Graph theory and epidemics

Berkeley Math Circle

15 April, 2020

About me

- Mark Penney
- Perimeter Institute / UWaterloo
- Visitor at MSRI

About me

- From Newfoundland, Canada
- Its nice enough

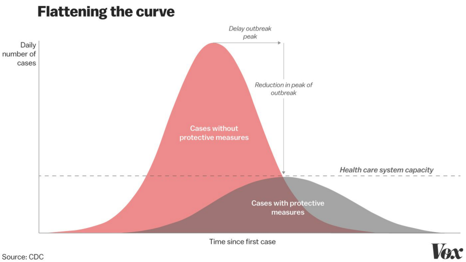
About me

- From Newfoundland, Canada
- Its nice enough
- ... Except in the winter

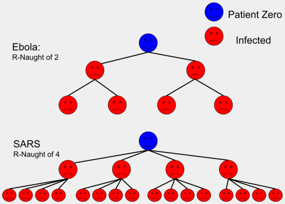
Mathematics of infectious disease

- I don't study infectious diseases
  - Topology (knots, surfaces,...) applied to theoretical physics (particle physics, quantum computing)
- Wanted to understand how we model epidemics
  - Personal reasons: Knowledge is power
  - Maybe contribute to COVID-19 research

COVID-19: Flattening the curve

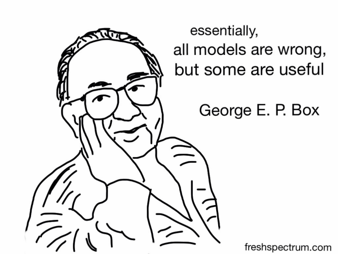


Fundamental number: Reproductive rate



- $R_0 =$ average number of new infections
- Phase transition at $R_0 = 1$:
  - $R_0 < 1$: disease dies off
  - $R_0 > 1$: epidemic outbreak
- Determines the trajectory of outbreak
- Flattening the curve $\Leftrightarrow$ reducing R_0

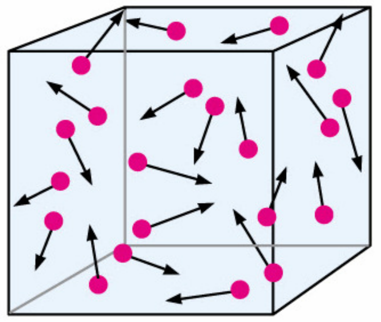
How do we determine R_0?



- Depends on assumptions of model
- Simplest model: $R_0 = \mu \cdot T$
  - $\mu =$ average number of contacts
  - $T =$ probability of spread
- To reduce R_0:
  - Reduce μ: social distancing
  - Reduce T: Handwashing, masks

What are the assumptions of this model?

Main assumption: 'Homogeneous mixing'. People interact with each other at random.
Does not take into account social structure!

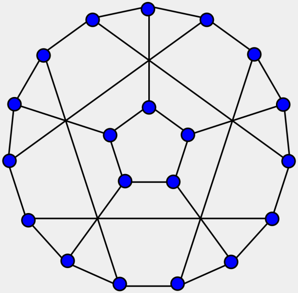


People are not like molecules of a gas!
More realistic: Understand the network or graph describing contacts

Contact networks

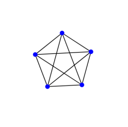
Recall: Graphs

A graph is a collection of vertices joined by edges

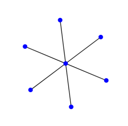


- Coronaviruses spread by close proximity interactions (CPIs)
- Build contact network for group of people:
  - Vertices: People in group
  - Edges: connect people having CPIs

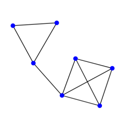
What sort of social structure leads to the following contact networks?



