TBSI 清华-伯克利深圳学院 Tsinghua-Berkeley Shenzhen Institute

Spread of Covid-19: Estimation and Simulation

Mathematical Statistics & Application in R Final Project

Date：2020 / 04 / 23

Table of Contents

Deterministic Model: SIR

TBSI 清华-伯克利深圳学院
Tsinghua-Berkeley Shenzhen Institute

$$S \xrightarrow{\beta} I \xrightarrow{\gamma} R$$

$$\frac{ds}{dt} = -\beta si$$

$$\frac{di}{dt} = \beta si - \gamma i$$

$$\frac{dr}{dt} = \gamma i$$

s : the ratio of susceptible people

i : the ratio of infected people

r : the ratio of recovered people

β : the rate of transmission

γ : the rate of recovery

$$s = s_0 e^{-\frac{\beta}{\gamma} r}$$

$$\frac{dr}{dt} = \gamma(1 - r - s_0 e^{-\frac{\beta}{\gamma} r})$$

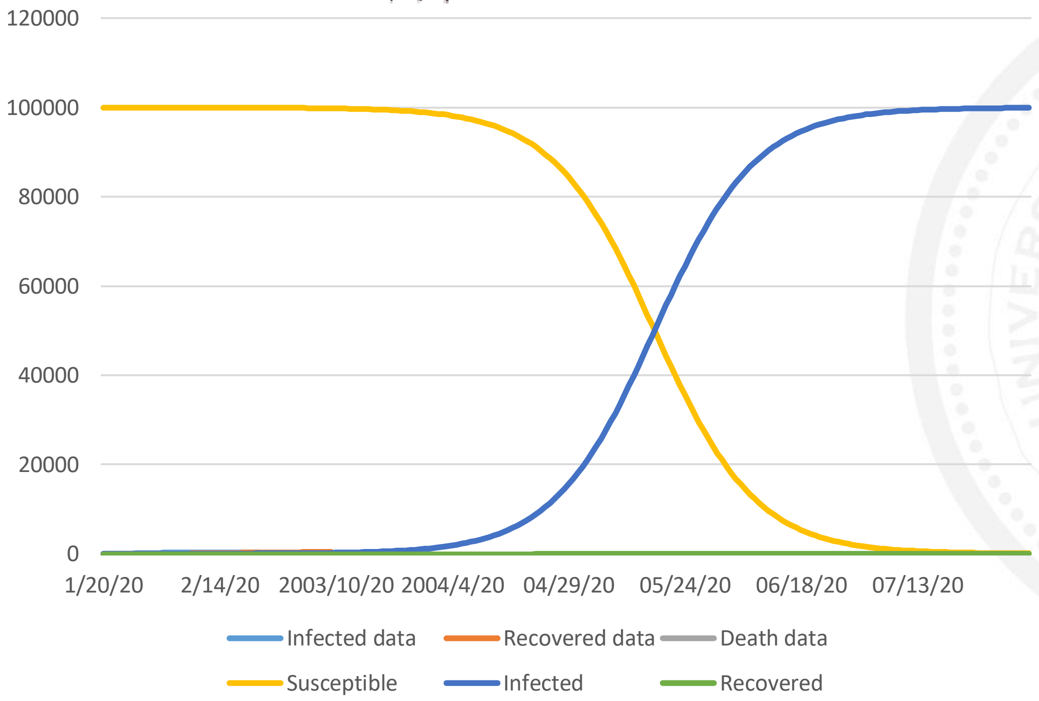
$$t = \frac{1}{\gamma} \int_0^r \frac{du}{1 - u - s_0 e^{-\frac{\beta}{\gamma} u}}$$

How to fit the real data into the model?

1. *Solve the differential equations at every time step;*
2. *Minimize the weighted MSE for s, i, and r*

$$e.g \; min \; \alpha_1 (\sum_{i=1}^{n} (\hat{s_i} - s_i)^2 + (1 - \alpha_1)(\sum_{i=1}^{n} (\hat{r_i} - r_i), set \; \alpha_1 = 0.1$$

3. *Get the estimated β, γ*



β=0.00000091, γ=0.00000024, r_0: 3.81126064.

