

# R とパンデミックの数理モデル新型コロナウイルス (**COVID-19**) 研究を例に

---

ill-identified

2020-02-29

- これまでのあらすじ:
  - 毎回間に合わなかったり遅刻したりしたらついにイベントの方から遠ざかってった
- 今回やる話:
  - 新型コロナウイルス (COVID-19) の時事ネタに乗っかって感染症の流行モデルの話.
- 当初やる予定だった話は Tokyo.R が再開したらやる予定
- SIR, SEIR という基本的な感染症流行モデルのシミュレーションの話
- 上記のさらに発展的な話
- 今回はちゃんと **R** の話もする

# イントロダクション

## 読む前から分かっていること

1. 中国の武漢中心に新型コロナウイルス (COVID-19) が流行
2. SIR/SEIR モデルは感染症の流行モデルで最も基本的
3. SIR モデルや SEIR モデルは計算が簡単
4. 今回はモデルから分かることを掘り下げる
5. 一度感染した人が再感染する場合を考慮したモデルがある

## これを読んで分かること

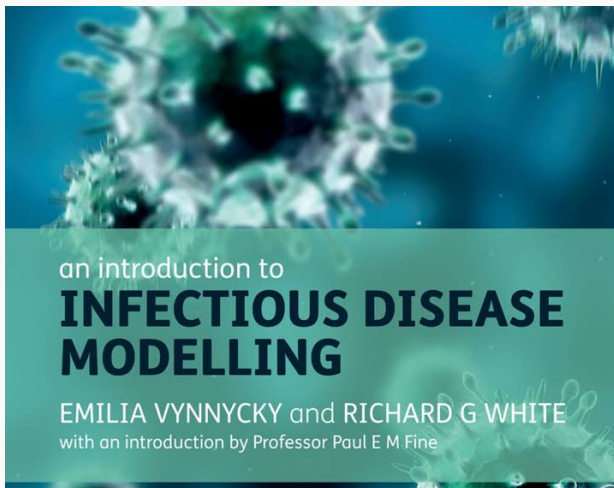
1. SIR/SEIR モデルは再感染を考慮しない. パンデミックが起こるのは1度きり
2. 感染予防はパンデミックの規模を軽減したり発生を遅らせたり
3. 再感染を考慮するモデルはもっと予想しづらい結果
4. SEIR モデルを拡張して COVID-19 の流行を分析した研究がもうある
5. 付録として具体的な数値計算の方法のヒントを書いた
6. お前, 得したな

# なぜシミュレーションか？

[1] > 統計モデルに対動学方程式の数理モデルは注目されないが、伝染病動学のより詳しいメカニズムを得ることができる

- 分析には必ず「ある仮定のもとで」という前置き
- 現実では人間が対策を立てることで結果を変えられる.
- 想定してない行動を取ることで結果は良い方にも悪い方にも変化
- 将来の予言ではなく「もしこの行動を取ったら？」の想像が重要

- 詳しくないので Vynnycky and White [2] を読む.
- 他にちょうどいい「入門」があるのかは知らない



# SIR モデル

- **SIR** モデルは人口全体を 3 つに分割
  - 感受性人口 (感染しうる人の数,  $S(t)$ )
  - 感染人口 ( $I(t)$ )
  - 回復人口 ( $R(t)$ )
- 一定確率で感受性, 感染, 回復へ順に遷移



