ISyE 6644 Mini-Project 1: Pandemic Flu Spread

Robert King, Joseph Richardson, Daniel Rodgers

June 16, 2021

***Abstract—***This project presents a simple model of the transmission of influenza through a grade school classroom, attempting to characterize the distribution of the duration and extent of spread through stochastic simulation. Given the stated assumptions of the model, a single case of influenza entering this classroom model is likely to lead to only 3 or 4 additional infections, with the spread occurring over about 7 days. [Still needs something, y’all…]

# PROBLEM DESCRIPTION

The prompt describes a model of a 21-child elementary school classroom in which one student starts the school day infectious with influenza. The other 20 children are assumed to be susceptible. An assumed amount of contact and probability of transmission are presented, as is an “infectious window” during which each child is capable of passing the virus. All children are presumed to show up for each day of school. We are asked to compute several statistical estimates based on these assumptions, then create a simulation of these conditions and use it to estimate the distribution of likely duration of the spread of virus through the class.

To be more specific, the assumptions of this model are as follows:

* The simulation is confined to this one classroom. No inputs or interactions with other classrooms need be considered.
* We are only concerned with these 21 children. Teacher and community contacts are not considered.
* Each child who has not yet been infected is equally susceptible.
* Each child who is currently infected will
  + Have an equal amount of contact with each susceptible child.
  + Be capable of spreading the virus for exactly 3 days
  + Show up for school despite being infected.
  + Shed virus to an equal degree.
* The above assumptions lead to the assumption that every contact between an infectious child and a susceptible one can be modeled as a random, independent [](https://www.codecogs.com/eqnedit.php?latex=Bernoulli(0.02)#0) trial.
* Once infected, a child will become contagious as of the next day, and as above, will remain so for 3 days--viral incubation time is negligible.
* After a child’s 3-day window of contagion, they will eventually recover from the flu and be fully resistant to it--reinfection and viral mutation are not considered.

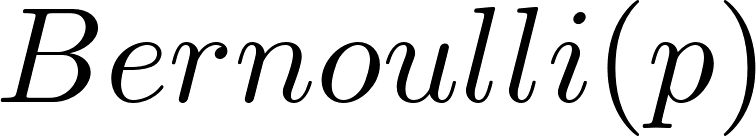
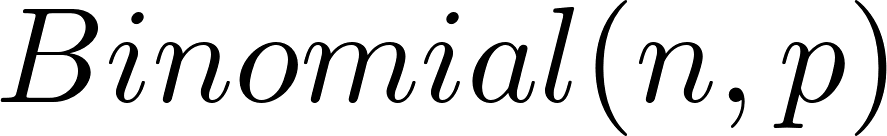
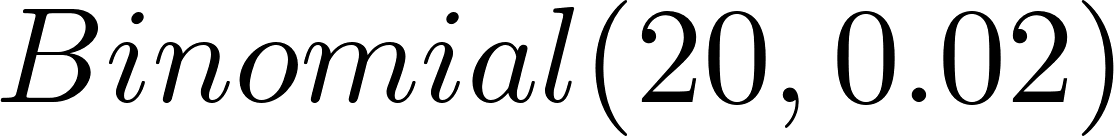
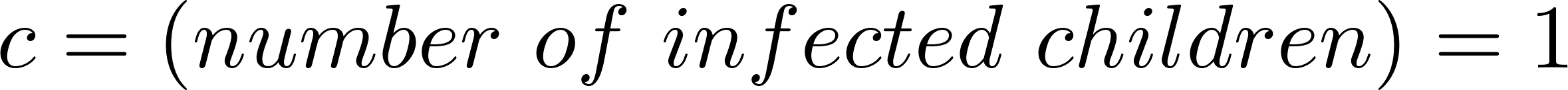
This problem is a small-scale variant of the well-established SIR model of epidemiology. SIR models, which typically operate at the large-population scale, attempt to categorize the populace into susceptible (S), infectious (I), and recovered (R), with population members moving from one category to the next according to fairly straightforward assumptions about the rate of new infections over time and the speed of recovery once infected. These models generally operate at such a large scale that the discrete variables underlying them (individual patients and units of time) can be treated as continuous. These simplifying assumptions make such models more analytic than stochastic.

Our model “errs” in the opposite direction, by using a small enough population that it is easy to track individuals and discrete events through the course of one simulated epidemic. The simplifying assumptions of our model are geared to make a stochastic approach more practical.

