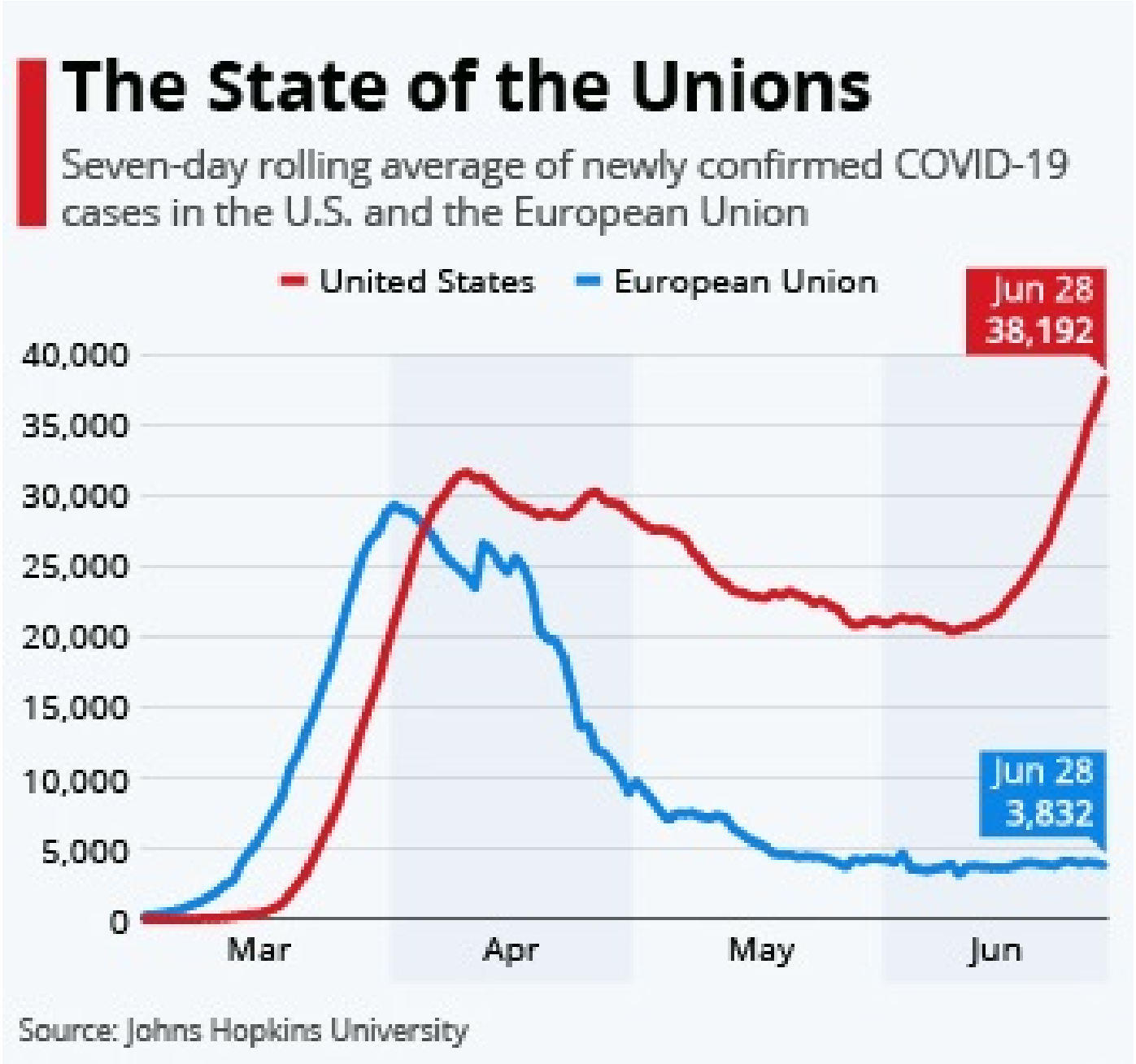
Introduction to the Mathematics of Epidemics on Graphs

Rick Durrett, Duke Math


The State of the Unions
Seven-day rolling average of newly confirmed COVID-19 cases in the U.S. and the European Union
United States
European Union
Jun 28
38,192
Jun 28
3,832
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
Mar
Apr
May
Jun
Source: Johns Hopkins University

Epidemics with fixed infection times

The action takes place on a graoh G in continuous time $t \in [0, \infty)$.