Deterministic Model for Disease Spreading

$$S \xrightarrow{\beta} E \xrightarrow{\alpha} I \xrightarrow{\gamma} R$$

$$\frac{ds}{dt} = -\beta s i$$
$$\frac{de}{dt} = \beta s i - \alpha e$$
$$\frac{di}{dt} = \alpha e - \gamma i$$
$$\frac{dr}{dt} = \gamma i$$

$s:$ *the ratio of susceptible people*
$e:$ *the ratio of exposed people (not infected but infective)*
$i:$ *the ratio of infected people*
$r:$ *the ratio of recovered people*
$\beta:$ *the rate of transmission*
$\alpha:$ *the rate of progression*
$\gamma:$ *the rate of recovery*

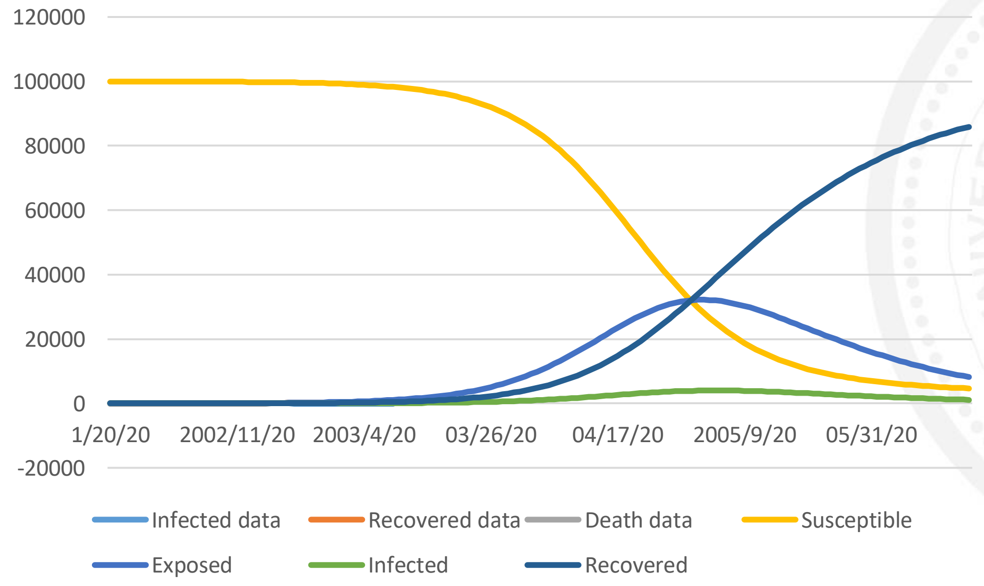
$$s = s_0 e^{-\frac{\beta}{\gamma} r}$$
$$\frac{dr}{dt} = \gamma(1 - e - r - s_0 e^{-\frac{\beta}{\gamma} r})$$
$$t = \frac{1}{\gamma} \int_0^r \frac{du}{1 - e - u - s_0 e^{-\frac{\beta}{\gamma} u}}$$

How to fit the real data into the model?
1. *Solve the differential equations at every time step;*
2. *Minimize the weighted MSE for s, i, and r*

$$e.g\ min\ \alpha_1 (\sum_{i=1}^{n} (\hat{s}_i - s_i)^2 + \alpha_2 (\sum_{i=1}^{n} (\hat{e}_i - e_i) + (1 - \alpha_1 - \alpha_2)(\sum_{i=1}^{n} (\hat{r}_i - r_i)$$

