

感染症数理モデルとCOVID-19

東京大学

稲葉 寿 氏

進行役 よろしくお願ひします。セッションB-3「感染症数理モデルとCOVID-19」を始めさせていただきます。なお、本セッションの発表者は、東京大学大学院数理研究所研究科の教授であらせられます、稲葉寿先生から発表をいただきます。

それでは、稲葉先生、よろしくお願ひいたします。

稲葉 ただ今ご紹介にあずかりました、稲葉です。本日このような機会をいただきまして、本当にありがとうございます。

実は10年ほど前に、やはりアクチュアリー会で感染症数理モデルについてお話しさせていただいたことがあるのです。その当時は、高病原性インフルエンザなどで非常に話題になっていたところですが、まさか10年後に、このCOVID-19のようなパンデミックが起こって、再びお話しすることになるとは予想もつかなかったところです。



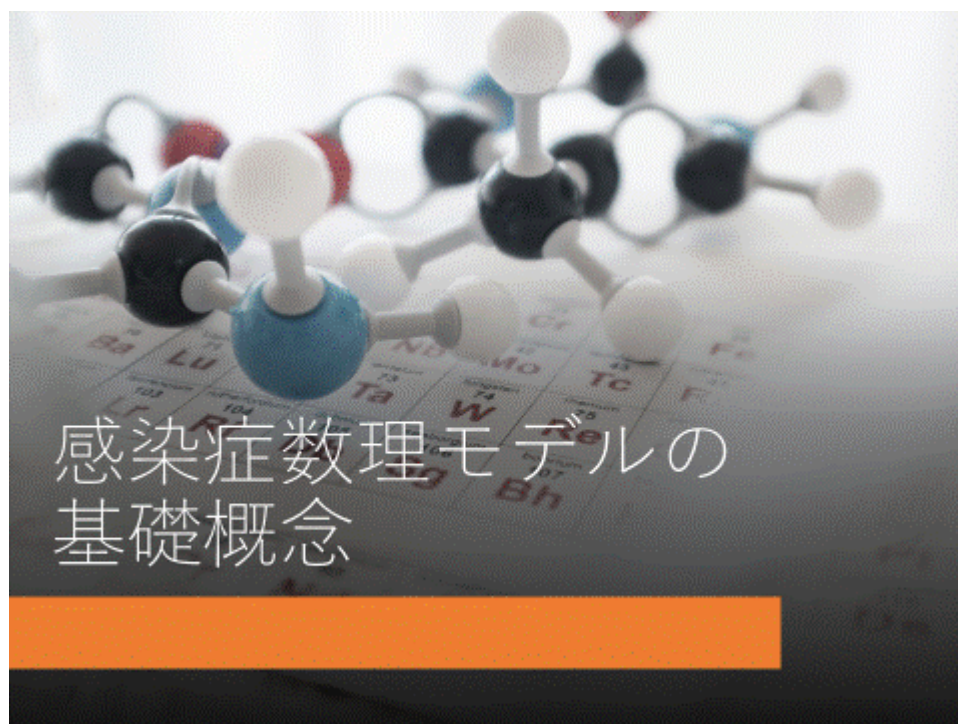
今日は、「感染症数理モデルとは、一体どのようなものか」ということを前半にお話しさせていただいて、後半では、「今度のCOVID-19で、どのようなことが問題になってきているか」を、少しお話ししたいと思います。

References

- 『数理人口学』(東京大学出版会, 2002年)
『感染症の数理モデル』(編著, 培風館, 2008年)
『人口と感染症の数理』(共著, 東京大学出版会, 2014年)
[Age-Strucured Population Dynamics in Demography and Epidemiology](#), (Springer, 2017)
- https://www.researchgate.net/profile/Hisashi_Inaba



恐らくアクチュアリーの皆様は、数学にかなり詳しいのではないかと思いますので、リファレンスとして、このような著書を出しております。『感染症の数理モデル』は、そのものずばりの本ですが、これは、今、絶版ですけれども、重版を作ってもうすぐ出版される予定です。数学的な詳細につきましては、このような本を参照していただけるといいと思います。



最初に、感染症の数理モデルの基礎的な概念についてです。実は、人口学、デモグラフィという学問分野があります。アクチュアリーの皆様は、生命表を通じてデモグラフィに対して、ある程度の知見を持っていらっしゃると思います。感染症モデルとデモグラフィには非常に近い関係があるので、そこに少しフォーカスして、お話ししたいと思います。

感染症数理モデルの歴史(1)

- ▶ 数理疫学の起源は18世紀のDaniel Bernoulliによる天然痘死亡率の寿命への影響に関する研究(1766)に遡る(compartment model).
- ▶ 流行曲線の数学的記述など、記述統計的な側面は19世紀中頃のWilliam Farr(1807~1883)にはじまる(Farr's law).
- ▶ 動的モデルとしての側面は、1920~30年代のKermack and McKendrickの研究(微分方程式モデル)に始まる。ダイナミカルモデルは、70年代に至るまであまり進歩がなかったが、80年代に数理生物学の勃興とともに、急激に進歩して、応用数理の有力な研究分野になった。
- ▶ 現在の数理疫学は、動的モデルをできる限り活用した統計的データサイエンスで、政策実装を目指している。

最初に少し歴史を話しますと、18世紀に Daniel Bernoulli が、天然痘死亡率の寿命への影響に関する研究をやっており、それが感染症数理モデルの、ほとんど最初の研究です。

19世紀になりますと、エピカーブですね、今でも盛んに COVID-19 で出てきますけれども、感染者数のカーブなどを記述統計的に考えることが William Farr によってはじまりました。この方は統計学で有名な人ですけれども、初めて一種の山型のエピカーブを出してきた人物です。

それから、皆さん、よくご存じの Kermack and McKendrick が 1920 年代から 30 年代に出てきまして、微分方程式のモデルとして感染症数理モデルを初めて明確に定式化したわけです。このことは、今回のコロナ禍でも Kermack and McKendrick モデル、あるいは S I R モデルが、あちらこちらで語られたので、皆さんも目にしておられると思います。いまだに、これが最も基本的な微分方程式モデルですが、あとで具体的にお示ししますが、Kermack and McKendrick はもっと一般的なモデルを提起していて、常微分方程式モデルはその特殊なケースに過ぎません。1970 年代までは、微分方程式モデルは非常にゆっくり進歩していたのですが、80 年代頃になりますと、数理生物学というものが成立してきて、その中では活発に研究される分野になりました。応用数理の有力な分野でもあるわけで、非線形力学系としてよく研究されてきています。

この春頃に、例の西浦先生の活躍で有名になりましたけれども、盛んに「数理疫学」などと言われたものは、このような動的な力学系的なモデルを活用しながら、同時に、データ・サイエンスとして感染症のいろいろなデータを解析して、そこから実態を探っていく、あるいは、政策実装を考えていく、そのような分野になってきているわけです。