SIR epidemic. Susceptibles become infected at rate λ times the number of infected neighbors. An infection lasts for time 1, after which the individual is "Removed," no longer can be infected.

Since the infection time is fixed the events x infects y_i are independent for all the neighbors y_i of x.

Let T be exponential with rate λ, i.e. $P(T > t) = e^{-\lambda t}$. Let

$$\tau_f = P(T \leq 1) = 1 - e^{-\lambda}$$

Epidemic = Percolation

Each edge will be $S - I$ (or $I - S$) only once. Flip a coin with probability τ_f of heads, and keep the edge if the coin shows heads, otherwise delete it. The individuals infected in an epidemic started by x = the connected component of the graph containing x.

Erdös-Renyi graph. There are n vertices. Each pair of vertices is indpendently connected by an edge with probability μ/n. The thinned graph is $ER(n, \mu\tau^f/n)$.

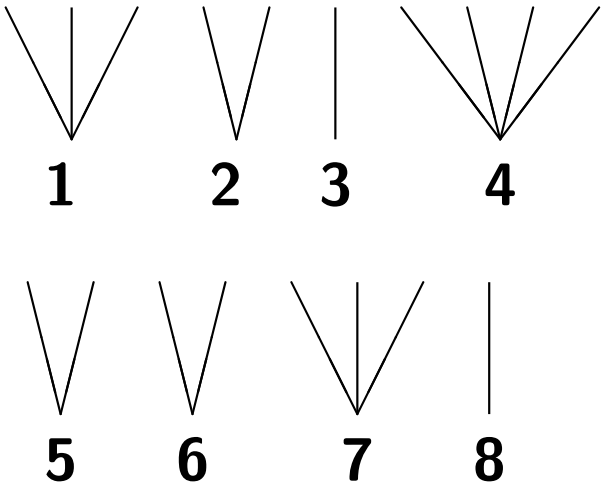
A large epidemic occurs with positive probability if $\mu\tau^f > 1$. If z_0 is the fixed point smaller than 1 of the generating function

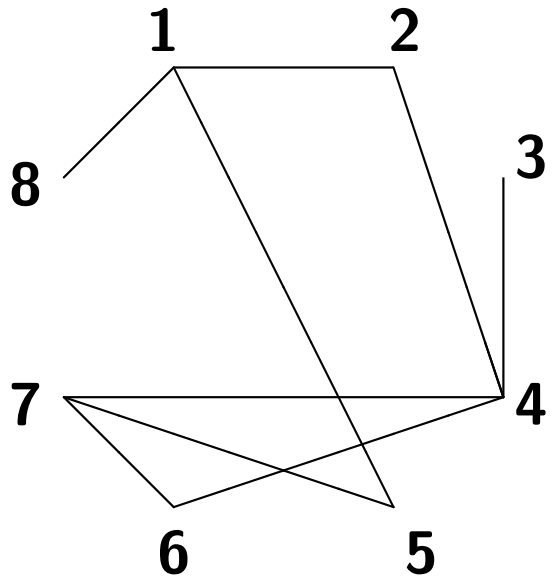
$$G(z) = \exp(-\mu\tau^f(1 - z)),$$

then $1 - z_0$ gives both the limiting probability an infected individual will start a large epidemic, and the fraction of individuals infected when a large epidemic occurs.