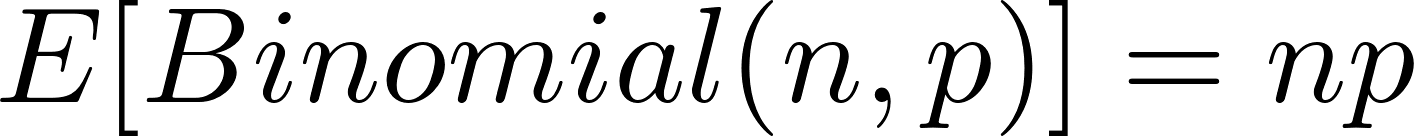
Another concept we are quietly addressing with this model is the basic reproduction number of an infectious disease, commonly abbreviated [](https://www.codecogs.com/eqnedit.php?latex=R_0#0). This is the number of new infections an infectious patient is expected to produce, all other factors being equal. This concept, which we allude to with our Bernoulli probability of infection, incorporates the inherent infectivity of the pathogen, its vectors and hardiness to environmental conditions, and the amount of pathogen a typical individual sheds. While this is a lot of heavy lifting for one number to do, in practice, it is a useful simplifying assumption, as changes in any of those factors can be modeled by a multiplier to [](https://www.codecogs.com/eqnedit.php?latex=R_0#0). Likewise, our model could simulate a behavioral intervention or mutation to the virus simply by changing the probability of a given contact being “successfully infectious”.

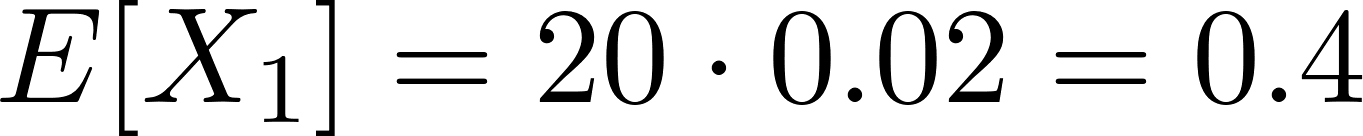
# SOLUTIONS

## A—What is the distribution of the number of kids that Tommy infects on Day 1?

Since each uninfected kid in the class represents one [](https://www.codecogs.com/eqnedit.php?latex=Bernoulli(p)#0) trial, at it is stipulated that we are modeling each "infection opportunity" as independent, the overall distribution of kids infected in one day will be [](https://www.codecogs.com/eqnedit.php?latex=Binomial(n%2C%20p)#0) where [](https://www.codecogs.com/eqnedit.php?latex=n#0) is the number of uninfected kids. Thus, [](https://www.codecogs.com/eqnedit.php?latex=Binomial(20%2C%200.02)#0) on day 1 with [](https://www.codecogs.com/eqnedit.php?latex=c%20%3D%20(number%5C%20of%5C%20infected%5C%20children)%20%3D%201#0).

## B—What is the expected number of kids that Tommy infects on Day 1?

[](https://www.codecogs.com/eqnedit.php?latex=E%5BBinomial(n%2Cp)%5D%3Dnp#0)

[](https://www.codecogs.com/eqnedit.php?latex=E%5BX_1%5D%3D20%20%5Ccdot%200.02%3D0.4#0)

## C—What is the expected number of kids that are infected by Day 2?

The random variable X represents the number of kids infected in a day. Let X1 and X2 represent the number of kids infected on day one and day two, respectively. Kids infected within a day are iid, however X1 and X2 aren’t independent since X2 changes based on the outcomes of X1.

For example, if Tommy infects one student on day 1, then there will be two spreaders on day two. If we assume that a child can only be infected once, then the probability that a kid is infected by both spreaders on day 2 is zero. Let A be the event that a kid is infected by spreader 1, and B the event that a kid is infected by spreader two. Then:

So, the probability of a new infection on day 2, given there was one infection on day 1, is 0.04. An equation for be can be generalized for any number of spreaders:

Where r1 is the number of kids infected on day 1. Therefore,

Let X\*2 represent the number of students infected by day 2. Since the number of kids who could be infected on day two depends on the number who were infected on day one, we’ll use conditional probabilities for day 2. For example, the probability that only 1 kid is infected by day 2 (excluding Tommy) is:

We then follow this method to calculate each value of X, from 0 to 20. Table 1, below, shows the probability distribution for X\*2, with Tommy included.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***X*** | ***1*** | ***2*** | ***3*** | ***4*** | ***5*** | ***6*** | ***7*** | ***8*** | ***9*** | ***10*** |
| ***P(X)*** | ***0.4457*** | ***0.30738*** | ***0.15194*** | ***0.06306*** | ***0.0224*** | ***0.00697*** | ***0.00193*** | ***0.00048*** | ***0.00011*** | ***0.00002*** |

***Table 1—***Selected values from the probability distribution for the number of infected kids by day 2 (including Tommy). For X > 10, the when rounded to the nearest hundred-thousandths place.