感染症数理モデルの歴史(2) since 1990

- ▶ 1990年代： **R_0 の定義の確立**とともに、生物数学的な研究（定性的・力学系的研究）は急速に拡大して、数理生物学の有力分野になった。欧米では、生物統計的研究とともに、モデルが実際の感染症流行の推計や予測にも広く使われるようになったのは90年代のエイズパンデミックの頃である。
- ▶ 2000年代：その後、2000年代にあいついだ新興・再興感染症対策で統計解析、データサイエンスとしての側面の研究が進展して、防疫政策への実装が進んできた。**この面では日本は乗り遅れた。**
- ▶ 今回のCOVID-19で、感染症制御・対策の手段として広く認知されるに至った。**データ分析方法等は進んだが**、予測・制御のツールとしては、新感染症への応答に関するホスト人口の情報が不十分という本質的な問題を抱えている。
- ▶ 政策実装をめざす実践的な統計・データサイエンスの部分では、**日本の教育・研究体制は非常に遅れている**のが現状。公衆衛生や疫学における適切な教育コースも存在しない。今回その弱点が露呈した！

もう少し詳しく申しますと、90年代頃に一つのエポックがありまして、今日お話しする basic reproduction number（基本再生産数）の定義が確立したのが90年頃だったのです。特に90年代にはエイズ・パンデミックがあったわけですが、それをきっかけにして感染症数理モデルがサイエンスとして発展したのです。しかし、残念ながら、日本では、世界的な動向には、いわば乗り遅れたところがありまして、このようなモデルを使って感染症対策を考えていく、あるいは統計解析をするという意味では、日本にそれを本当にやっていくグループや大学の講座などはなかったわけです。その弱点が、今回のCOVID-19で出てきたと思います。日本では、そういった世界水準のデータ・サイエンスとしての数理疫学をやっているところが西浦先生のチームぐらいしかないわけです。

Bernoulli's model (Dietz and Heesterbeek MB2002)



$$\frac{du}{da} = -[\lambda(a) + \mu(a)]u, \quad (1)$$

with the initial condition $u(0) = 1$.

The probability $w(a)$ to be immune and alive is given by

$$\frac{dw}{da} = [1 - c(a)]\lambda(a)u(a) - \mu(a)w, \quad (2)$$

with the initial condition $w(0) = 0$.

The solutions of these equations are

$$u(a) = \exp \{ -[A(a) + M(a)] \}, \quad (3)$$

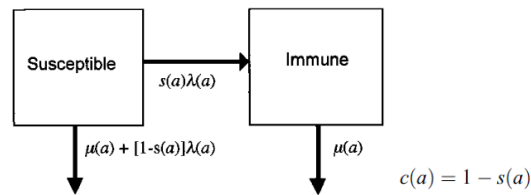


Fig. 2. States and transitions of Bernoulli's epidemiological model for an immunizing infection in a cohort which is in equilibrium with respect to time. $s(a)$ = probability of surviving the infection. $\lambda(a)$ = force of infection; $\mu(a)$ = death-rate due to other diseases.

先ほど言った Bernoulli のモデルを少しだけお見せします。これは面白くて、これから説明するモデルの原型になっているのですけれども、コンパートメント・モデルといって、人口の中を感染状態で分けて、その間の遷移を考えています。Bernoulli は、人の一生という長い年月を考えたので、寿命に対する影響に興味がある。感染している期間は非常に短いので、感染状態にある期間は無視しています。つまり、Susceptible（感受性）—まだかかっていない人—の割合が $u(a)$ で、感染した時点で、死んでしまう人と治ってしまう人の二つに分け、治った人は、すぐ回復状態になるという単純なモデルです。 $1-s(a)$ が天然痘の致死率です。このようなモデルを考えて、生残率 u の微分方程式を立て、生残率曲線を作るわけです。

天然痘死亡率が除去された生残率曲線

8

K. Dietz, J.A.P. Heesterbeek / *Mathematical Biosciences* 180 (2002) 1–21

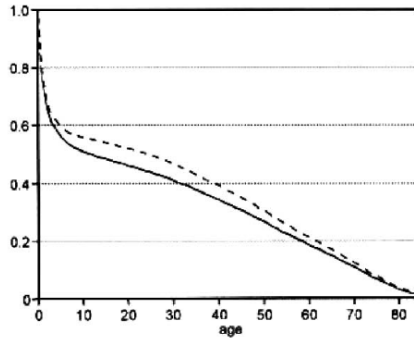


Fig. 4. The life table based on Halley in the state with smallpox (continuous line) and without smallpox (broken line). The median age would increase by 14 years from about 11.5 to 25.5 years!

この実線は非常に古い 17 世紀のハレーの生命表による生残率です。天然痘がある状態での生命表生残率です。そして、点線は天然痘をなくしたらどのような生残率になるかという生残率です。皆さんご存じのように、生残率曲線の積分が寿命ですから、この差の面積が延びた寿命です。「天然痘がなくなると、これだけ寿命が延びる」という考え方をしているわけです。このように、人口を感染状態によって分けて、その間を個体群が遷移するということが、基本的な考え方になっています。

感染症数理モデルの問題意識

力学系としての数理解析（数理生物学）

- 流行が起きるための条件（侵入の閾値現象）を、基本再生産数 R_0 を定めることで明らかにする。
- 再感染やホスト人口の demography を考慮した長期の問題において、エンデミック定常解の存在や安定性を明らかにする。
- 人口学的・社会的・疫学的異質性などホスト個体群の構造をモデルに取り入れて、感染流行のより詳細な数理構造を明らかにする。

データサイエンスと政策実装（数理疫学）

- 報告データから新規感染者数の推定、 R_0 、 R_t や世代時間分布等を具体的に計算・推定する、感染パターンの解析、接触調査や隔離等、介入効果の定量化、流行規模予測、など実践的要求に応える。

問題意識としては、一つには、流行が起きるための条件を、幾つか基本的なパラメータ、基

本再生産数を定めて明らかにすることです。更に、1回の流行だけではなくて、感受性人口の補充などがありますと、ずっと人口の中に感染症が定在するなどということにもなってくるのですが、そのようなことを明らかにする。つまり、流行の数学的な構造を明らかにすることが、力学系としての数理解析の問題意識です。

もう一つは、先ほど言ったデータ・サイエンスと政策実装で、実際に基本再生産数や、いろいろなパラメータを計算して、それに対する政策の効果を考える。それで、どのような政策が望ましいかと考えていく。つまり、より実践的な要求に応えるようなデータ・サイエンスとしての問題意識があるわけです。


Key idea

基本再生産数 R_0
(R-naught)

なんらかの病原体(ウイルスや細菌など) に対してすべてが感受性(susceptible)を有する個体からなるホスト(宿主)人口(個体群)集団において典型的な1人の感染者が、その全感染期間において再生産する2次感染者の期待数を**基本再生産数(basic reproduction number)**とよび、 R_0 で表す。

感染人口を世代毎にみて、1次(初期)感染者(primary cases)、2次感染者(secondary cases)、3次感染者等を継続的に考えた場合、 R_0 は等比級数的に変化する各世代の感染者サイズの(漸近的な)公比である。

Diekmann, Heesterbeek & Metzによって1990年に数学的な定義(次世代作用素法)が確立。 J. Math. Biol. (1990) 30: 303–322



On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations

O. Diekmann^{1,2}, J. A. P. Heesterbeek¹, and J. A. J. Metz¹
¹Centre for Mathematics and Computer Science, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands
²Institute of Theoretical Biology, Leiden University, Kalendsteed 43, 2311 GP Leiden, The Netherlands

Abstract. The expected number of secondary cases produced by a typical infected individual during its entire period of infectiousness in a completely susceptible population is mathematically defined as the dominant eigenvalue of a positive linear operator. It is shown that in certain special cases one can easily compute or estimate this eigenvalue. Several examples involving various structuring variables like age, sexual disposition and activity are presented.