Configuration model

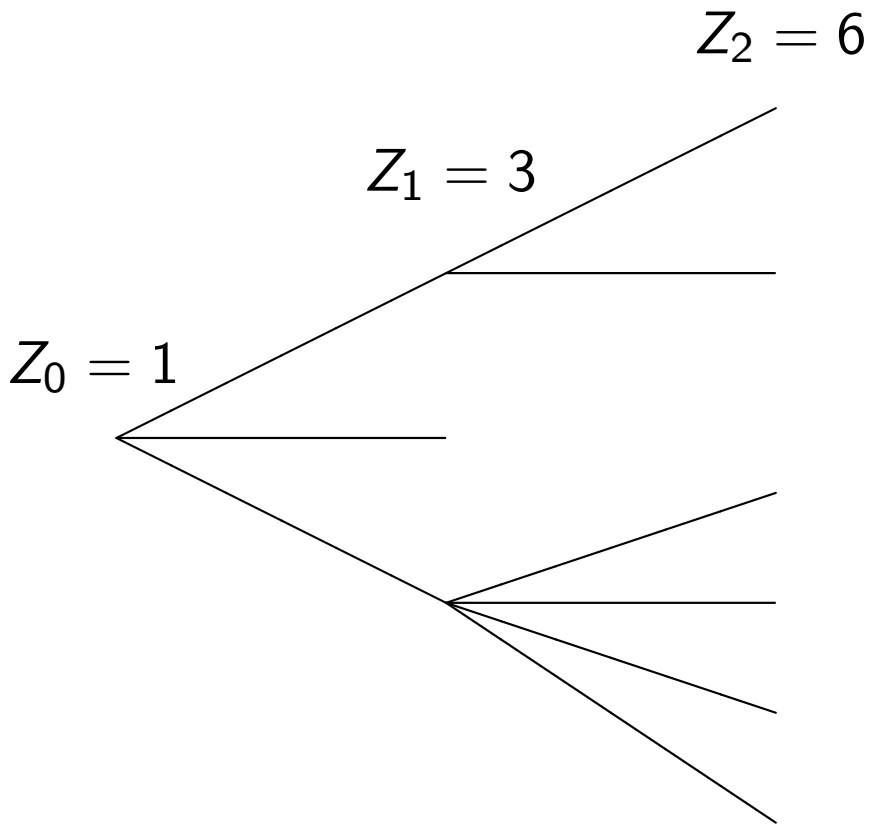
Let $d_1, \ldots d_n$ be independent and have $P(d_i = k) = p_k$. Since we want d_i to be the degree of vertex i, we condition on $E_n = \{d_1 + \cdots + d_n \text{ is even}\}$. To build the graph we think of d_i half-edges attached to i and then pair the half-edges at random.





Branching process viewpoint

Pick a vertex x to start from. Let Z_m be the number of vertices at distance m from x. In the Erdös-Renyi case, when m is small, Z_m is a branching process in which each individual has a Poisson(μ) number of children.



Branching process view of configuration model

First vertex chosen has j neighbors with probability $p_j = P(D_i = j)$.

Since we connect half-edges at random, a first generation vertex with degree k is k times as likely to be chosen as one with degree 1, so the distribution of the number of children of a first generation vertex is

$$q_{k-1} = \frac{kp_k}{\mu} \quad \text{for } k \geq 1 \text{ where} \quad \mu = \sum_k kp_k$$

The $k-1$ on the left-hand side comes from the fact that we used up one edge connecting to the vertex. If p has finite second moment, q has finite mean $\nu = \sum_k k(k-1)p_k/\mu$.

If $\nu > 1$ then there is a component of order n with positive probability.

General infection times

Suppose now that infections have duration S with density function $f_S(s)$. A common choice is exponential(γ) so that the set of infected sites is a Markov process.

Let $\tau_c = P(T < S)$. Infections from x to its neighbors y_i are no longer independent, but that does not matter, since our branching process will be supercritical if the mean number of children $\nu\tau_c > 1$.

$$\hat{G}_1(z) = \int_0^\infty ds\, f_S(s) \sum_{j=0}^{\infty} z^j \sum_{k=j}^{\infty} q_k \binom{k}{j} (1 - e^{-\lambda s})^j (e^{-\lambda s})^{k-j}$$

If $q_k \to p_k$ call this $\hat{G}_0$. When $\nu > 1$ the probability of a large epidemic is $1 - \hat{G}_0(\zeta)$ where $\hat{G}_1(\zeta) = \zeta$ is the fixed point < 1.

Probability of large epidemic $\neq$ final size.

For general infection times, the relevant graph for percolation has oriented edges x to neighbors y.

Probability of a giant component when we start at a randomly chosen vertex is percolation probability when we follow edges in the direction of their orientation. (Edges from x are dependent.)

Final size is probability of a giant component when we start at a randomly chosen vertex and follow edges in the direction OPPOSITE to their orientation. (Edges to x are independent.)

Homework: Compute the probability of a large epidemic and the final size on the complete graph when $S =$ exponential(γ). Out degree is shifted geometric, in degree is Poisson.

Power law random graphs

$p_k \sim Ck^{-\alpha}$, $q_k \sim C'k^{1-\alpha}$. If $2 < \alpha < 3$ then $\sum_k kp_k < \infty$ but $\sum_k kq_k = \infty$, i.e., the branching process has infinite mean.

Internet (physical network of machines) $\alpha = 2.16$ Faloutos[3] (1999)

Movie actor network (connected by movies) $\alpha = 2.3$

Liljeros et al (2001) Sexual network of 4,781 Swedes.
$\alpha_{male} = 3.3$, $\alpha_{female} = 3.5$.