Figure 1: SIR モデルの構造

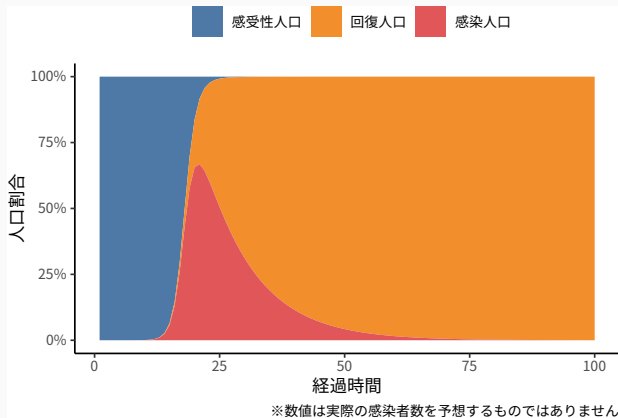


# SIR モデルの数式

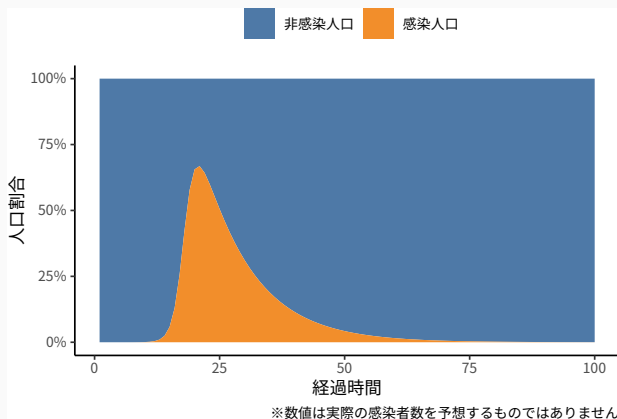
- $\beta, \gamma$  はパラメータ
1. 一定期間ごとに未感染者から  $\beta$  の割合で感染者が発生
  2. 感染者は  $\gamma$  の割合で感染から回復

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta S(t)I(t), \\ \frac{dI(t)}{dt} &= \beta S(t)I(t) - \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t)\end{aligned}$$

# 計算結果のプロット

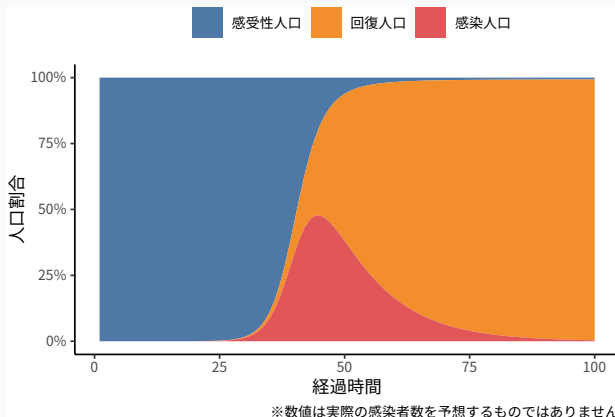


# 感染・非感染の二分割表示



# パラメータを変えると？

- さっきより感染力パラメータを小さくする
- 確かに感染増加のタイミングが遅れた
- ピーク時の感染者数が低下した



# SIR モデルの性質

- あるタイミングで爆発的に感染者が増加
- ピークに達した後は減少する
- 感染力が低いと感染規模・感染スピード低下

# パンデミックの発生条件

- 感染者より回復者のほうが増えるスピードが速ければ感染者は増えないので
- 感染初期はほとんど感染者がいない
  - この条件をもとに微分方程式を解く [4].

$$I(t) = I(0) \exp((\beta N - \gamma)t)$$

- 以下が増加の条件

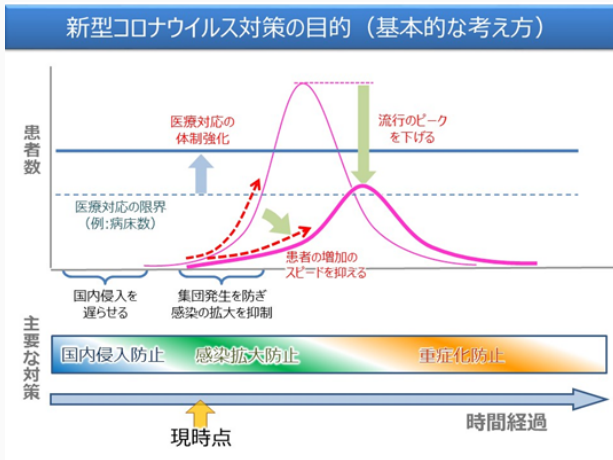
$$R_0 = \frac{\beta N}{\gamma} > 1$$

## 現実置き換えると

- COVID-19 は 飛沫感染と接触感染のリスクが高いとされる (厚労省)
- 人間同士の距離が近く, より多くの人間と接触があるほど感染力が  $\beta$  が増加
- $R_0 > 1$  は人口密度が一定の値を超えるかどうかを表している
- 外出を避け繁華街を出歩かないようにすることはパンデミックの抑止に効果がある

# 厚労省の図

- 厚生労働省のサイトでも似たような図
- ここまでの SIR モデルの話と合致する





# Rでのやりかた

- グラフ図は ggdag パッケージで作成
- 人口構成のグラフはもちろん ggplot2
- SIR モデルのシミュレーションは常微分方程式を解く必要
  - `deSolve::ode()` を使う
  - 以下も参考になる <http://statmodeling.hatenablog.com/entry/sir-model-ode-1>
- 今回紹介したものは基本的に全部これで解ける

## deSolve::ode()

- 引数: 初期値, 区間, 方程式を表す関数, パラメータ
  - 関数はパラメータと  $t$  に対する
$$dS(t)/dt, dI(t)/dt, dR(t)/dt$$
をリストで返す必要
  - with() を使うと見やすい