- 5 people all sharing CPIs
  - Household, group of friends, small sports team



- 6 people having CPI with 1 other
  - Doctor visiting patients, tutor

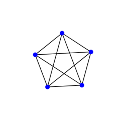


- 7 people in 2 clusters
  - The households of 2 friends

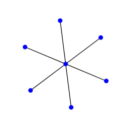
Degrees of contact networks

Recall: Degree

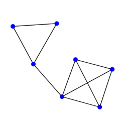
The degree of a vertex in the number of edges connected to it



- All vertices have degree 4
  - Every family member is in contact with the remaining 4



- Central vertex has degree 6, remainder degree 1
  - Doctor as 6 contacts, patients have 1

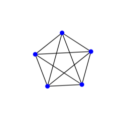


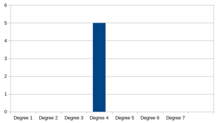
- Within each household, each family member is in contact with the remainder. Friends have higher degree

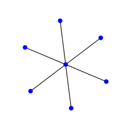
Degree distribution

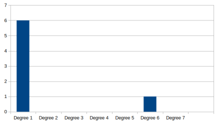
Takeaway message

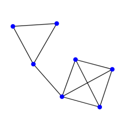
Social structures lead to varying degrees in contact network
Helpful to plot number of vertices having each degree

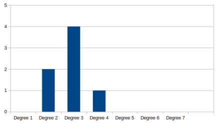












Degree distribution

Definition

The degree distribution of a graph is

$$d \mapsto p_d = \frac{\#\text{ vertices of degree } d}{\#\text{ vertices}}$$

Probability that randomly chosen vertex has degree d
Obtained from previous graphs by dividing by $\#$ vertices

Exercise

Compute the average degree of our 3 examples. Can you find a simple, general formula?

Average degree

Recall: Average of probability distribution

Given probability distribution p_d, the average is:

$$\mu = \langle d \rangle = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + \ldots = \sum_d d \cdot p_d$$

The average degrees are:

- ‘Household’ network: $\mu = 4$
- ‘Doctor’ network: $\mu = 1 \cdot \frac{6}{7} + 6 \cdot \frac{1}{7} = \frac{12}{7} = 1.7\ldots$
- ‘Friends’ network: $\mu = 2 \cdot \frac{2}{7} + 3 \cdot \frac{4}{7} + 4 \cdot \frac{1}{7} = \frac{20}{7} = 2.9\ldots$

Average degree formula

Proposition

The average degree of a graph is

$$\mu = \frac{2 \cdot \#\ \text{edges}}{\#\ \text{vertices}}$$

Proof

$$\mu = \frac{\sum_d d \cdot \ \#\ \text{vertices of degree d}}{\#\ \text{vertices}}$$

$$\sum_d d \cdot \ \#\ \text{vertices of degree k} = \sum_{\text{vertices } v} \#\ \text{edges meeting } v$$

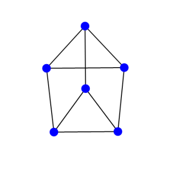
This sum counts each edge twice

Regular graph

A graph is called d-regular if each vertex has degree d.

Homogeneous mixing

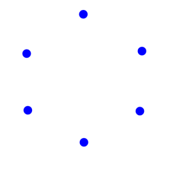
Assumption lead to $R_0 = \mu \cdot T$. Assumes that contact network is d-regular with $d \approx \mu$



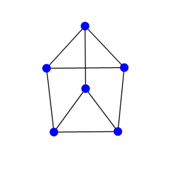
- Can you make a 3-regular graph with 5 vertices? 7?
- No! Must have $3 = \frac{2 \cdot \# \text{ edges}}{\# \text{ vertices}}$
- Must $d \cdot \#$ vertices is even!

Question

Can you make a 4-regular graph with 6 vertices?



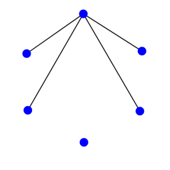
- Arrange vertices in circle



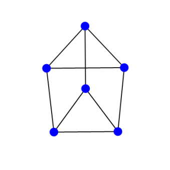
- Can you make a 3-regular graph with 5 vertices? 7?
- No! Must have $3 = \frac{2 \cdot \#\text{ edges}}{\#\text{ vertices}}$
- Must $d \cdot \#$ vertices is even!