"Internet is robust yet fragile." Percolation probability for random attacks $p_c = 0$ but is easily disconnected by targeting high degree nodes.

For proofs, Bollobás and Riordan (2004) Internet Math 1, 1–35

SIS model

Nodes become infected at rate λ. At rate 1, infected nodes become healthy (and again susceptible to disease)

If infected $\to$ occupied, and susceptible $\to$ vacant this is the contact process invented by Ted Harris in 1974 (on $\mathbb{Z}^d$).

λ_c is threshold for survival, $\theta(\lambda) =$ survival probability, $\theta(\lambda) \sim C(\lambda - \lambda_c)^\beta$

According to degree-based mean field calculations

If $\alpha \le 3$, $\lambda_c = 0$, $\beta = 1/(3-\alpha)$
If $3 < \alpha < 4$, $\lambda_c > 0$ but $\beta = 1/(\alpha - 3) > 1$
If $\alpha \ge 4$, $\lambda_c > 0$. $\beta = 1$

Epidemics in Complex Networks *Rev. Modern Phys.* 87, 925–979

Berger, Borgs, Chayes and Saberi (2005)

Considered the contact process on Barabási-Albert preferential attachment graph, has a power law degree distribution with $\alpha = 3$.

Theorem. (a) The probability that the process will survive from a randomly chosen vertex is $\lambda^{\Theta(1)}$ and hence $\lambda_c = 0$.
(b) With probability $1 - O(\lambda^2)$ the survival probability is

$$\lambda^{\Theta[\log(1/\lambda)/\log\log(1/\lambda)]}$$

Idea: If the infection can reach a vertex with degree $C\lambda^{-2}$ there is survival

Lemma. On the star graph with k leaves, if $k\lambda^2 \to \infty$

$$P\left(\xi^0_{\exp(k\lambda^2/10)} \neq \emptyset\right) \to 1$$

(Shirshendu) Chatterjee and Durrett (2009)

Suppose $\alpha > 3$, $P(d_i \leq 2) = 0$, and G_n has no self-loops or parallel edges.

Theorem. Let ξ_t^1, $t \geq 0$ denote the contact process on the random graph G_n starting from all sites occupied. Then for any $\lambda > 0$, there is a positive constant $p(\lambda)$ so that for any $\delta > 0$

$$\inf_{t \leq \exp(n^{1-\delta})} P\left(\frac{|\xi_t^1|}{n} \geq p(\lambda)\right) \to 1 \quad \text{as } n \to \infty.$$

Main idea: Look at all the vertices with $d_i \geq n^\epsilon$. The infection survives for a long time on their stars. The graph has diameter $O(\log n)$ so if a star becomes healthy it is easy to push infection from another star to it.

Bounds were also proved on the critical exponent β.