$set\ \alpha_1 = 0.1, \alpha_2 = 0.1$

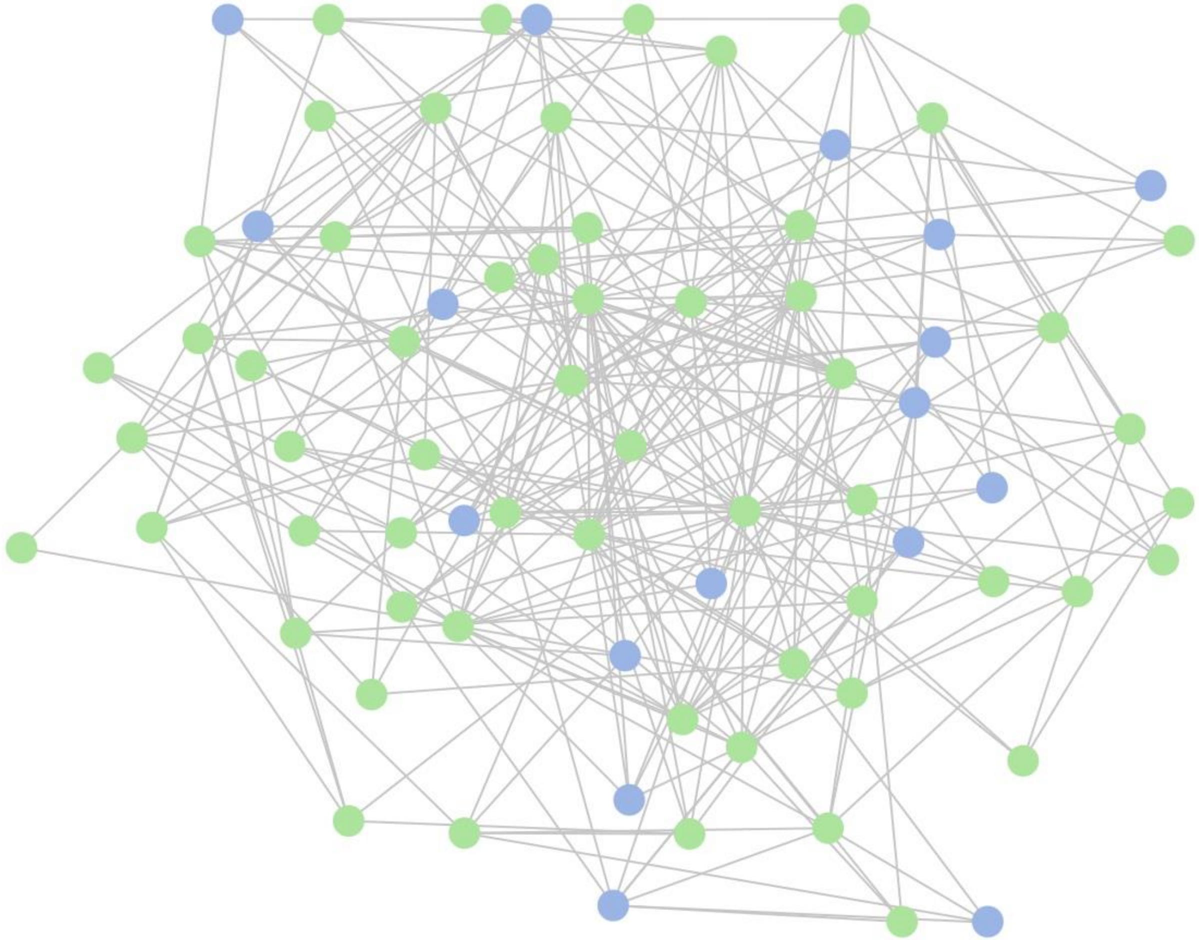
3. *Get the estimated* β, α, γ



β=0.00001423, α=0.05080899, γ=0.40000000, r_0: 0.00003558.

MGF: Scale-free Network Case

Network assumption:
1) Neighbor for every selected node is independent;
2) The network is huge.



Features:
1) is preferential attachment;
2) follows a power-law distribution or clustering coefficient distribution;
3) has a heavy tail..

