

Introduction to SEIR modelling

What is a SEIR model?

- **S** (Susceptible): People that are susceptible to the disease, or people who might get the disease
- **E** (Exposed): People who were infected by the virus (currently has the virus), and can transmit the virus to other susceptible, but who has not developed symptoms (carriers)
- **I** (Infected): Similar to exposed, infected are who has the virus, transmissible, and might have or not have symptom.
 - I_p (preclinical): Symptomatic patients who has not yet developed the symptom
 - I_s (subclinical): Asymptomatic patients
 - I_c (clinical): Symptomatic patients who has the symptom
- **R** (removed): People who has recovered/demised

Geographical Stratification

- Ontario has
- 26 Public Health Units
- 50 administrative districts
- 528 counties

We have following data of each county:

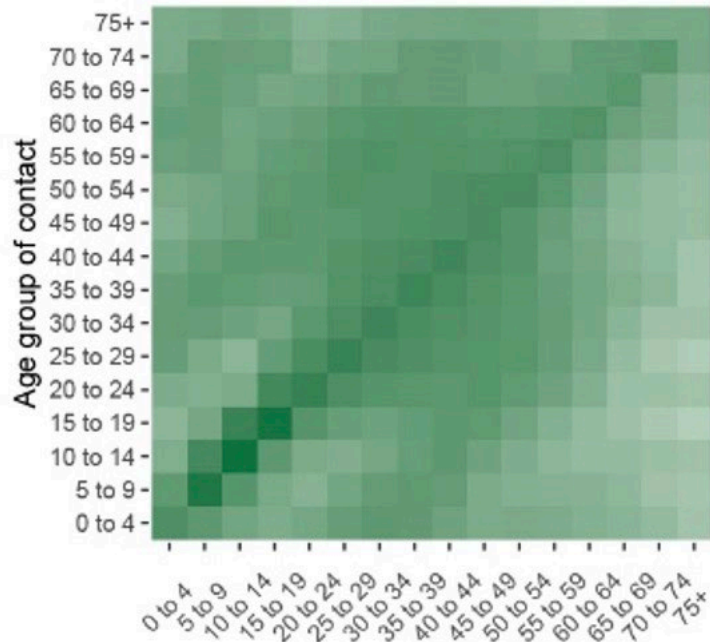
- Population
- Commuting matrix

Age stratification

- We stratified the Ontario population into 16 age-bands, as
 {0 to 4, 5 to 9, ... , 70 to 74, 75 +}
- To correspond with the contact matrix, case-fatality-rate, susceptibility, and vaccination level

Contact matrix

- A 16×16 matrix.
- Each row represents a five-year age band of contacters
- Each column represents a five-year age band of contactees



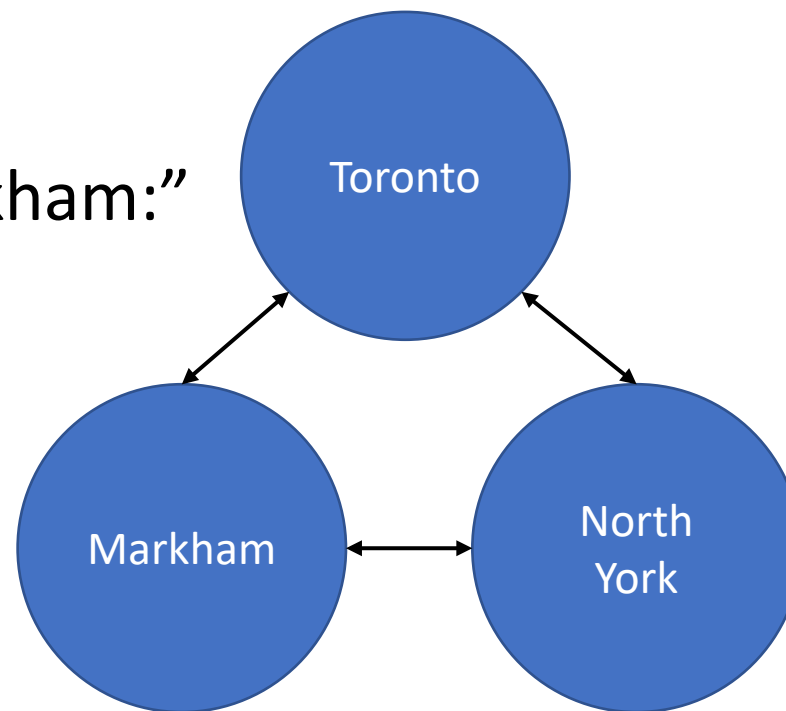
Commuting matrix

- A 528×528 matrix.
- Each row represents a county of residence
- Each column represents a county of work

A simpler view

$$\begin{pmatrix} 10000 & 6000 & 4000 \\ 2000 & 8000 & 3000 \\ 1500 & 3000 & 5000 \end{pmatrix}$$

- Let's consider the case where there are only three counties.
- Here, we know
- No. people live & work in Toronto: 10000
- No. people live in Toronto & work in Markham:"
2000
Etc.