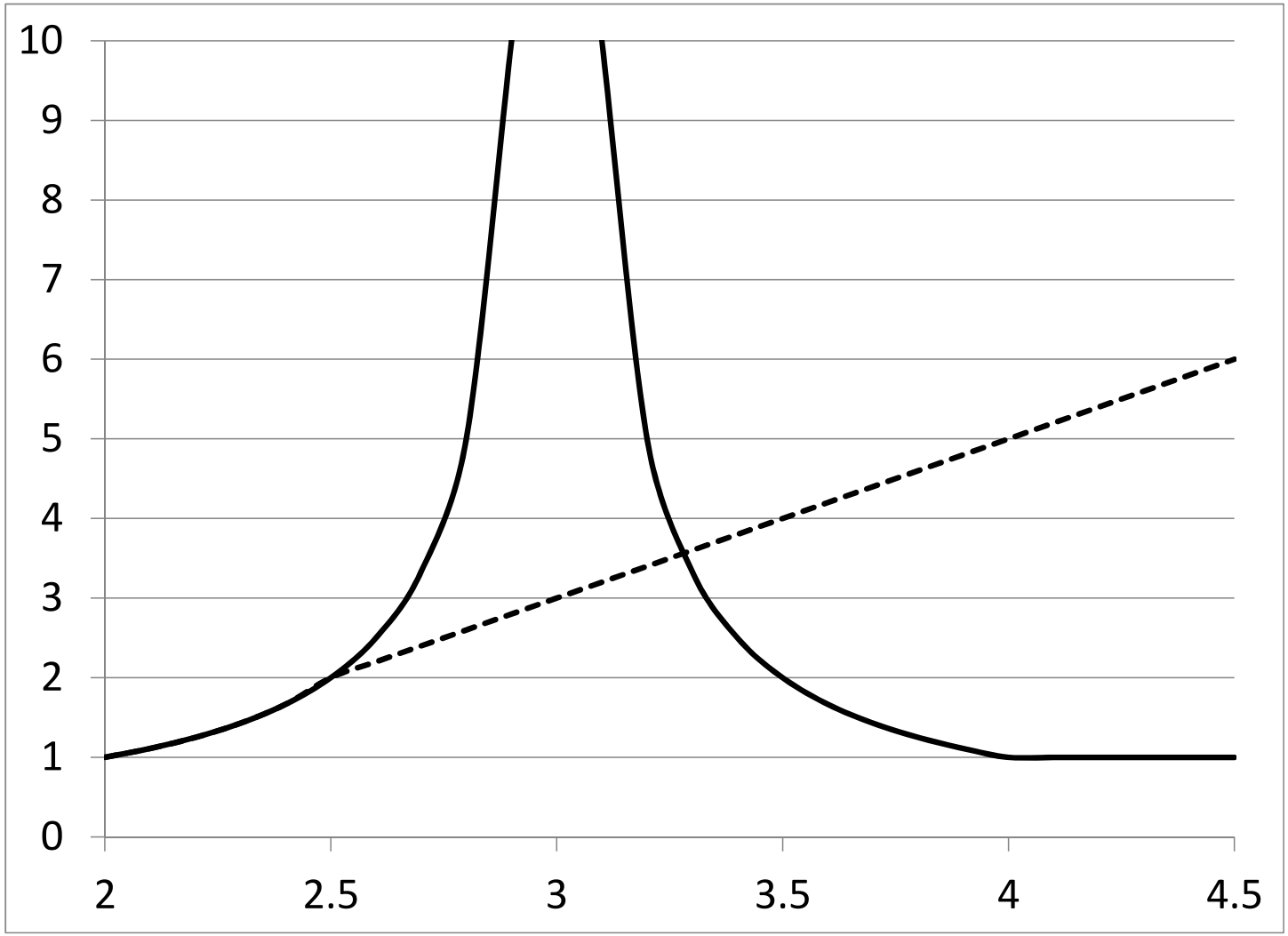
Mountford and friends (2013, 2016)

Mountford, T., Valesin, D., and Yao, Q. (2013) extended the result to cover $2 < \alpha \leq 3$ and proved upper and lower bounds on survival probability, that are the same up to a constant.

$$\theta(\lambda) = \begin{cases} \lambda^{1/(3-\alpha)} & 2 < a \leq 5/2 \\ \lambda^{2\alpha-3} \log^{2-\alpha}(1/\lambda) & 5/2 < \alpha \leq 3 \\ \lambda^{2\alpha-3} \log^{4-2\alpha}(1/\lambda) & 3 < \alpha \end{cases}$$

Mountford, T., Mourrat, J-C, Valesin, D., and Yao, Q. (2016) showed that survival holds for times $t \leq e^{cn}$.

Physics versus Rigorous Critical Exponents



Contact process on Galton-Watson trees.

It is easier to work on the infinite object because there is survival for all time rather than just metastability. On trees there are two critical values: if $\lambda > \lambda_1$ the process does not die out. If $\lambda > \lambda_2$ the root is occupied infinitely often.

Theorem. Pemantle (1992) If the offspring distribution in the Galton-Watson tree is a stretched exponential $p_k = c_\gamma \exp(-k^\gamma)$ with $\gamma < 1$ and has mean $\mu > 1$ then $\lambda_2 = 0$.

Theorem Zoe Huang and Durrett (arXiv:1810.06040) If the offspring distribution p_k for a Galton-Watson tree is subexponential

$$\limsup_{k\to\infty} (1/k) \log p_k = 0.$$

and has mean $\mu > 1$ then $\lambda_2 = 0$.

Bhamidi, Nam, Nguyen, and Sly arXiv:1902.03263

Theorem. Consider the contact process on the Galton-Watson tree with offspring distribution ζ with $E(\exp(c\zeta)) < \infty$ for some $c > 0$, then $\lambda_1 > 0$.

Proof is based on a very simple recursive equation. They also get bounds on the critical value and on survival times. To lower bound survival times they use subtrees with good expansion properties.