Scale-free Network Case: Estimation

TBSI 清华-伯克利深圳学院 Tsinghua-Berkeley Shenzhen Institute

$$Rewrite\ power-law\ distribution$$

$$p_k = \frac{\nu-1}{k_{min}}\left(\frac{k}{k_{min}}\right)^{-\nu} where\ c = \frac{\nu-1}{k_{min}^{-\nu+1}}$$

$$z_m = \int_{x_{min}}^{\infty} x^m \frac{\nu-1}{x_{min}}\left(\frac{x}{x_{min}}\right)^{-\nu} = \frac{\nu-1}{\nu-1-m}x_{min}^m$$

$$For\ method\ of\ moments$$

$$E[x] = \frac{\nu-1}{\nu-2}x_{min}$$

$$E[x^2] = \frac{\nu-1}{\nu-3}x_{min}^2$$

$$\therefore \hat{x}_{min} = \frac{\sum x_i^2(\nu-3)}{\sum x_i(\nu-2)}$$

$$\therefore \hat{v} = \frac{-2(\sum x_i)^2 + 2n\sum x_i^2 \pm \sqrt{-n(\sum x_i)^2\sum x_i^2 + n^2(\sum x_i^2)^2}}{-(\sum x_i)^2 + n\sum x_i^2}$$

$$For\ maximizing\ log-likelihood\ estimator$$

$$With\ continus\ assumption$$

$$p(x) = \prod_{i=1}^{n}\frac{\nu-1}{x_{min}}\left(\frac{x}{x_{min}}\right)^{-\nu}$$

$$logp(x) = nlog(\nu-1) - nlogx_{min} - \nu\sum_{i=1}^{\infty} log\frac{x_i}{x_{min}}$$

$$\frac{\partial logp(x)}{\partial x_{min}} = \frac{-n}{x_{min}} + \frac{\nu}{x_{min}}\sum_{i=1}^{\infty}\frac{1}{x_i} = 0$$

$$\frac{\partial logp(x)}{\partial \nu} = \frac{n}{\nu-1} - \sum_{i=1}^{\infty} log\frac{x_i}{x_{min}}$$

$$\nu = 1 + n\left[\sum_{i=1}^{n} ln\frac{x_i}{x_{min}}\right]^{-1}$$

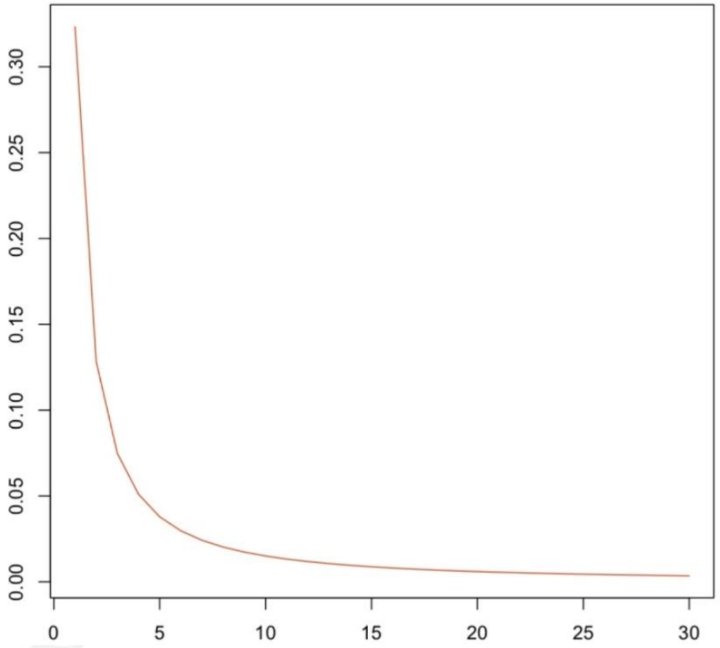
$$With\ discrete\ assumtion$$

$$\nu = 1 + n\left[\sum_{i=1}^{n} ln\frac{x_i}{x_{min}-\frac{1}{2}}\right]^{-1}$$