Modelling infection – naïve

Suppose we have:

- Total number of transmissible individuals on a specific day: $T = I + E$
- Average infection rate: β
- Average number of contacts in a day: c

Then, the number of newly infected individuals is:

$$I_n = c\beta T$$

Naïve model – cont.

- The distribution function of a removal:
- We use ReLU after a specific date, according to Lancet paper, to represent the probability distribution of a specific patient that will recover/decease after n days of initial infection.

How does our model look like?

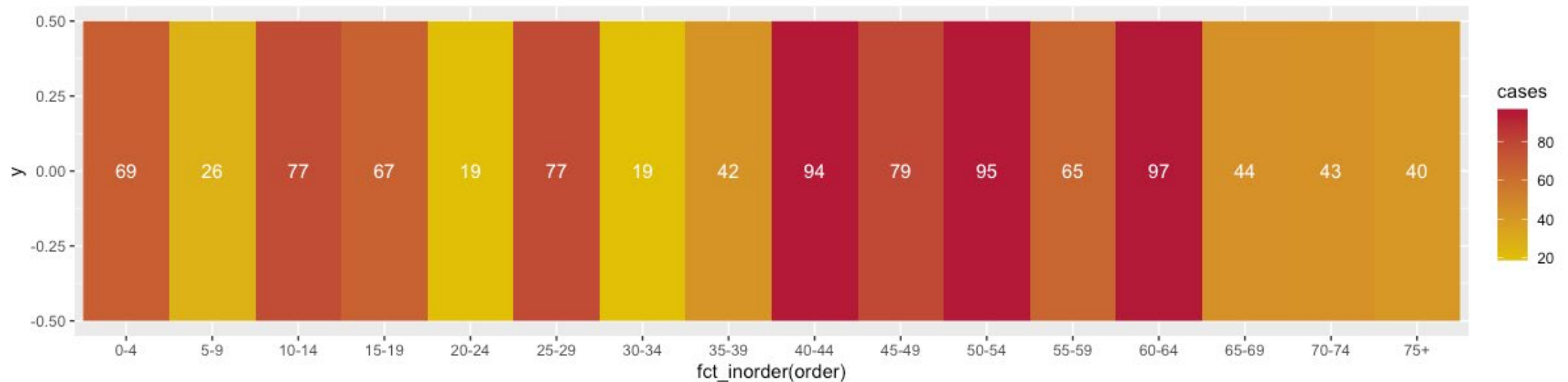
- Our model uses a double-valued vector $\in \mathbb{R}^{16}$ to store the basic epidemic data of a day, where each entry represents a corresponding age-band.
- For example, the number of transmissive individuals in a day looks like:

$$v_T = [152, 543, \dots, 240]$$

- Here, there are 152 new cases for the age band '0 to 4'

How does our model look like?

- This is how the number of transmissible individuals might look like.
- Color indicates the number of cases.



How does our model look like? – cont.

- Intuitively, the number of newly exposed/infected individuals is essentially another vector.
- According to our naïve model, the vector of newly exposed individuals looks like:

$$v_{\Delta E} = \beta[152, 542, \dots, 240]$$

- For example, we set $\beta = 2$, we have:

$$v_{\Delta E} = [304, 1084, \dots, 480]$$



How does our model look like? – cont.

- However, not all individuals have the same susceptibility (i.e, some individuals are more “resistant”/immune against the transmission), and the number of contacts between each age band in a day are also different.
- Thus, we introduce the idea of effective contact ratio (these are also vectors!), as

$$\mathbb{R}_+^{16} \ni \beta = \eta \circ \rho$$

- Where η is the (age-specific) susceptibility, ρ is the (age-specific) immunity level, β is the (age-specific) infection rate

How does our model look like? – cont.

- Now, we can model the number of newly exposed individuals using such quantities!

$$\mathbb{R}_+^{16} \ni dE = \overbrace{\beta \circ \mathcal{M}(C)}^{\text{Effective matrix}} T$$

- Here, T is the (age-specific) number of transmissible individual, $M(C)$ is the contact matrix, $\beta \circ M(C)$ is the effective contact matrix (effective matrix).

How does our model look like? – cont.