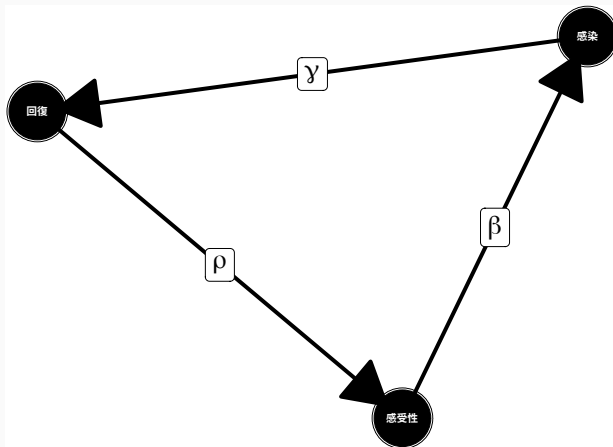
```
sir <- function(t, y, params){  
  with(  
    as.list(c(params, y)),  
    list(c(  
      S = -beta * S * I,  
      I = beta * S * I - gamma * I,  
      R = gamma * I)  
    ))  
}
```

# より現実を反映したモデル

- SIR モデルでは回復した人間は二度と感染しない
- - 新規感染者数は感受性人口と感染人口と感染力  $\beta$  の積で決定
  - 回復人口が増えると新規感染者数が減る
- SIR モデルのピークが一度きりの理由
- 当てはまりやすい例: はしか
- 全ての伝染病に当てはまるわけではない

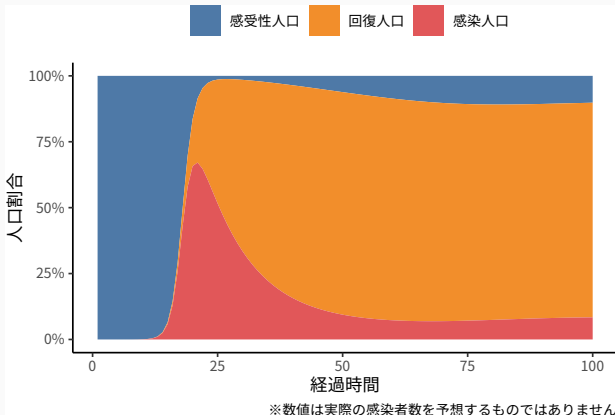
# 再感染の可能性を考える

- 回復した人が再感染する場合を考える
- 免疫を失う人の割合を  $\rho$  として, SIR モデルを改造



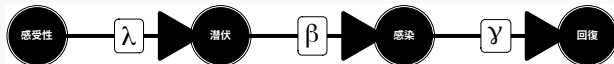
# 再感染のシミュレーション例

- 1% の確率で免疫を失う場合



# SEIR モデル

- S と I の間に E を追加
- E は潜伏期間
  - 感染したが発症も他人に伝染してもいない
- 正確なパラメータがないと SIR と大差ない



# COVID-19 のシミュレーション

- 早くも 2 つの研究が arXiv に投稿されている
  - Zhou et al. [3](YJ-SEIR), Peng et al. [1](一般化 SEIR)
- 中国国内の COVID-19 流行をシミュレーション
- いずれも隔離政策の影響を考慮した修正版 SIER モデル

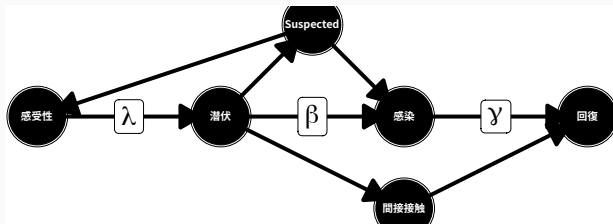
# YJ-SEIR モデル

- 「隔離政策以降, 新規患者数増加が鈍化している」
- 現在わかっていること:
  1. 武漢市では交通を遮断, 住民の外出も制限
  2. 臨時の救急医療体制を構築し, 効果が出ている
  3. 交通を遮断する以前には多くの人間の流入出があった
  4. COVID-19 は間接接触でも感染する可能性
- 隔離を行った後でも (3), (4) で感染拡大の可能性
- 「これまでの感染数推移は **SEIR** モデルの結果と異なる」