Problem. Is $\lambda_1 < \lambda_2$?

DOMath project, Duke Summer 2020

- Molly Borowiak (Apex, NC)
- Fayfay Ning (Long Island, NY)
- Justin Pei (Taipei)
- Sarah Zhang (Birmingham, AL)
- graduate assistant: Hwai-Ray Tung

Two Topics

Duke is unusual because students are required to live on campus 3 out of 4 years, so 80% of students live on campus. How much would spread in the dormitories be reduced if all students had single rooms?

Gressman and Peck (arXiv:2006.03175) have used a stochastic agent-based model of a university to argue that to reduce spread all classes larger than 30 should be online only. We will use a much simpler models to show that the answer depends on how contagious the disease is.

Household model

This well-studied model is one step more realistic than homogeneously mixing. See e.g., Ball, Mollison, and Scalia-Tomba (1997) Ann. Prob.

- $p_H = P$(infect a given person within the house)
- $p_G =$ (G is for glabal) P(infect a given person outside the house)
- $G_H(z) =$ g.f. of the size of the epidemic within the house started by one person
- $G_\nu(z) = \exp(Np_G(z-1)$ g.f. of number of long distance infections caused by one person, distribution Poisson(Np_G)

Probability of a large epidemic is the positive solution of

$$1 - \zeta = G_H(G_\nu(\zeta))$$

Dorm with all single rooms

- $G_D^1(z) =$ g.f. of the size of the epidemic within the dorm started by one person
- $G_\nu(z) = \exp(Np_G(z-1))$ g.f. of number of long distance infections caused by one person, distribution Poisson(Np_G)

Spread of the epidemic in the dorm is an branching process with progeny Poisson(λ_1), $\lambda_1 = np_D$. Let G_b^1 be the g.f. of total progeny

Mean number infected in dorm: $(1-\lambda_1)^{-1}$

G_D^1 solves $G_D^1(z) = zG_b^1(G_D^1(z))$

Solve by iteration. Probability of a large epidemic solves

$$1-\zeta = G_D^1(G_\nu(\zeta)) \quad \text{where} \quad G_\nu(z) = \exp(Np_G(z-1)$$

Dorm with all double rooms

p_L = probability you infect your roommate

Epidemic in the dorm is an branching process with progeny

$$p_L \text{Poisson}(2 \cdot 2n_1 p_D) + (1 - p_L)\text{Poisson}(2n_1 p_D)$$

Let G_b^2 be the generating function of total progeny

Let $\lambda_2 = (1 + p_L)2n_1 p_D$. Mean number infected: $(1 - \lambda_2)^{-1}$

g.f. of size of epidemic within the dorm solves $G_D^2(z) = zG_b^2(G_D^2(z))$

Probability of a large epidemic solves

$$1 - \zeta = G_D^2(G_\nu(\zeta)) \quad \text{where} \quad G_\nu(z) = \exp(Np_G(z - 1)$$

Comparison of epidemic probabilities

$p_L = 0.7$, $n = 2n_1$ so same number of students per dorm.