この「基本再生産数(basic reproduction number)」という概念がキー概念です。ここに定義を書きましたが、「なんらかの病原体が、すべて感受性を有するような個体群に侵入したときに、典型的な1人の感染者が、その全感染期間において再生産する2次感染者の期待数」です。ラフに言うと、1人の感染者が、自分が感染者である間に何人の人にうつすか、という数字です。

ジェネレーションとして感染の進行を見てみますと、この基本再生産数は、一つに感染世代と、それが作り出した次世代の感染人口のサイズ比、漸近的にそのようなサイズ比になっていくことが分かります。ですから、世代的に見れば感染人口は等比級数的に増えていくか減少していくか、どちらかです。

R₀の推定値例

現在では、従来考えられたよりも数学の分散はずっとおおきく、桁違いの数字まで観測される。生物的にきまる感染確率・感染性期間と環境によって決まる接触頻度の積なので、数学的定義はともかく、実際の流行への応用に際して、非常に解釈が難しい数字でもある。

感染症	Location and time period	R ₀
麻疹	England and Wales, 1947-50	13-14
	Ghana, 1960-68	14-15
おたふく風邪	England and Wales, 1960-80	11-14
	Netherlands, 1970-80	11-14
風疹	England and Wales, 1960-70	6-7
	Poland, 1970-77	11-12
	Gambia, 1976	15-16
百日咳	England and Wales, 1944-78	16-18
	USA, 1943	16-17
ポリオ	Netherlands, 1960	6-7
天然痘	West Africa, 1968-73	3-5
HIV	Uganda, 1985-87	10-11

出典: Modelling Epidemics, P. Farrington, The Open University 2008.

これは、先ほどの人口学での例えで言いますと、考え方としては、女性が女兒を何人産むかという純再生産率(net reproduction rate)と同じです。自分の子孫、子どもを、どれだけつくるかということです。

この数値例は、いろいろなテキストに出ていますが、幅をもって表されています。それは、基本再生産数が、物理定数のようにシャープに決まるわけではなく、モデルと状況に依存しているためです。とはいえ、これを見ると、それほど大きな幅ではないようにも見える。しかし、最近の研究によると、数値幅は実はもっと広いと言われています。後でお示ししますように、基本再生産数は、生物学的に決まるパラメータ部分と、社会的に集団の特性として決まる部分がありますので、そのような社会集団の特性の部分は非常にレンジが大きいわけです。ですので、実際にそれを計測することも、それを適用して何かの結果を出して結論付けることも、なかなか難しいところがあります。

表 10.1: 主な国・地域の COVID-19 の基本再生産数 R_0 の推定値

国	基本再生産数 R_0	文献
アメリカ	4.02 (95%CI, 3.69–5.15)	[10]
アフリカ	2.37 (95%CI, 2.22–2.51)	[24]
EU	4.22 ± 1.69	[21]
イラン	2.7 (95%CI, 2.1–3.4)	[15]
インド	5.218 (95%CI, 4.56–5.81)	[25]
韓国	2.6 (95%CI, 2.3–2.9) or 3.2 (95%CI, 2.9–3.5)	[35]
中国	3.28	[23]
日本	1.954 (95%CI, 1.851–2.025)	[6]
日本	1.49 (95%CI, 1.30–1.70)	[16]
日本	2.86 (95%CI, 2.73–2.97)	[20]
日本	2.50 (95%CI, 2.43–2.55)	[31]
ブラジル	3.1 (95%CI, 2.4–5.5)	[29]
ロシア	4.234 (95%CI, 3.764–4.7)	[25]

次の表は、COVID-19 です。いろいろな文献で、各国でどれくらいの基本再生産数が計測されたかを載せています。1.5～2 ぐらいのところから 4 以上まで、結構、ばらけています。このように、いろいろな状況に応じて数字が変わってしまうところが、まず一つの大きな問題です。

感染症流行の閾値原理

—力学系としての挙動—

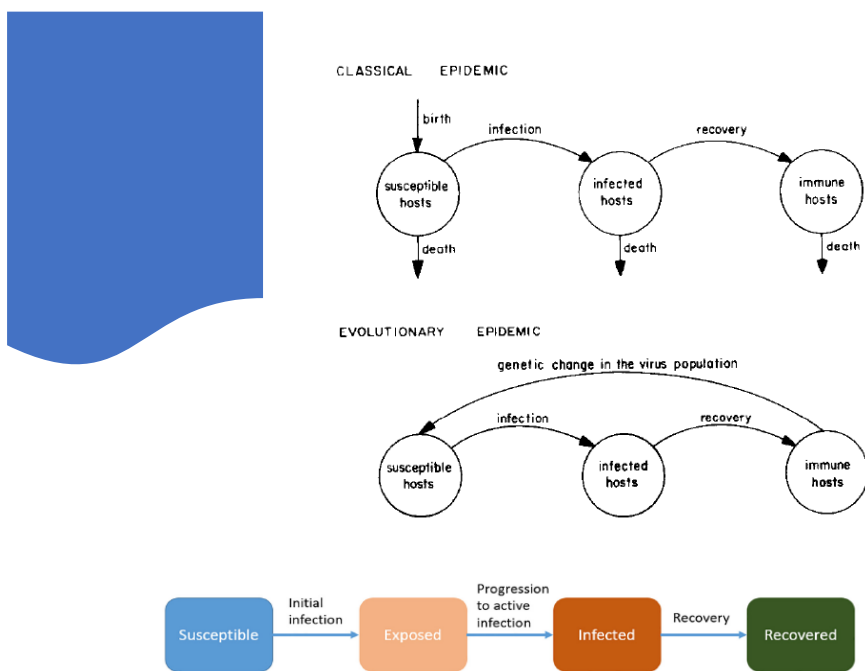
- 感染状態のベクトルがゼロであるような自明な定常解 (disease-free steady state; DFSS) がある。
- **基本再生産数 R_0** が定義される。DFSS は $R_0 < 1$ では局所安定で流行はおきないが、 $R_0 > 1$ では不安定で、流行が起きる (**侵入閾値**)
- $R_0 > 1$ では DFSS が不安定化して、非自明な定常解 (endemic steady state; ESS) が分岐する。このとき、システムは弱パーシステントである (**エンデミック閾値**) 。
- $R_0 < 1$ で DFSS が大域安定ならば、基本再生産数を 1 未満にすることが、防疫上の almighty な原則になる。しかし後退分岐が起きて、 $R_0 < 1$ でも複数のエンデミック定常解が存在することがある。