Question

Can you make a 4-regular graph with 6 vertices?



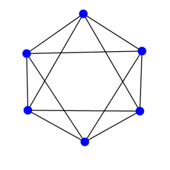
- Arrange vertices in circle
- Join each vertex with 2 nearest on each side



- Can you make a 3-regular graph with 5 vertices? 7?
- No! Must have $3 = \frac{2 \cdot \# \text{ edges}}{\# \text{ vertices}}$
- Must $d \cdot \#$ vertices is even!

Question

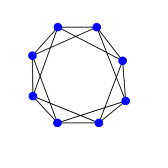
Can you make a 4-regular graph with 6 vertices?



- Arrange vertices in circle
- Join each vertex with 2 nearest on each side

Question

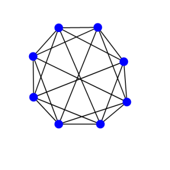
Can you make a 5-regular graph with 8 vertices?



- Arrange vertices in circle
- Join each vertex with 2 nearest on each side

Question

Can you make a 5-regular graph with 8 vertices?



- Arrange vertices in circle
- Join each vertex with 2 nearest on each side
- Join opposite vertices

Observation

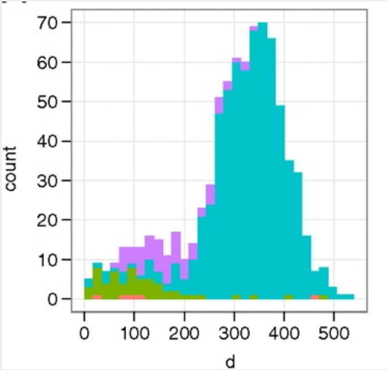
Can use these constructions to make d-regular graph with V vertices, assuming $d \cdot V$ is even and V is large enough

Realistic degree distributions

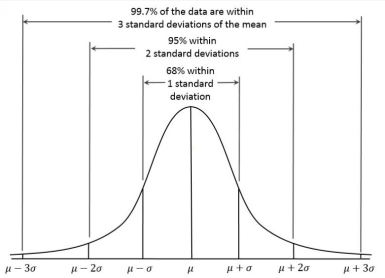
- Simplest model assumes regular contact network
- Saw that even simple contact networks are far from regular

Question

What are the degree distributions of realistic contact networks?



- Number of daily CPIs in US high school
- Blue = students, pink = teacher, green = staff
- What other statistical ideas might be useful here?



- Standard deviation σ measure of 'width' of distribution

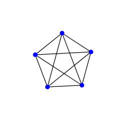
$$\sigma^2 = (0-\mu)^2 p_0 + (1-\mu)^2 p_1 + (2-\mu)^2 p_2 + \cdots = \sum_d (d-\mu)^2 p_d$$

- σ^2 = Average of (distance from μ)²

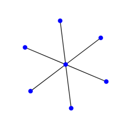
Computing standard deviation

Formula

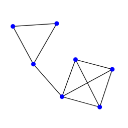
$\sigma^2 = (0-\mu)^2 p_0 + (1-\mu)^2 p_1 + (2-\mu)^2 p_2 + \cdots = \sum_d (d-\mu)^2 p_d$



- $\mu = 4$
  - $\sigma^2 = (4-4)^2 \frac{5}{5} = 0$



- $\mu = \frac{12}{7}$
  - $\sigma^2 = \left(1 - \frac{12}{7}\right)^2 \frac{6}{7} + \left(6 - \frac{12}{7}\right)^2 \frac{1}{7} = \frac{150}{49} = 3.06\ldots$



- $\mu = \frac{20}{7}$
  - $\sigma^2 = \left(2 - \frac{20}{7}\right)^2 \frac{2}{7} + \left(3 - \frac{20}{7}\right)^2 \frac{4}{7} + \left(4 - \frac{20}{7}\right)^2 \frac{1}{7} = \frac{20}{49} = 0.408\ldots$