We now use the expected value formula to find the expected number of kids infected by day 2:

## D—Simulate the number of kids that are infected on Days 1,2,. . . . Do this many times. What are the (estimated) expected numbers of kids that are infected by Day i, i= 1,2, . . .? Produce a histogram detailing how long the “epidemic” will last.

See the appendix for the full Python code of a Monte Carlo simulation. The simulation abstracts students as columns in a matrix, storing both whether they are currently infected and how long their infection has lasted. It then loops for up to 63 simulated days, incrementing counters for number of days since start of run, “infection timer” for each infected student, and day of week. If the day of week parses as “weekend”, students are considered to be out of class, so no random infection trials are conducted. Otherwise, on each day an internal loop conducts a random Bernoulli(0.02) trial for each susceptible student, provided there is still at least one contagious student in the class. Note that this adds the further assumption, not mentioned above, that the school follows a typical American calendar with students attending class Monday through Friday, with Saturday and Sunday off. Further, all simulations assumed a “Monday start”: the simulation always begins with patient zero arriving on a Monday, with that being day 1 of their contagious period.

The simulation keeps a running total of the number of infected since the start of that run, and runs for 63 simulated days or until there are no more contagious students, whichever comes first. The code runs 10,000 replications. The duration of the epidemic for that run is computed from the difference between the number of days since Tommy’s infection began and the duration of the last (shortest) infection. A histogram of lengths appears in the appendix. The mean duration was 8.5084 days, with a median of 7 days. The mean number of infections over the course of the epidemic was 5.9568, with median 4 infections. Clearly, given the stated assumptions, the infection is unlikely to make its way through every child in the class. Nor would it be expected to last more than a few days. In many runs of the sim, the weekend rule did not even take effect before the run ended for lack of contagious children.

The appendix includes a table of the mean number of infected children by day of duration. In the rare cases where more than a few children were infected, the epidemic naturally tended to last longer overall, but there appeared to be an asymptote at 32 days’ duration and mean 5.968 total infected students. Even in the worst circumstances, this infection appears unlikely to last longer than a month or infect more than 6 students total.

# CONCLUSIONS

This simulation modeled the extent and duration of an infectious disease spreading through a small population.

As mentioned above, this is a small scale example based on a number of simplifying assumptions. Adjusting some of the core assumptions of this model could lead it in interesting new directions. The most obvious might be to “scale it up”: to expand the simulation to encompass several classrooms at once, including some modeling of how separate classrooms interact with one another. A further scaling up could model each “classroom” as one community in a network of adjacent towns, although that would alter some other underlying assumptions, and likely the time scale of the simulation.

A variation of this simulation in Arena could assign differing susceptibility to each entity in the system, modeling individual variations in health status and behavior. The “shape” of the distributions of duration and extent of infection in our model are highly dependent on the simulated [](https://www.codecogs.com/eqnedit.php?latex=R_0#0) of our virus. An alteration of its basic reproduction number or the behavioral characteristics that factor into the probability of transmission could radically change the outcome. Such a “tweak” to transmission probability could also be used to model behavioral or health interventions, like social distancing, hand washing, or a vaccine.

# APPENDIX

[histogram of days’ duration]

[table and/or plot of number of infected per day]

[code, with or without comments]

# REFERENCES

1. Hethcote H.W. (1989) Three Basic Epidemiological Models. In: Levin S.A., Hallam T.G., Gross L.J. (eds) Applied Mathematical Ecology. Biomathematics, vol 18. Springer, Berlin, Heidelberg. Accessed at <https://www.mtholyoke.edu/~ahoyerle/math333/ThreeBasicModels.pdf> on 6/16/21
2. Jones, J.H. Notes on R0. Dept. of Anthropological Sciences, Stanford University. (2007). Accessed at <https://web.stanford.edu/~jhj1/teachingdocs/Jones-on-R0.pdf> on 6/19/21.
3. Liu, J.M., Xia, S. *Computational Epidemiology*, Health Information Science. Springer Nature Switzerland AG. (2020).
4. Taghvaei, A., Georgiou, T.T., Norton, L. *et al.* Fractional SIR epidemiological models. *Sci Rep* 10, 20882 (2020). <https://doi.org/10.1038/s41598-020-77849-7>