最初に、この基本再生産数 R_0 を使った閾値原理を述べておきます。これは微分方程式の言葉で書いてありますが、一つの力学系を考えて、感染症がない定常状態、何も病気がはやっていない状態、そのような感染者がゼロであるような定常解があって、そこで感染者が発生した

ときに、定常解がいわば摂動を受けるわけです。そのときに、それが力学系として安定か不安定か、ということに対応しているわけです。ですから、基本再生産数 R_0 が 1 より大きければ不安定で、感染人口が成長を始めるわけです。 R_0 が 1 より小さければ、そこは、ローカルには安定ですから、そこで自然に減衰して流行は起こらないということになるわけです。

更に、今、侵入の閾値性ということだけを言いましたけれども、外から感受性の人口を補充してやる。例えば、出産がある、人口移動があるなどしますと、感染者が、ずっとその中に定在するような定常解ができます。それは、先ほどの感染症のない定常状態から新たな定常解が分岐してくるとい現象が起こるわけです。基本再生産数は、そのような閾値にもなっています。

少し先走って言えば、仮にローカルに R_0 が 1 より小さく、disease-free の状態が大域安定であれば、 R_0 を 1 より小さくせよということが、オールマイティーな疫学上の指針になるわけです。けれども、力学系として見ると、実はもう少し複雑で、そのような基本再生産数が 1 より小さい劣臨界の状態でもエンデミックな状態があるような力学系もあるのです。そのような場合には、 R_0 を 1 より小さくしても、なかなか根絶がうまくいかないことも予想されるわけです。



SIR とよく言われている数理モデルは、ホスト人口集団の中を、susceptible (感受性) と infected (感染) と immune (回復・免疫) の三つに分けて、その間を、感染したら susceptible から infected に移る、治ったら immune に行く、などというようにしているわけです。それから、図の中段、immune から susceptible に行っている矢印がありますが、これは、一旦治って immunity を持っても、また再感染が発生する、感受性が戻ってくる場合に対応しています。これは、COVID-19 でも疑われています。再感染が起きる可能性があるということです。そうすると、力学系としての性質は、また変わってきてしまうわけです。

下の方のカラーで書いてあるコンパートメントの流れ図は、今、COVID-19 でよく使

われているモデルで、Exposed クラス（潜伏期）は Infected ですが、まだ感染性を持たないという状態です。このような状態を間にかませることもあります。あるいは、asymptomatic（無症候性）で、まだ症状が出ないものと出るものに分けるなど、用途によってコンパートメントの分け方は、いろいろあります。

The origin of SIR modeling:
-1st Kermack and McKendrick model-

- ▶Kermack and McKendrick I (1927): one outbreak model
- ▶Kendall ('56), Reddingius ('71), Metz ('78), Diekmann ('77)

$$\begin{aligned} \frac{dS(t)}{dt} &= -\lambda(t)S(t), \\ \frac{\partial i(t, \tau)}{\partial t} + \frac{\partial i(t, \tau)}{\partial \tau} &= -\gamma(\tau)i(t, \tau), \\ i(t, 0) &= \lambda(t)S(t), \\ \frac{dR(t)}{dt} &= \int_0^\infty \gamma(\tau)i(t, \tau)d\tau, \\ \lambda(t) &= \int_0^\infty \beta(\tau)i(t, \tau)d\tau. \end{aligned}$$

これは、さらりと見ていただければいいです。Kermack and McKendrick が書いた 1927 年の論文には何が書いてあったかを、今風に書くと、こうだという話です。Kermack and McKendrick のモデルはよく引用されますけれども、原文をきちんと読んだ人は、恐らくほとんどいないでしょう。今風に書いていないので、実は、非常に難しい論文です。それを現代的な力学系として構成してみると、このようなことを言っています。

感受性の人 λ という感染力で感染して減少します。そして $i(t, \tau)$ は、時刻 t で感染してから経過時間が τ であるような人口の密度です。その人たちは、 $\tau = 0$ 、境界値のところに新しく感染した人として現れるわけです。だから、 $i(t, 0) = \lambda(t)S(t)$ です。 $\gamma(\tau)$ は、感染してから経過時間に依存する回復率です。回復した人は R のクラスに入って来て、 R が増えていくという式です。さて λ をどうするかが問題ですが、 $i(t, \tau)$ という時刻 t 、感染からの経過時間 τ の感染者が $\beta(\tau)$ という感染率を持っているとします。その正体が何であるかは、また問題ですが、 λ はある種の重みを付けた感染人口サイズに比例しているということです。感染人口が増えれば感染力が増えるということです。感受性人口と感染人口の間に非線形の相互作用があるわけです。

最も単純な SIR モデル

いまサイズ N の封鎖人口を想定して、 $I(t)$ を時刻 t の感染人口サイズとすれば、よく知られたように、

$$\begin{aligned}\frac{dS}{dt} &= -\beta I(t)S(t) \\ \frac{dI}{dt} &= \beta I(t)S(t) - \gamma I(t) \\ \frac{dR}{dt} &= \gamma I(t)\end{aligned}\tag{1}$$

という SIR モデル（前期ケルマック・マッケンドリックモデル）が定式化できる。 β のかわりに N を外に出して β/N を用いる場合もあるが、 N は定数なので定性的な差はない。

ここで、 S は感受性人口サイズであり、 R は回復人口サイズ、 γ は回復率、 β は伝達係数である。 $N = S + I + R$ は定数で、 βN は感染1個体の単位時間における平均接触数と一回の接触あたりの感染成功確率の積であるので、単位時間に生産する二次感染者数を与える。

このモデルは非常によく検討されているわけですがけれども、これは、よく皆さんが見る簡単な S I R モデルです。先ほどの内部の感染期間（経過時間 τ ）というパラメータへの依存性を無視してしまいますと、普通の常微分方程式のモデルになります。ここでは β が伝達係数で、この $\beta I(t)$ が先ほどの感染力に相当しています。ですから、感染力は感染人口サイズに比例している、ということが基本的な仮定です。 S は、それで減少します。今、外との出入りを考えない、短い時間スケールの流行を考えています。また、感染人口は $\beta I(t) S(t)$ で増えて、 $\gamma I(t)$ で減っていくということです。そのようなモデルになります。 $S(t)$ と $I(t)$ と $R(t)$ を足しますと、全人口になりますが、それは保存されているという系です。