$$Let\ us\ assume\ x_{min} = 1, \hat{\nu} = 1 + \frac{n}{ln(\sum x_i)}$$

Scale-free Network Case

$$P_k = C \cdot k^{-v} \quad \rightarrow \quad \hat{v}_{MLE} = 1.33241$$



$$G^{(neighbor)} = G_0(G_1(\cdots G_1(x)\cdots)) \quad \ldots where \quad G_1(x) = \frac{G_0'(x)}{G_0'(1)}$$

$$Z_m = \frac{d}{dx} G^{(neighbor)}\Big|_{x=1} = G_0'(1)\left[G_1'(1)\right]^{m-1} \quad \rightarrow \quad Z_2 = 1321.40592$$

$$T_c = \frac{1}{G_1'(1)} = \frac{G_0'(1)}{G_0''(1)} \qquad = 0.0123764$$

According to the public data, we know that T = 0.22.

By substituting T into the following formula, we get:

$$E[s] = 1 + \frac{T(\xi(\nu-1))^2}{\xi(\nu)[(T+1)\xi(\nu-1) - T\xi(\nu-2)]}$$

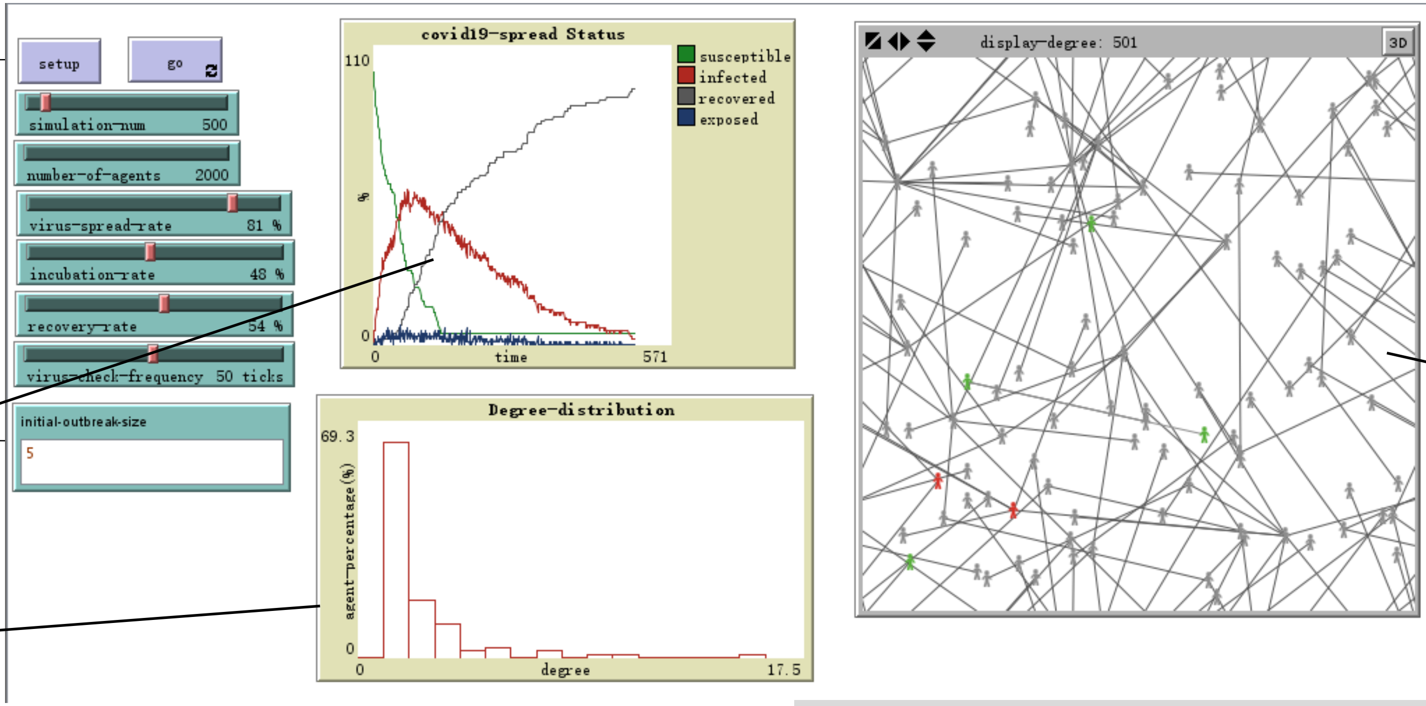
$$= 0.7855$$

Scale-free Network Case: Simulation

Parameters input

Plot the result

Degree distribution



Simulation world

Simulation for parameter sensitivity test :

