations requires two program components:

1. A function that
   - Accepts input values for
     - Disease parameters, like γ and β,
     - Scenario parameters, like initial and terminal conditions,

     and
   - Returns the time history of the class counts.

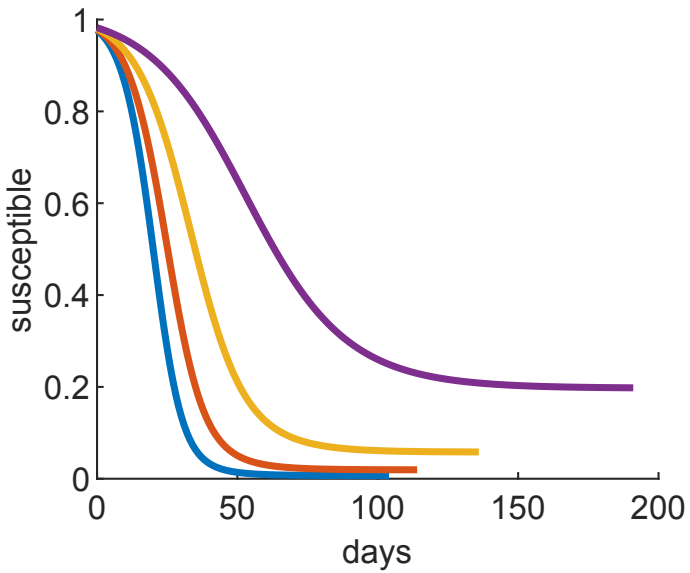
2. A driver script that
   - Organizes an experiment,
   - Utilizes the function to obtain outcomes, and
   - Displays the results.

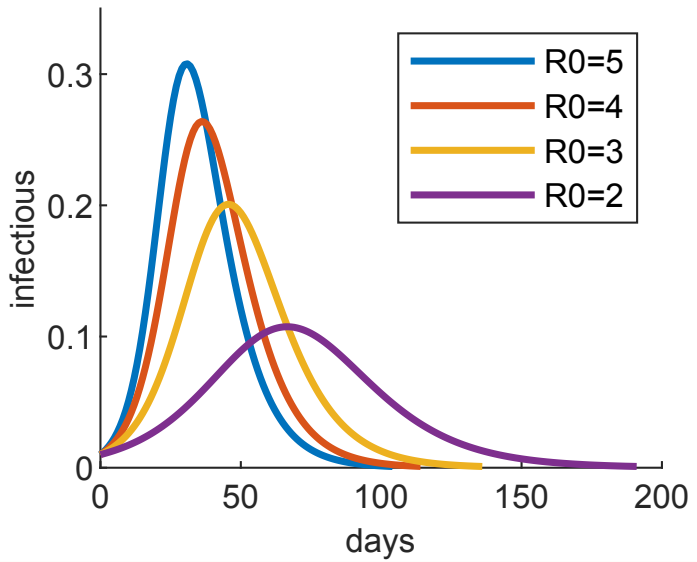
4.2 SIR_paramstudy.R

This driver script plots various outcomes as a function of a parameter. Script elements:

1. Define function sir_sim(beta,gamma,I0,V,target).
2. Define default scenario values.
3. Prescribe range and count of study parameter values.
4. Set up data structures.
5. Run loop:
   - Collect study parameter value.
   - Compute derived parameter values.
   - Use function sir_sim to collect results.
   - Add results to a data structure.
6. Create plots.

4.3 Simulations can suggest conjectures





We can prove these conjectures (Homework problem 3)

1. The fraction of susceptibles is continually decreasing, but is bounded away from 0.

2. The epidemic ends with $I = 0$ and $S = s_\infty$ for some $s_\infty > 0$.

Homework

Problems 1–3 are essential. Problem 4 is 'extra credit.'

1. Use characterization to find a fatal flaw in a commonly-used model.
2. Use an R program to produce the parameter study plot on the 4.1 slide. You will just need to make a few changes to SIR_paramstudy.R.
3. Use mathematics to prove the conjectures on the 4.3 slide.
4. Work through a model research study that could have been done by undergraduates (but without any prompting other than a general question).