初期の侵入：線形化方程式と R_0

このとき感染者や回復者のいない定常状態は $(S, I, R) = (N, 0, 0)$ であるが、そこでの線形化方程式は、

$$\frac{dy}{dt} = \beta N y - \gamma y = (\beta N - \gamma)y \quad (2)$$

となる。 y は初期の感染人口サイズであり、基本再生産数は

$$R_0 = \frac{\beta N}{\gamma} = [\text{単位時間あたりの二次感染者産生数}] \times [\text{平均感染性期間}] \quad (3)$$

としてえられる。ここで、平均感染性期間は $T = \frac{1}{\gamma}$ である。

R_0 は一人の感染者がその全感染性期間のあいだに生産する二次感染者の平均数に他ならない。初期感染者の増加率は

$$\lambda = \beta N - \gamma = \gamma(R_0 - 1)$$

符号関係

$$\text{sign}(R_0 - 1) = \text{sign}(\lambda)$$

がなりたつ。すなわち、 $R_0 > 1$ であれば、感染症のない平衡状態 $(N, 0, 0)$ は不安定になるが、 $R_0 < 1$ であれば安定である。これが閾値原理に他ならない。ただし、この侵入閾値は局所的なもので、大域的なものではない。感受性の補充があれば劣臨界でもエンデミックな定常解は存在しうる。

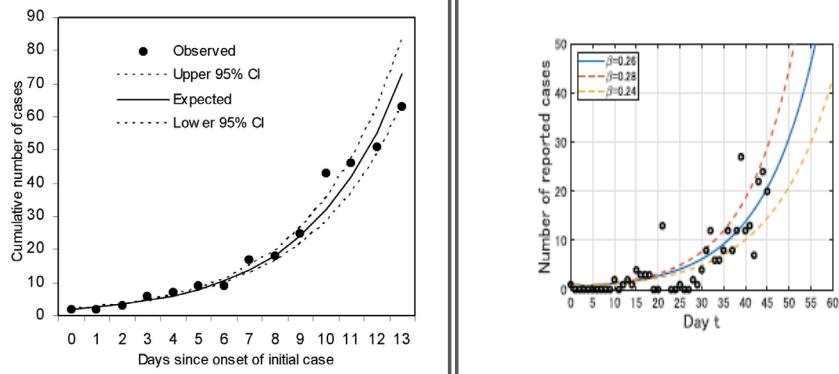
最初に、「全体のホストが全部感受性だったら、どうか」ということを考えます。そのときに、 (S, I, R) の状態は $(N, 0, 0)$ です。 N は全人口サイズです。「 $(N, 0, 0)$ という状態に少数の感染例が発生したら、どうなるか」を考えるということで、その定常状態での線形化方程式を考えなさいということです。だからこの(2)のように、—この y は摂動された少数の感染人口サイズなので記号を I から y に変えていますけれども— S は全部 N に置き換えてやりますと、このような線形の方程式が出てきます。ですから、ここの $\beta N - \gamma$ が、初期の侵入した感染者の成長率になります。この $\beta N - \gamma$ を変形して $\beta N / \gamma$ という数字を作ります。これが基本再生産数になっています。

なぜかと言いますと、 γ は減衰率ですから、 $1 / \gamma$ は感染者としての滞在時間の平均値です。だから、単位時間あたりに βN だけの2次感染者を作り出しますから、それに $1 / \gamma$ を掛けると、自分が感染者である間に作り出した2次感染者の数が出てきます。それが R_0 です。 R_0 を使いますと、先ほどの $\beta N - \gamma$ は $\gamma \times (R_0 - 1)$ ですから、 $R_0 - 1$ と成長率 λ の符号は同じです。ですから、 R_0 が1より大きければ成長率 λ はプラス、 R_0 が1より小さければ成長率はマイナス、という対応関係が出てきます。これが R_0 によって表現された感染症侵入の閾値条件です。

では、成長率を最初から使えばいいではないか、と思うかもしれないですけれども、 R_0 は個体の特徴づける再生産のパラメータからできています。「自分が、どれくらいの感染者としての寿命を持っているか」、「単位時間に何人にうつすか」、つまり、そのような個体が、どのような再生産をするかを決めていて、それがマクロで見た全体の成長率と、どのような関係にあるかを示しているわけです。感染性の期間を短くしたり、感染率を下げるなど、いろいろな介入行為は個体に対して働くものです。それが集団の成長率にどのように影響するかが、この R_0 を通じて定式化できることになります。

それは人口問題でも同じです。人口学における R_0 は人間の純再生産率です。どのように産

み方を変えれば、全体の人口の成長率はどのように変わるか、ということを決めているのが R_0 の関係式（特性方程式）です。成長率はマクロ・パラメータですから、直接何かコントロールできるわけではない。個体の行動変容を通じてマクロが変わってくる、そのようなところを捉えていることが大事です。



流行初期の感染人口のマルサスの成長の例：
 Spanish influenza in Maryland, 1918 (左) (Nishiura, 2008),
 and COVID-19 in Japan 2020 (右) (T. Kuniya, 2020)

「では、実際、それほど指数関数的成長をしているのですか？」と言うと、この左側はスペイン・インフルエンザの例ですが、Maryland の 100 年前の例です。右側は、今度の COVID-19 で、神戸大学の國谷さんの論文から取りました。数が少ないので凸凹しているのですが、指数的な成長が観察できるわけです。

臨界免疫化割合

$\epsilon \in [0, 1]$ を免疫化割合として、ワクチンによって、初期状態が $((1-\epsilon)N, 0, \epsilon N)$ に変化したとしよう。この部分免疫化された定常状態における線形化方程式は、

$$\frac{dy}{dt} = \beta(1-\epsilon)Ny - \gamma y \quad (4)$$

となるから、その実効再生産数（なんらかの介入行為がある場合の感染者再生産数）は、

$$R_e = (1-\epsilon)\frac{\beta N}{\gamma} = (1-\epsilon)R_0 \quad (5)$$

となる。したがって、 $R_e < 1$ となるように ϵ を選べば、感染流行の拡大は防ぐことができることになるから、

$$\epsilon > \epsilon^* := 1 - \frac{1}{R_0} \quad (6)$$

となればよい。このとき、 ϵ^* を臨界免疫化割合とよぶ。すなわち ϵ^* より大きな割合の感受性ホスト人口が免疫化されれば、流行はおきない。

もう一つ、そこから出てくる重要な概念があって、臨界免疫化割合というものです。これは、最初のホスト人口の一定の割合を免疫化します。ワクチンを投与して、もう感染しないようにするわけです。そうすると、線形化方程式は(4)のようになります。 ϵ が免疫化した割合です。したがって、感受性サイズは $(1 - \epsilon) \times N$ になります。このような方程式になりますから、このときの基本再生産数は介入行為のある場合の再生産数という意味で実効再生産数 (effective reproduction number) とよばれますが、これは R_e と書いてありますが、元の R_0 の $1 - \epsilon$ 倍です。