Reproductive rate v2

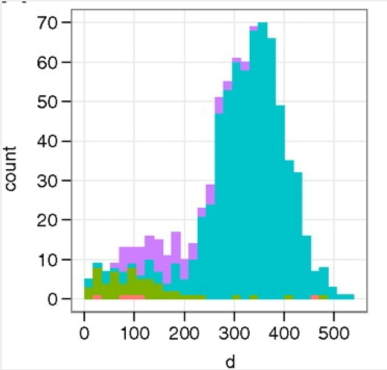
Better model for R_0

Correction factor accounting for width

$$R_0 = T\mu\left(1 + \frac{\sigma^2}{\mu^2}\right)$$

- 'Household' network: $\frac{\sigma^2}{\mu^2} = \frac{0}{4^2} = 0$
  - For regular graphs $\sigma = 0$ so simple = better model
- 'Doctor' network: $\frac{\sigma^2}{\mu^2} = \left(\frac{150}{49}\right)^2\left(\frac{7}{12}\right)^2 = \frac{625}{196} = 3.19\ldots$
  - Better model gives R_0 4 times larger!
- 'Friends' network: $\frac{\sigma^2}{\mu^2} = \left(\frac{20}{49}\right)^2\left(\frac{7}{20}\right)^2 = \frac{1}{49} = 0.02\ldots$
  - Better model is only slightly larger than simple model

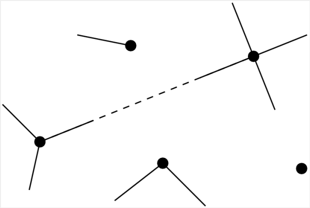
Reproductive rate v2



- $\frac{\sigma^2}{\mu^2} = 0.118$
- Better model for R_0 is almost 12% higher

Takeaway messages

- Since $\frac{\sigma^2}{\mu^2} \geq 0$, simple $R_0 \leq$ better R_0
- Wider degree distribution $\Rightarrow$ faster spread of disease



How to make a random graph

1. Create vertices with chosen degrees
2. Connect the edges at random

Theorem

On average, a contact has $\mu\left(1+\frac{\sigma^2}{\mu^2}\right)$ contacts
$\Rightarrow$ since $\geq \mu$, your friends have more friends than you!

Proof

Question:

What is the probability that an edge meets a vertex of degree d?

Answer:

$$\frac{\#\text{edges meeting degree } d}{2E} = \frac{d \cdot \#\text{vertices of degree } d}{2E}$$

$$= \left(\frac{d}{2E/V}\right)\left(\frac{\#\text{vertices of degree } d}{V}\right) = \frac{d \cdot p_d}{\mu}$$

The average degree of a vertex meeting a random edge is

$$\sum_d d \cdot \left(\frac{d \cdot p_d}{\mu}\right) = \frac{1}{\mu}\sum_d d^2 \cdot p_d$$

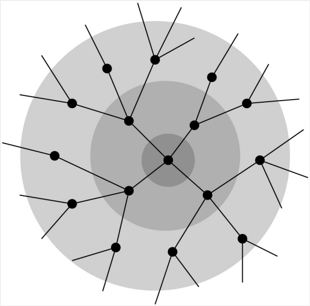
Proof

Observe that,

$$\begin{aligned}\sigma^2 &= \sum_d (d-\mu)^2 p_d = \sum_d d^2 \cdot p_d - 2\mu \sum_d d \cdot p_d + \mu^2 \sum_d p_d \\ &= \sum_d d^2 \cdot p_d - \mu^2\end{aligned}$$

Therefore,

$$\mu\left(1 + \frac{\sigma^2}{\mu^2}\right) = \frac{\mu^2 + \sigma^2}{\mu} = \frac{1}{\mu}\sum_d d^2 \cdot p_d$$



- Reproductive rate R_0 determines spread of disease
- Good model: $R_0 = T\left[\mu\left(1+\frac{\sigma^2}{\mu^2}\right)\right]$
  - Second factor is average degree of a neighbor
- To flatten the curve need to also reduce 'width' of contact network
  - e.g., 'super-spreaders', large gatherings, schools

Thank you very much for your attention!