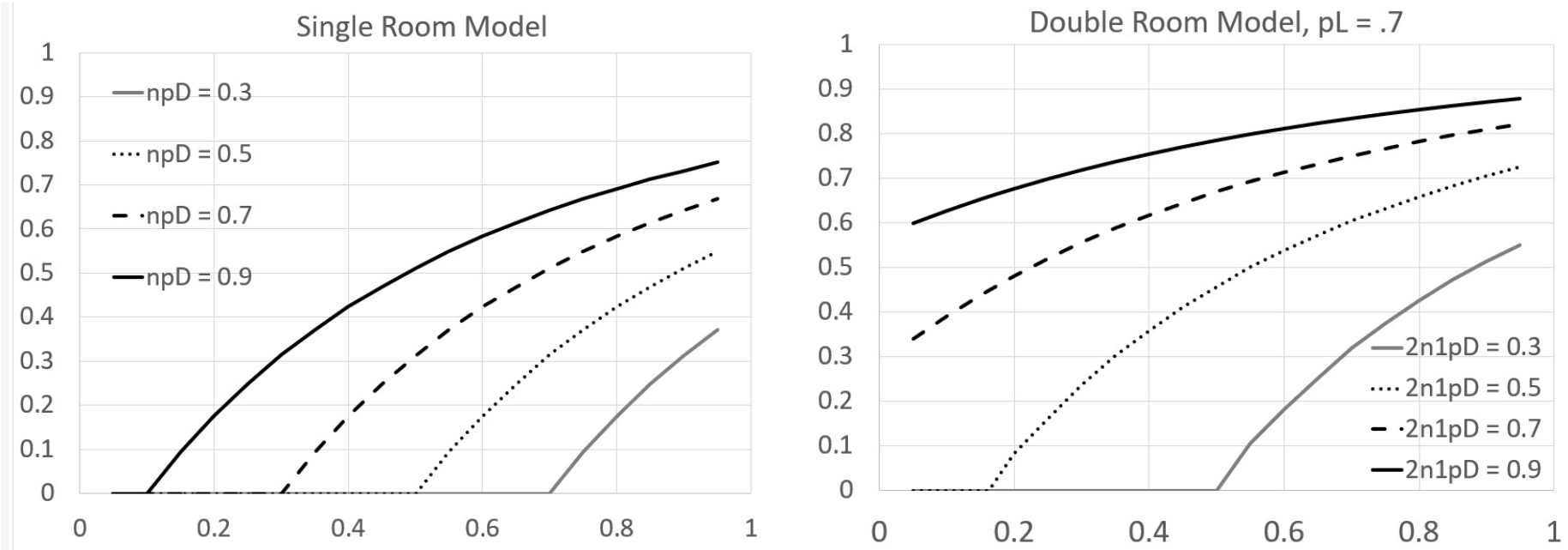


Figure: When $2n_1p_D = 0.7$ or 0.9 there is an epidemic even when $p_G = 0$. When $2n_1p_D = 0.3$ or 0.5 there is a positive threshold for p_G but it is 50% larger in the single room situation.

Class model

We consider some simplifieid models for a university where each student takes three classes. The situation can be described by a graph in which there are n students and m classes with sizes c_i $1 \leq i \leq m$, and there is an edge from each student to the three classes they are enrolled in. Each classroom is assumed to be homogeneously mixing.

An individual that is infected in class i at time t, adds his two alter egos (his presence in other classes) to the infected population at time $t + 1/2$ and then infections at time $t + 1$ are produced. None of the time $t + 1/2$ infecteds are present at time $t + 1$. Two time steps = one week.

Scenario 1. 100 classes of size 30. Each student takes 3 classes so this means 1000 students total. An individual infected in class 1 at time 0 will produce a Poisson($87p$) infecteds in class 1 at time 1.

Scenario 2

25 classes of size 60 ($1 \leq i \leq 25$) and 75 of size 20 ($26 \leq i \leq 100$), so again 1000 students. $m(i,j)$ expected number of infections in class j caused by one infected in in class i.

$$m_{i,j} = \begin{cases} (c_i - 1)p & j = i \\ 2 \cdot \frac{c_j}{3000 - c_i} \cdot (c_j - 1)p & j \neq i \end{cases}$$

Since maximum eigenvalue has $x_i = a$ for $1 \leq i \leq 25$ and $x_i = b$ for $26 \leq i \leq 100$, the eigenvalue problem for the 100×100 matrix can be reduced to one for a 2×2 matrix and we have

$$R_0 = 129.38p$$

which is 48% larger than all classes of size 30.