1.*Set ranges for parameters;*

2.*Calculate the minimum, maximum and mean results;*

3.*Draw the experimental distribution of the outputs.*

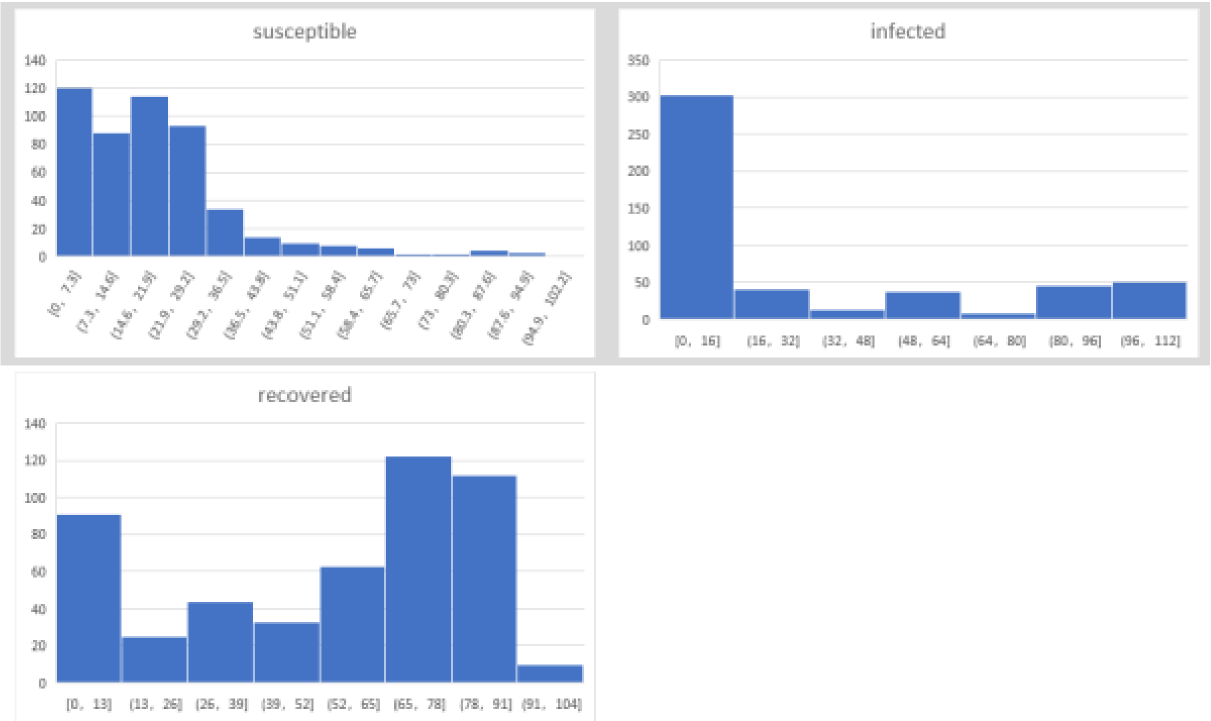
PARAMETERS AND INITIALIZATION	RANGES
initial-outbreak-size	[1 1 5]
recovery-rate	[1 10 100]
virus-spread-rate	[1 10 100]
run the model multiple times	500
agents numbger	2000
tick-check frequency	50

Scale-free Network Case: Simulation

AVERAGE(%)	MIN	MAX	MEAN
susceptible	9.9720	97.0000	19.7972
infected	1.5560	78.2940	44.6495
recovered	0.0000	53.4540	35.5533
virus	1.5560	78.2940	44.6495

- The mean ratio of susceptible people, infected people and recovered people are around 19.80%, 44.65%, 35.55%;
- The experimental distribution of the ratio of susceptible people and infected people were similar to the power-law distribution while ratio of recovered people is not.