これが1より小さくなるような臨界値は、 $(1 - \epsilon) \times R_0 = 1$ を解いて、 ϵ は $1 - (1/R_0)$ であればいいということです。最近、COVID-19 では $R_0 = 2.5$ が使われましたけれども、そうすると $\epsilon = 60\%$ です。それぐらいが免疫化すると、もう感染人口は成長しないです。それが臨界免疫化割合になります。流行が起きなくなるような免疫化割合の臨界値です。

流行による集団免疫

(1)において、 dI/dt を dS/dt でわれば、 (S, I) 相平面上の軌道方程式として、

$$\frac{dI}{dS} = -1 + \frac{\gamma}{\beta S} \quad (7)$$

が得られる。このとき、初期の実効再生産数 $\beta S(0)/\gamma$ が1を超えていれば、解軌道 $I = I(S)$ は時間とともに増加して、 $S_t := \frac{\gamma}{\beta}$ において最大値をとり、その後減少に転ずる。この感染者人口がピークに達する時間はほぼ

$$t^* \sim \frac{\log N}{\lambda}$$

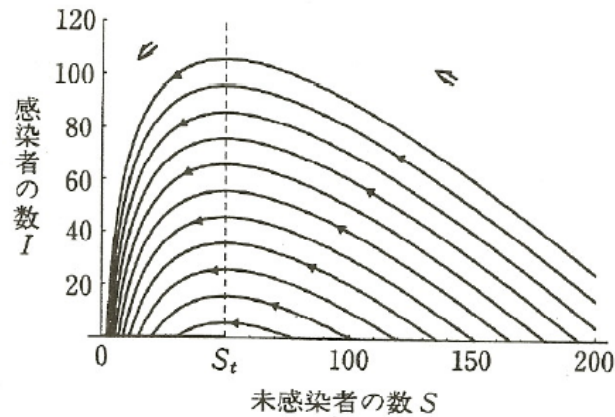
であることが知られている。さて、このピークにおいて感受性人口の割合は

$$\frac{S_t}{N} = \frac{1}{R_0} \quad (8)$$

になる。すなわち、全人口が感受性の状態から、感受性人口割合が $1/R_0$ 以下になるということは、流行による獲得免疫割合が $1 - 1/R_0 = \epsilon^*$ になることによって、流行は自然減衰にむかうことを意味している。

SI平面（相平面）での軌道

(S で $R_t = \beta S/\gamma = 1$ で集団免疫が成立、 R_t は S に比例して減少)

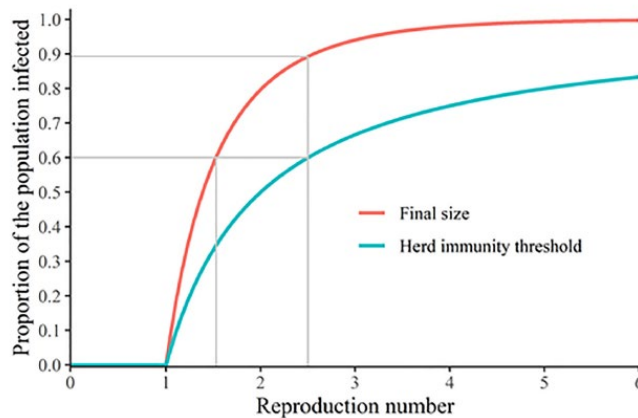


では、ワクチンではなくて、病気が流行して感染が広がって、治った人が免疫化していく場合のプロセスを見てみます。そうすると、この dI/dt を dS/dt で割ると、 (S, I) 相平面上での軌道の方程式が出てきます。 I が S の関数として出てくるわけです。このときに、次のスライドですが、 $S-I$ 平面上にこのようなグラフが描けるのです。この、ちょうど山のピークのところに S_t と書いてありますけれども、ここまで流行が進んでくると、後は減衰していきます。割合としては、(8)の式のように $S_t/N = 1/R_0$ です。

ですから、この山のピークに着くまでに $1 - 1/R_0$ の割合の人が感染して、免疫化されて、

除外されます。そうすると、後は減衰していくだけということになっているのです。このような臨界免疫化が流行のピークで達成されていることとなります。そこに到達するまでの時間が、 t^* と書いてありますが、初期成長率で人口サイズの自然対数を割ったものに大体近いことが分かっています。これは面白い話です。要するに、「Nに到達するまで、成長率 λ で指数関数的に、そのまま走れ」といったときの到達時間とほとんど同じです。ですから、そのホストサイズや初期成長率を分かっていると、いつ頃感染ピークが来るかが決まってしまうという構造になっています。

Article in Paediatric respiratory reviews - June 2020 Modelling insights into the COVID-19 pandemic
DOI: 10.1016/j.prrv.2020.06.014



$$p(t) := \frac{S(0) - S(t)}{S(0)} = 1 - \frac{S(t)}{S(0)}$$

$$p(\infty) = \lim_{t \rightarrow \infty} p(t), \quad p = \lim_{S(0) \uparrow N} p(\infty)$$

$$1 - p = e^{-R_0 p}$$

このスライドは何かと言うと、流行の強度を計算しています（スライド下部の数式参照）。 $p(t)$ は $S(0)$ から $S(t)$ を引いて、初期のホストサイズ $S(0)$ で割り算しているのです。要するに、 $S(0)$ という初期の感受性人口に対して、時刻 t までにどれくらいその中の人々が感染したか、という割合になっています。 $p(\infty)$ は、最終的に感染した人の割合になります。

更に、（添字なしの） p は、 $p(\infty)$ で初期の感染者のサイズを零に持っていきます。力学系の初期値問題としては、 $i = 0$ 、初期の感染者が0であっては流行がスタートしませんけれども、初期データを無限小に持っていくのです。最初は非常に無限小に小さい感染者しかいなかったという極限的なケースを考えたときの最終的な感染者割合、流行強度が p です。この p を満たす方程式が $1 - p = \exp(-R_0 \cdot p)$ という最終規模方程式(final size equation)になっていて、 R_0 が決まるとこの最終規模(final size)が決まってしまうという関係になっています。

最終規模 p は $p(\infty)$ の下限です。それがつまり流行の強度です。そのグラフが、この赤線です。これは、基本再生産数が1を超えないとポジティブに立ち上がってきません。ここで面白いことは、 R_0 が1より小さければ、初期感染者サイズを零に収束させれば、最終規模も零に収束してしまうのですが、 R_0 が1を超えていると、もう最終規模は0に行かないのです。だから、最初の感染規模がどれだけ小さくとも、 $R_0 > 1$ であれば必ず大きな流行が起きるということがいえます。これが R_0 のもつ閾値性です。

もう一つは、先ほどの集団免疫閾値 (Herd immunity threshold)、あるいは臨界免疫化割合が、この水色の線です。例えば、ここの 2.5 のところを見てみますと、集団免疫閾値は 60% です。一方、最終規模は 90% です。60% のところで流行が減衰し始めるのですが、流行が全部終わるまでに、更に 30% 感染してしまう。だから、「集団免疫達成したから、これでオーケー」というわけではないのです。ほっておけばその後もずっと流行が続いて、9 割以上の人が最終的に罹患してしまうわけです。

