Teaching Modeling with Mosquito-Borne Disease Epidemics*

Jeff Knisley, Mathematics
East Tennessee State University
April 8

*Part of the Symbiosis Project funded by the Howard Hughes Medical Institute
Overall Goals

• To create a context which fosters the symbiosis of biology, statistics, and mathematics
  – Multiple entry points for many different students
  – Multiple activities possible and available
• To create a context that can be used with (and become familiar to) a wide range of students
  – Freshman use it superficially; Seniors more depth
  – An “F=ma”-like alternative for the Life Sciences
SIR Epidemics

• A population of constant size $N$ is partitioned into $S =$ Susceptibles, $I =$ Infecteds, $R =$ Recovered

• $C(I) =$ Contact rate of infection (as a function of $I$)
• $r =$ rate of recovery
• There are more sophisticated models
  – We stay simple for pedagogical purposes
  – Simple models are still highly applicable!
SIR Epidemic Model

- SIR models are either Systems of DE’s or Systems of Difference Equations

\[
\frac{dS}{dt} = -C S I
\]

\[
\frac{dI}{dt} = C S I - rI
\]

\[
\frac{dR}{dt} = rI
\]

\[
S_{t+1} = S_t \Delta t - C \left( \sum_{i=0}^{t} S_i \right) - rI_t
\]

\[
I_{t+1} = I_t \Delta t + C \left( \sum_{i=0}^{t} S_i \right) - rI_t
\]

\[
R_{t+1} = R_t \Delta t + rI_t
\]
Types of SIR Epidemics

- Epidemic type often defined by how the disease is spread (Human-Human contact, Insect-borne, water borne, airborne, etc)

- Human-Human Contact: \( C(I) = bI \)
- Mosquito-Human Contact: \( C(I) = c \)

Assuming homogenous mixing of the populations
Which Epidemic is it??

• There have been many instances where one epidemic type has been confused for another
  – Yellow fever is Mosquito-borne (Carlos Finlay, 1889),
    • Proven to be Mosquito-borne in Cuba in 1900 (Maj. Walter Reed). Experiment cost Jesse Lazear his life.
    • Practically no cases of Yellow fever in Cuba since 1901
  – Yellow fever officially remains a human-human transmitted disease (via poor sanitation) until 1905.
    • Last Yellow Fever epidemic, New Orleans, 1905
    • Extensive outbreak in spite of elaborate sanitation efforts
Which Epidemic is it??

• This is still an Important Question – St. Louis Encephalitis, Dengue Fever, etc.

• And is a question we can explore
  – Pedagogically: Via Simulations (Netlogo) that generate data for each epidemic type (via models)
  – Statistically: Form Hypothesis, Analyze Data, Get $p$-value, Infer epidemic type
  – Mathematically: Why we should not be surprised that epidemic types are often confused?
Pedagogically: The Simulations

• Simple S-I-R interactions among N agents
  – All Simulations: Recovery with a fixed rate $r$ depending on how long an individual is ill (on average)
  – Human-Human: Infections with a given probability when a Susceptible meets an Infected
  – Mosquito-Human: Infections with a given probability when a Susceptible meets a Mosquito (assuming all Mosquitoes are infected)

• We simulate a “village” near a river (N = 500)
A white susceptible (S) becomes a pink infected with "ProbabilityOfInfection" upon meeting an infected. Infecteds (I) turn red while symptomatic and then turn dark green once they recover (R).

S (yellow), I (red), and R

Number

0 500

0 hours 775
Pedagogically: The Simulations

• But there is more to the story
  – People work, eat, drink water during the day
  – People sleep at night (violates homogeneity assumptions)
  – When they are sick enough, they want medical attention (hospital)

• Given a randomly-generated epidemic, can a student determine what type of epidemic it is?
  – Our “real world” is not so pristine as the models
  – But features of the models do “survive”
A white susceptible (S) becomes a pink infected with "ProbabilityOfInfection" upon meeting an infected (I). Once symptomatic, infecteds turn red and head to the Hospital. Red infecteds turn dark green once they recover (R).
Statistically: The Data

• Mosquito-Human epidemics are likewise set in “real world” circumstances
  – Model assumptions are violated
  – But same assumptions violated for all models!

• Although the models are developed mathematically, they are compared statistically
  – Especially given that assumptions are often violated
  – And that epidemics of different types tend to “look alike” (later)
Once bitten, a white susceptible (S) becomes a pink infected (I) with "ProbabilityOfInfection". Once symptomatic, infecteds turn red and head to the Hospital. Red infecteds turn dark green once they recover (R).
Statistically: Analyzing The Data

• EpidemicVector simulation randomly simulates one of 4 epidemic types (Human-Human, Mosquito-Human, Food Borne, Water Borne)
  – Based on a 5 digit number the student provides
  – They must determine which epidemic is generated

• Simulation mechanics:
  – People eat/drink during the day, sleep in huts at night
  – Once sufficiently ill, they go to the hospital
  – Only the patient count per hour is available for analysis
Epidemic Simulations: 4 possible Methods of Transmission

- Set Up
- Go
- Ticks Per Hour: 5
- On/Off: Reduce Mosquitos
- On/Off: Provide Meals
- On/Off: Provide Water

Hospital Patients Per Hour:
- Number of Patients
- Hours: 0 to 240
Statistically: Analyzing The Data

• Students can apply any of 3 interventions
  – Reduce the mosquito population
  – Provide untainted food to some of the population
  – Provide sterile water to some of the population

• Question: Which intervention reduces severity of the epidemic?
  – Simple approach: collect “max patient counts” in each simulation
  – Test for significance of difference between means of Experimental (with intervention) and Control
  – We have them do so with the randomization test
Result of Simulation with the data associated with the button **Significant**
The Randomization Test

Result of Simulation with the data associated with the button **InSignificant**

Proportion of the percentage or trial means whose magnitude exceeds that of the original input data.
Mathematically: Linearization

- Mosquito-Human model is the *linearization* of the Human-Human model (in simple forms)

\[
\begin{align*}
\frac{dS}{dt} &= -bIS \\
\frac{dI}{dt} &= bIS - rI \\
\frac{dR}{dt} &= rI \\
\frac{dS}{dt} &= -cS \\
\frac{dI}{dt} &= cS - rI \\
\frac{dR}{dt} &= rI
\end{align*}
\]

\( c = b \left( I_0 + S_0 \right) \)
Mathematically: Linearization

• Linearization is ‘valid’ as long as $R(t)$ is small
  – Thus, epidemic types often look “the same” for relatively long periods of time
  – Long term: Linearization (Mosquito) eventually infects everyone in the population (but may take “arbitrarily long” to do so )

• Deep Mathematical Question: On what interval is a linearization a ‘valid’ approximation?
Goal: Contexts that become “familiar” and can be used in many situations

- Freshman, General Majors: Simulation generates data for statistical testing
- Sophomore, Math Majors: Linear difference equation as both (a) programming assignment and (b) excuse to study sequences and series
- Junior/Senior: Systems of DE’s, Linearization of DE’s, statistical modeling, time series analysis
Thank you!

Website: http://math.etsu.edu/Symbiosis/epidemics

- Data available from command line in Netlogo
- Netlogo interface with R can be used with Desktop Version of Netlogo (not enabled with applets)