- Similarly, the number of newly infected individual is

$$dI = dI_c + dI_s \quad \text{where} \quad \begin{cases} dI_c = \mu \left(\overbrace{(1, \dots, 1)}^{\in \mathbb{R}^{16}} - a \right) \circ E - dR_c \\ dI_s = \mu a \circ E - dR_s \end{cases}$$

- Where I is the (age-specific) number of infected individuals I_s is the number of subclinical (asymptomatic) I_c is the number of clinical (symptomatic), μ is a constant scalar-valued probability function of infection rate, a is the age-specific asymptomatic ratio.

How does our model look like? – cont.

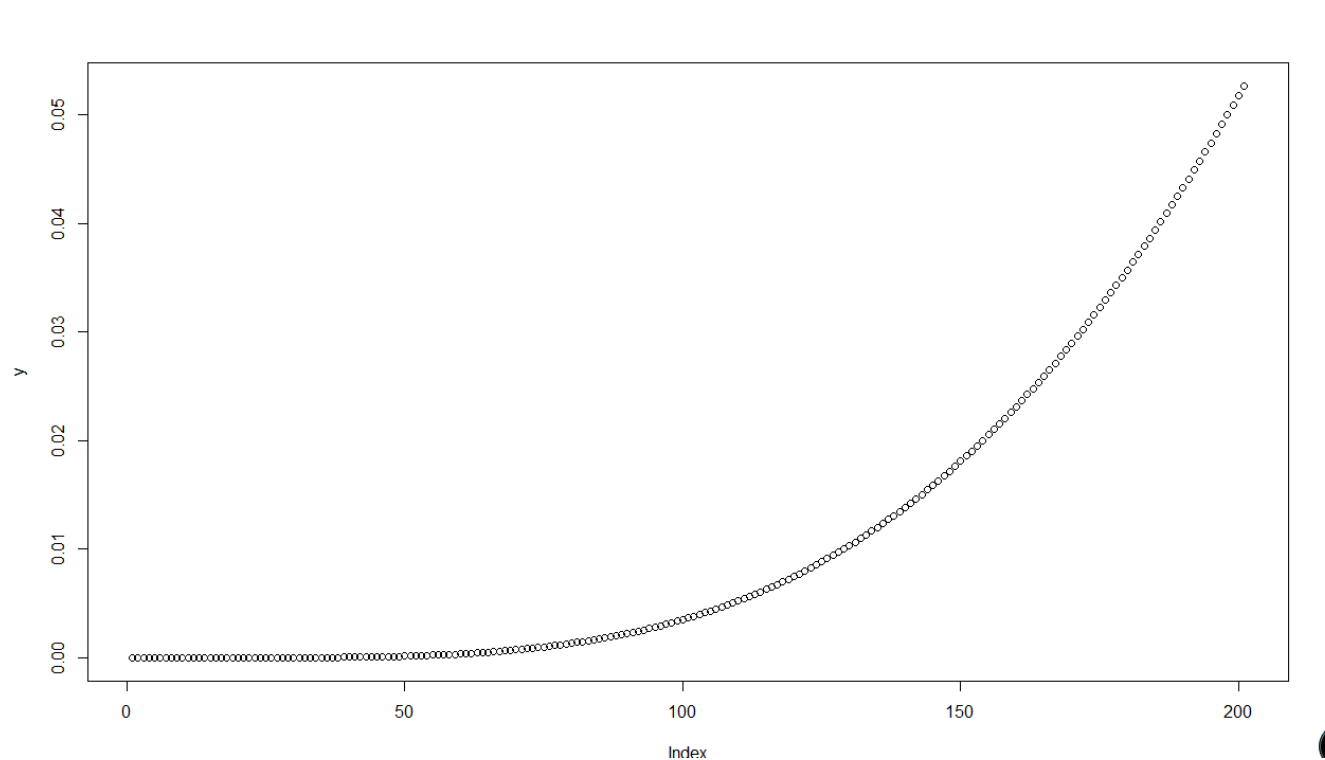
- Similarly, the number of recovered/demised individuals is:

$$dR = dD + dr \quad \text{where} \quad \begin{cases} dD = \gamma CFR \circ I \\ dr = \gamma(1 - CFR) \circ I \end{cases}$$

- Where R is the (age-specific) number of removed individuals. D is the (age-specific) number of demised individuals, r is the (age-specific) number of recovered individuals, CFR is the (age-specific) case-fatality rate.