感染者数のグラフ

感受性サイズ N , R_0 によって流行規模、ピークタイムは決まる
(Example: $R_0=3.77$)

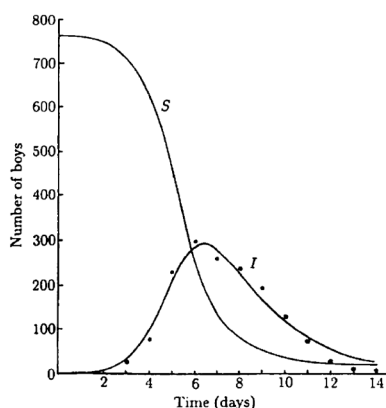


Figure 10.3. Influenza epidemic data (*) for a boys' boarding school as reported in the British medical journal, *The Lancet*, 4th March 1978. The continuous curves for the infectives (I) and susceptibles (S) were obtained from a best fit numerical solution of the SIR system (10.1)–(10.3); parameter values $N = 763$, $S_0 = 762$, $I = 1$, $\rho = 202$, $r = 2.18 \times 10^{-3}/\text{day}$. The conditions for an epidemic to occur, namely, $S_0 > \rho$, are clearly satisfied and the epidemic is severe since R/ρ is not small.

これは、有名なマレー (J. D. Murray) の教科書「数理生物学入門」に出ているものですが、S I R のモデルをグラフで描いてみたものと、イングランドのある寄宿舎でのインフルエンザ流行のデータが、「これは、よく合いました」という話になっています。

ピークタイムと R_0

初期成長率 (内的成長率) と R_0 の関係は

$$\lambda = \beta N - \gamma = \gamma(R_0 - 1)$$

であったから、ピークタイムと R_0 の関係は、

$$t^* \sim \frac{\log N}{\gamma(R_0 - 1)}$$

となる。たとえば、 $R_0 = 2.5, \gamma = 1/14, N = 1.2 \times 10^8$ とすれば、 $t^* \sim 174$ で、およそ、半年後くらいにピークが来ることになる。

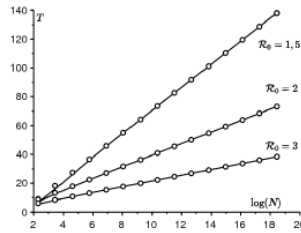


Figure 1. The date T of the epidemic peak according to $\log N$ (solid line) according to (small circles). Parameter values: $R(0) = 1, b = 1/4, \gamma_0 = 1/14 \in (1, 1.5, 2, 3)$.

先ほどのピークタイムとの関係ですけれども、 $t^* \sim \log N / \lambda$ の λ を R_0 の式に置き換えますと、ピークタイムと R_0 の関係が出てきます。サイズ N とピークタイムのグラフは、このように R_0 が大きいほど下の方に来ていますから、同じ N のサイズでも、 R_0 が大きいほど、早くピークに到達するということです。

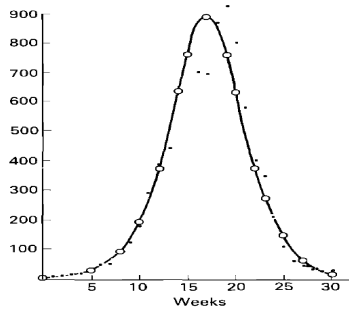


Figure 1. Deaths from plague in the island of Bombay over the period 17 December 1905 to 21 July 1906. The ordinate represents the number of deaths per week, and the abscissa denotes the time in weeks. As at least 80-90% of the cases reported terminate fatally, the ordinate may be taken as approximately representing dz/dt as a function of t . The calculated curve is drawn from the formula:

$$y = \frac{dz}{dt} = 890 \operatorname{sech}^2(0.2t - 3.4).$$

流行曲線 ($dR/dt = \gamma I$ の曲線) の例: ボンベイ (現ムンバイ) のペスト 1905-06

上の図は、Kermack and McKendrick の論文に出ているボンベイ (今のムンバイ) のペストのデータで、これも (モデルとデータが) 非常によく合う。これは、 dR/dt をエピカーブとして描いており、それはほぼ死者の数ですけれども、合うわけです。

ところが、ここで面白いことがあって、その当時のデータは、もう残っていないですけれども、実はこのエピカーブから、このモデルを適用したサイズ N はどれくらいか推定できます。フランスのバカエル (N. Bacaër) という数学者が計算したところ、このエピカーブを信用すると、元の N は数万のオーダーだったのです。しかし、当時のボンベイの人口は 100 万人はいたはずで、100 万の人口に、モデルをそのまま適用したわけではないということです。そこに一つの大きな問題があります。

COVID-19でのODEモデルの適用例 (N は全国人口, SEIR model, by T. Kuniya)

$$\begin{cases} S'(t) = -\beta S(t)I(t), & E'(t) = \beta S(t)I(t) - \varepsilon E(t), \\ I'(t) = \varepsilon E(t) - \gamma I(t), & R'(t) = \gamma I(t), \quad t > 0, \end{cases}$$

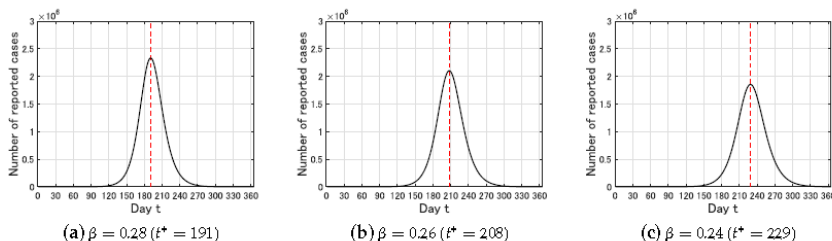


Figure 2. Time variation of the number $Y(t)$ of infective individuals who are identified at time t ($0 \leq t \leq 365$) for $p = 0.1$. The dot lines represent the epidemic peak t^* .