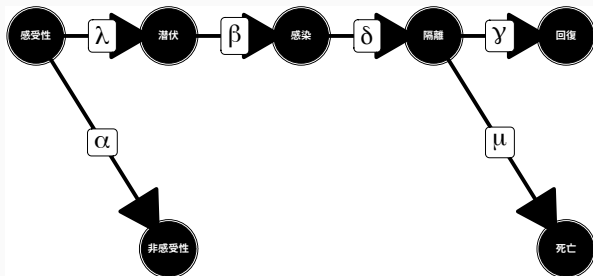
# YJ-SEIR モデルの構造

- 2つのグループを追加
  - 感染が確認されていないため未隔離の人間
  - 間接接触で感染した人間
- 残存した COVID-19 による汚染の影響を考慮



# 一般化 SEIR モデル

- YJ-SEIR よりもシンプル
- 以下のグループを追加
  - 感染の恐れがなくなった非感受状態  $P$
  - 感染がわかると隔離状態  $Q$  に移行
  - 死亡状態  $D$  を明示的に回復人口から分離

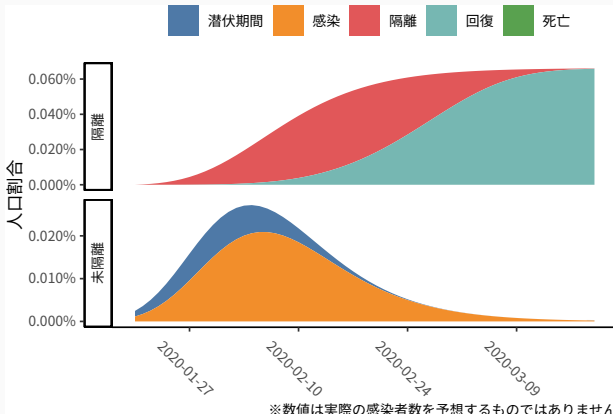


# G-SEIR モデルの再現

- 著者らの主目的は
- COVID-19 が発生した時期
- アウトブレイク終息時期の推定
- まずモデルのパラメータを正確に推定する必要
  - 具体的な数値が掲載されていたので真似してみる
  - 湖北省 (除外武漢市) の推定結果を使う

# G-SEIR モデルの結果

- だいたい 50 日くらいで終息するという結果に



## 補足 1

- 回復率・死亡率パラメータは時間によって変化
  - 徐々に治療体制が確立されているため
  - 回復率は増加, 死亡率は低下
- プログラムではパラメータも変数として計算する
- 具体的な数値はなかったので適当に設定

$$\mu(t) = \exp(-0.9(t + 3)) + 0.001,$$

$$\gamma(t) = \textit{logistic}((t - 60)/10)$$

## 補足 2

- このシミュレーションは日本にも適用できるか？
- 適切にパラメータを推定すればできる
- 感染力パラメータの意味
  1. ウィルスの特性
  2. 自然環境の特性
  3. 社会環境の特性
    - (2), (3) は日中でかなり違う
    - 数字をそのまま持ってきてても無意味
- 専門的な話は **WHO** とか **NIID** とか厚労省とかの発表を待て
- マルチエージェントシミュレーションの話もやりたかったけど時間が足りなかった

## わかったことのまとめ

- SIR モデルでは感染 > 回復でパンデミックが始まる.
  - 感染増加を抑えるとパンデミックの開始を遅らせたり, ピーク時の感染者数を軽減できる.
- 感染力変化モデル, 再感染モデル, マルチエージェントシミュレーションなどがある
- 今回の COVID-19 の隔離措置を考慮したシミュレーションした研究がある
- 実際に感染がどれだけ広がるのか, それはまだ混沌の中...
- それが....., パンデミック...!

# References

---

- [1] Liangrong Peng, Wuyue Yang, Dongyan Zhang, Changjing Zhuge, and Liu Hong. Epidemic analysis of COVID-19 in China by dynamical modeling. February 2020. URL <http://arxiv.org/abs/2002.06563>.
- [2] Emilia Vynnycky and Richard G. White. *An Introduction to Infectious Disease Modelling*. Oxford University Press, New York, 2010. ISBN 978-0-19-856576-5.



- [3] Yimin Zhou, Zuguo Chen, Xiangdong Wu, Zengwu Tian, Liang Cheng, and Lingjian Ye. The Outbreak Evaluation of COVID-19 in Wuhan District of China. February 2020. URL <http://arxiv.org/abs/2002.09640>.
- [4] 寿稲葉. ケルマック-マッケンドリック伝染病モデルの再検討. 2002. URL [https://www.ms.u-tokyo.ac.jp/~inaba/inaba2002\\_KMmodel.pdf](https://www.ms.u-tokyo.ac.jp/~inaba/inaba2002_KMmodel.pdf).