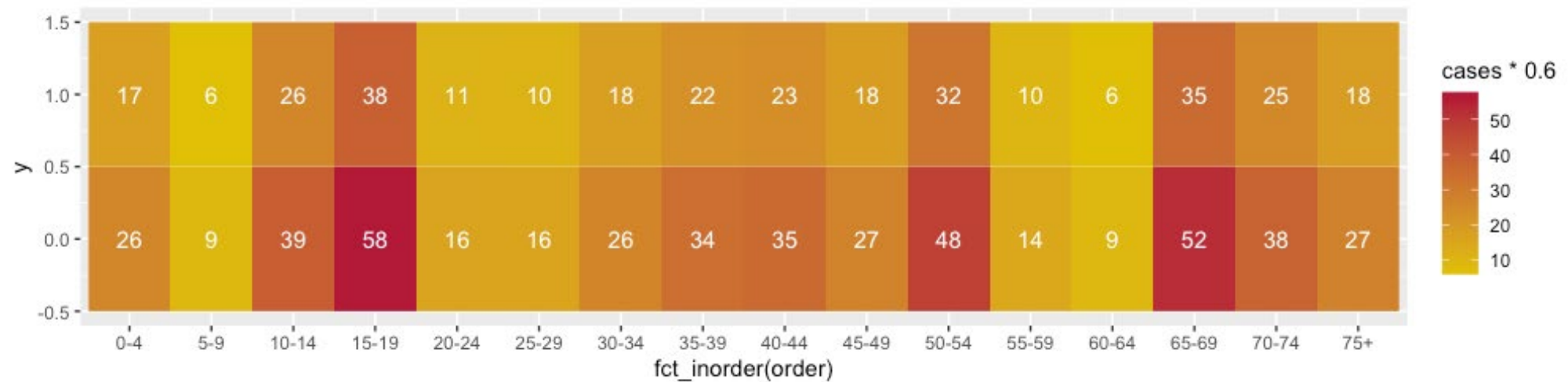
infected -> recovered & exposed -> infected

- A gamma distribution! [REF], it's something like this



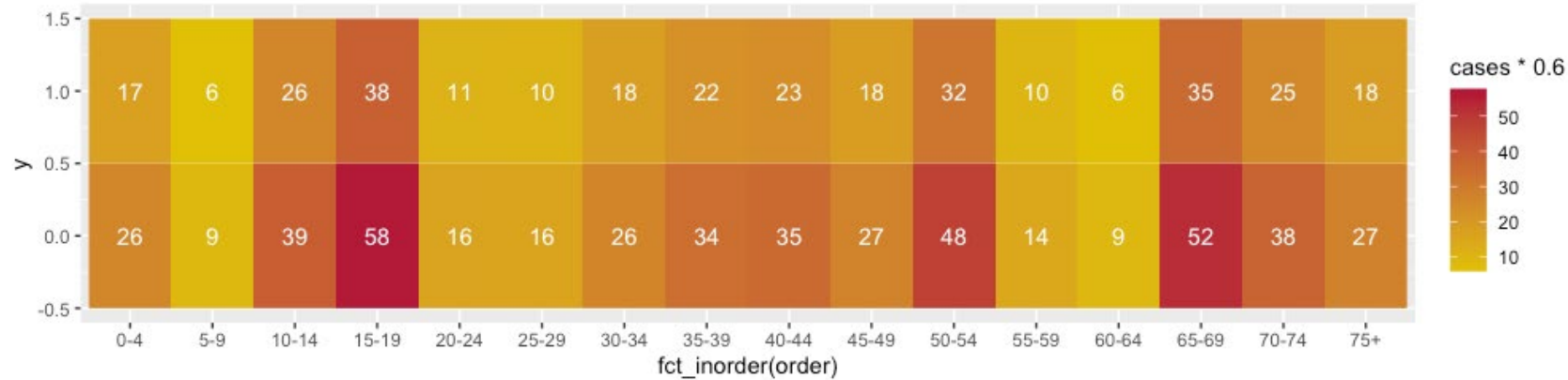
Commutation

- We modelled the number of commute let's consider the following array.



- Assume the first row is the number of local-worker, and the second array is. The number of commuting workers.

Commutation – cont.



- Now, we reshape the second row into a $528 * 16$ np.array, where the sum of each row is corresponding to the entry of the, for example:

$$\text{Age and destination specific transmissible individuals} = \overbrace{\begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{527,1} & a_{528,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{1,16} & a_{2,16} & \cdots & a_{527,16} & a_{528,16} \end{pmatrix}}^{528 \times 16} \in \mathbb{R}_+^{528 \times 16}$$

where

$$j \in \{1, \dots, 16\} \quad \sum_{i=1}^{528} a_{i,j} = v_i \quad \text{i.e.} \quad \sum_{i=1}^{528} a_{i,2} = 9$$

Commutation – cont.

- Intuitively, the in-coming commuters is also a $528 * 16$ matrix. Such matrix will be reshaped (merged) into a 16-dimensional vector and will be used to calculate the number of new cases.
- The number of new cases will then be re-distributed according to the weight of each entry (detailed description in supplementary material).
- Essentially, just
reshape -> travel -> reshape -> transmission -> return -> reshape