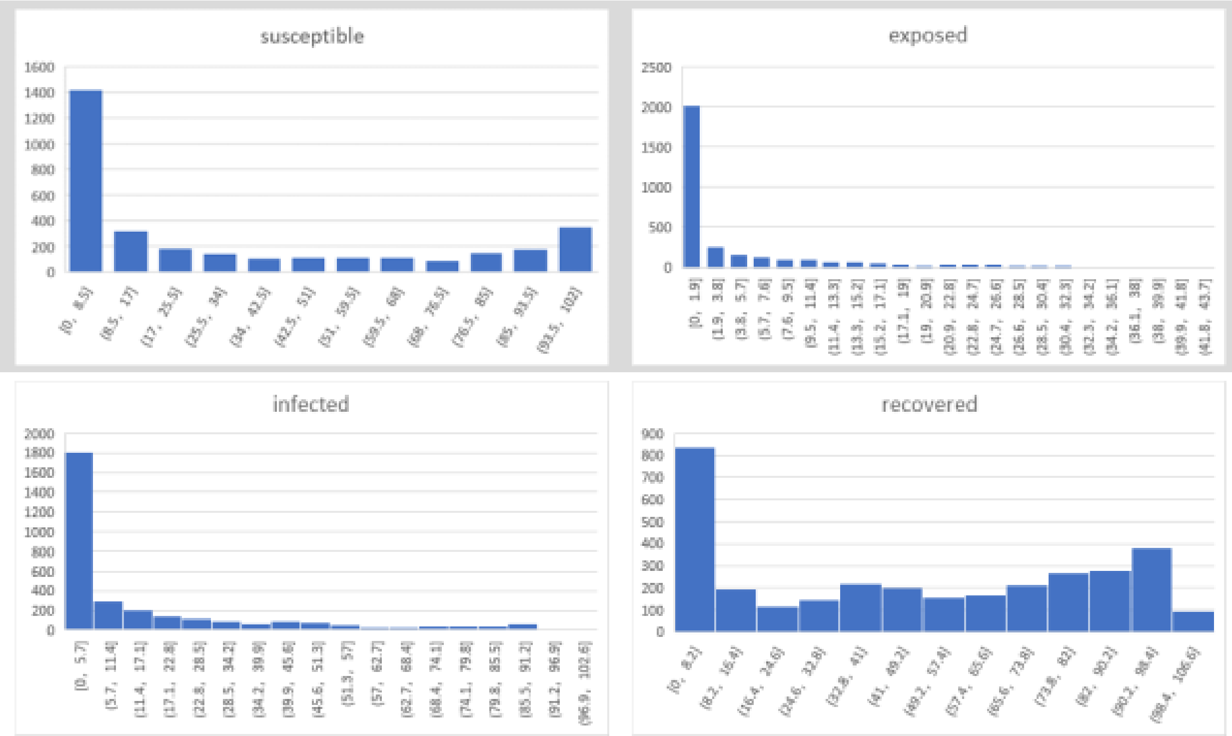
The experimental distribution of SIR



Scale-free Network Case: Simulation

AVERAGE(%)	MIN	MAX	MEAN
susceptible	32.3918	97.8309	43.9674
exposed	0.0000	14.3373	5.7009
infected	1.2026	41.5529	21.8660
recovered	0.0000	46.8978	28.4640
virus	1.3938	48.6874	27.5569

The experimental distribution of SEIR



- The mean ratio of exposed people was 5.7% which was lower than other outputs.
- This corresponds to our intuitive in the case of Covid-19.

Conclusion

MGF → E(s) = 0.7855

SIR → max(infected %) = 0.78294

SEIR → max(infected %) = 0.41553

p.s. An additional exposure state has been included in the transition state without any given data. The estimated value will be affected during solving additional one differential equations.