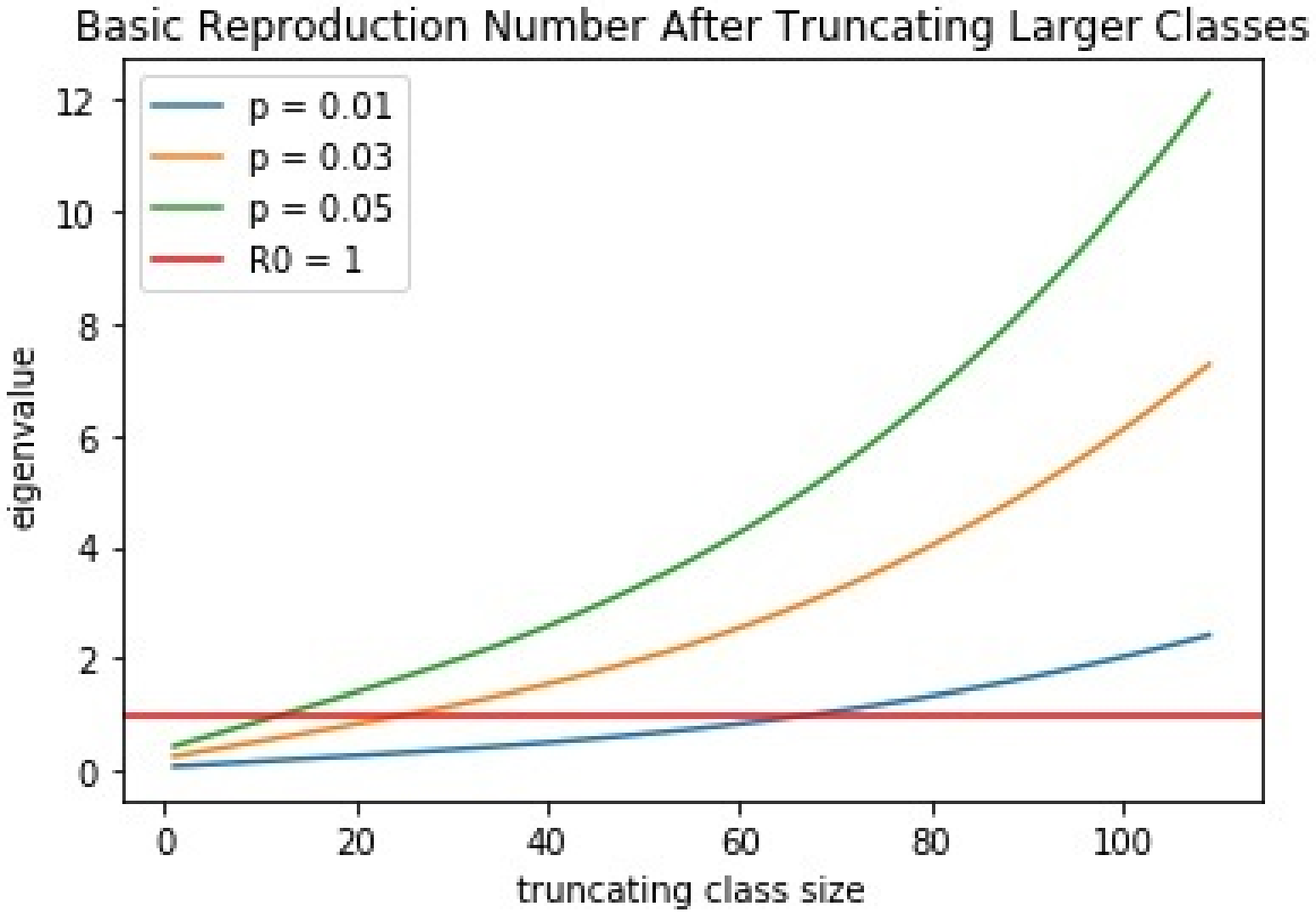
Scenario 3

Consider a university with 111 classes, one each of sizes ranging from 10, 11... 120. Total enrollment in classes is $10 + 11 + ... + 120 = 7215$, or 2405 students. $m(i, j)$ is the expected number of infections in class j caused by one infected student in i.

$$m_{i,j} = \begin{cases} (j+8)p & i = j \\ \frac{2(j+9)}{7215-(i+9)} \cdot (j+8)p & i \neq j \end{cases}$$

The maximum eigenvalue $R_0 = 251.5p$

If all classes of size $> k$ are moved online, $m(i, j) = 0$ for $j > k$.


Basic Reproduction Number After Truncating Larger Classes
p = 0.01
p = 0.03
p = 0.05
R0 = 1
eigenvalue
truncating class size
0
2
4
6
8
10
12
0
20
40
60
80
100


Thank You
Danke
Maake
Kiitos
Vinaka
Suksama
Suksama
köszönöm
спасибо
Dankscheen
Asante
ارك شكرا
Dhanyavadagalu
Manana
Dankon
Diolch i Chi
Grazie
Shukria
Biyan
Dank Je
Najis Tuke
Dziekuje
Chokrane
Khap
Tibi
Taiku
cảm ơn bạn
நன்றி
Juspaxar
Arigato
Gratias
Paldies
Mauruuru
Gracias
Obrigado
감사합니다
Bedankt
D'akujem
谢谢
Merci
Kop Khun
Tack
धन्यवाद
Děkuji
ありがとう
Ua Tsaug Rau Koj
Matur Nuwun
Rahmat
Grazas
Di Ou Mèsi
Go Raibh Maith Agat
Tingki
Eskerrik Asko
хвала
Nirringrazzjak
Hvala
ขอขอบคุณคุณ
Terima Kasih
Salamat