もう一つの例を挙げてから結論を申しますが、これは COVID-19 で、1 億 2,000 万に対して SEIR モデルを適用したときのエピカーブです。2 月初めからスタートすると、半年ぐらいのところにピークが来ているのです。当時、オリンピックの開催が議論になりましたけれども、オリンピックをやっている頃にピークが来てしまうということで、「これは、だめではないか」という一つの根拠になったわけです。しかし一方で、このような推計を、そのままナイーブに信じていいのかという問題があるわけです。

SIRモデルの問題点

- 単純なSIRモデルが適用されるべき（一様なミキシングという仮定が適切である）**感受性人口サイズ**がどのくらいかは、あらかじめ決めることができない。感受性の分布もわからない。
- 封鎖された一様な小人口で、介入がなければ良い予測が期待できる。しかし**複雑大規模な人口**にそのまま適用すれば、良い予測は得られない。回避行動のないデータはあまりない。全人口に単一のR0を適用した一様混合モデルは、往々にして巨大な流行を予想することになる。
- COVID-19のように、無症候性の感染者が、大きな割合で存在する場合、観測されたデータから、モデルによって流行の広がりや推定することは困難であろう。疫学的調査が必要。

つまり、このSIRという非常に単純なモデルは、全体のホストの人口が一様に混合されているような状態を仮定していますが、感受性人口のサイズが大きい場合どれほどの人口のサイズにモデルを適用すべきか、モデルを見ても分からないわけです。「1億2,000万が一様にミキシングすることは考えられない」、それで、「それをすぐ適用するのは、少しおかしい」という話です。

先の例でも出てきた、小さな同質的集団に対しては、実際に観測しても確かにうまく合います。数百のオーダーでは問題なく当たるのですけれども、非常に大きなナショナル・レベルで、そのまま適用していいかは、大きな論点であると思います。ですから、一種の試算に過ぎないと思われるわけです。また、何らの介入行為や回避行動のないデータも滅多にないので、それらを確認することもなかなかできないわけです。

COVID-19の場合には、観測されているデータが、そもそも症候性の感染者のデータだけで、無症候の人の割合は分かっていないという問題点ももちろんあります。やはり疫学的な調査がないと分からないと思います。

実効再生産数の理論 (リアルタイム分析の指標)

再生方程式

SIR モデル (1) の感染者に関する方程式を非同次方程式とみなして定数変化法の公式を用いれば、

$$I(t) = I(0)e^{-\gamma t} + \int_0^t e^{-\gamma(t-\sigma)} \beta S(\sigma) I(\sigma) d\sigma \quad (9)$$

$B(t) := \beta S(t)I(t)$ とおけば、以下の再生方程式 (renewal equation) が得られる:

$$B(t) = \beta S(t)I(0)e^{-\gamma t} + \beta S(t) \int_0^t e^{-\gamma\sigma} B(t-\sigma) d\sigma \quad (10)$$

これは人口学におけるロトカの積分方程式と同様に、新規の感染者が自己再生産する様子を記述している。このとき、積分核

$$\Psi(t, \sigma) = \beta S(t)e^{-\gamma\sigma}$$

が、人口学における「純再生産関数」に相当する。その総和が時刻 t における純再生産率である:

$$R_t = \int_0^\infty \beta S(t)e^{-\gamma\sigma} = \frac{\beta S(t)}{\gamma} \quad (11)$$

とくに、全人口が感受性である場合、 $S(t)$ を N でおきかえた場合が、基本再生産数である。

少し難しい話になってしまうのですが、もう一つ非常に大事なアイデアがあって、それが renewal equation (再生方程式) です。SIR モデルの I の方程式を非同次の線型方程式だと思って定数変化法の公式で積分方程式に書き直しています。そうすると、(9) のような式が出てきます。

これに $\beta S(t)$ を掛けてやると、 $B(t) := \beta S(t)I(t)$ と置いています。B は何かと言うと、これは、ちょうど時刻 t で発生した新規感染者です。その B の積分方程式が導けますが、それが (10) のヴォルテラ型の積分方程式です。解釈は簡単で、時間 σ だけ前に発生した感染者が $\exp(-\gamma\sigma)$ という生存率で生残してきて $\beta S(t)$ だけの 2 次感染者を作りました、それが時刻 t での感染者です、ということです。また、初期感染者から直接発生した二次感染者数は、第 1 項の初期データとして表現されています。

結局、これは感染者が自己再生産する様子を記述している。皆さんは、恐らく、人口学でロトカの積分方程式をやられたと思うのです。子どもが生まれて一定時間生存してきて、そこで、出生率で子どもを産む。形式的には、そのプロセスと全く同じです。

この積分核に相当する $\beta S(t) \times \exp(-\gamma\sigma)$ 、これは人口学でいう出生率 \times 生残率のようなものです。それを、ある時刻 t のところで積分したものが (11) の式、 $\beta S(t) / \gamma$ になります。これがいわば、その時間での瞬間的な再生産率 R_t です。3 月、4 月頃、実効再生産率と話題になりましたけれども、あれは、正にこれです。人口学で言うと、 R_t はある時刻で観測された純再生産率のようなものにあたります。これは S という感受性人口サイズが分かっているなくても B のデータがわかれば測定されるわけです。そのときの時々刻々変わってくるデータから測定されているものがこの R_t で、これはもちろん R_0 と違う値で、時間によって変動しているわけです。

感染者の人口学

もっと一般に、感染からの経過時間に依存した感染率 β と (感染状態への) 生残率

$$\ell(\sigma) := \exp\left(-\int_0^\sigma \gamma(x) dx\right)$$

を想定すれば、以下の再生方程式を得る：

$$B(t) = G(t) + S(t) \int_0^t \beta(\sigma) \ell(\sigma) B(t - \sigma) d\sigma \quad (12)$$

ここで、 $G(t)$ は初期時刻に存在した感染者が生み出す単位時間あたりの二次感染者数である。ここで、感受性人口サイズは、本来の SIR モデル (閉鎖系) では、

$$S(t) = S(0) - \int_0^t B(\sigma) d\sigma$$

となつて、非線形の積分方程式の解として B が決定される。しかし、この方程式を (様々な介入や行動変容で) $S(t)$ が不明で、外部から与えられた関数と考えれば、非自律的な線形積分方程式とみなせる。そのとき、「実効再生産数」は、

$$R(t) = S(t) \int_0^\infty \beta(\sigma) \ell(\sigma) d\sigma \quad (13)$$

となる。これは人口学でいえば、時刻 t での期間出生率に相当する。

上のスライドには「感染者の人口学」と書きましたが、renewal equation の立場から見ますと、「どのように感染者が感染者を生み出す」というプロセスとして流行を考えることができます。ここでは、先ほどの生残率を一般化して一般の生命表生残率のように ℓ と書きます。一般の生残率関数 $\ell(t)$ を使って書くと実効再生産数は、時刻 t では (13) のような (生残率) \times (二次感染者発生率) の積分になっているわけです。

実効再生産数

そこで、二次感染者を生み出す感染齢 (世代時間) の分布を

$$\Phi(\sigma) = \frac{\beta(\sigma) \ell(\sigma)}{\int_0^\infty \beta(x) \ell(x) dx}$$

とおく。初期条件の影響が消えた段階では、

$$B(t) = R(t) \int_0^\infty \Phi(\sigma) B(t - \sigma) d\sigma \quad (14)$$

となるから、 Φ は実測できないが、発症待機時間分布などで代替すれば、

$$R(t) = \frac{B(t)}{\int_0^\infty \Phi(\sigma) B(t - \sigma) d\sigma} \quad (15)$$

よって、感染者数 $B(t)$ (これもしばしば発症者数) を観測すれば、感受性人口サイズが不明の場合の「実効再生産数」が計算されることになる。 $R(t) < 1$ が持続すれば、 B は減衰していると推定できる。

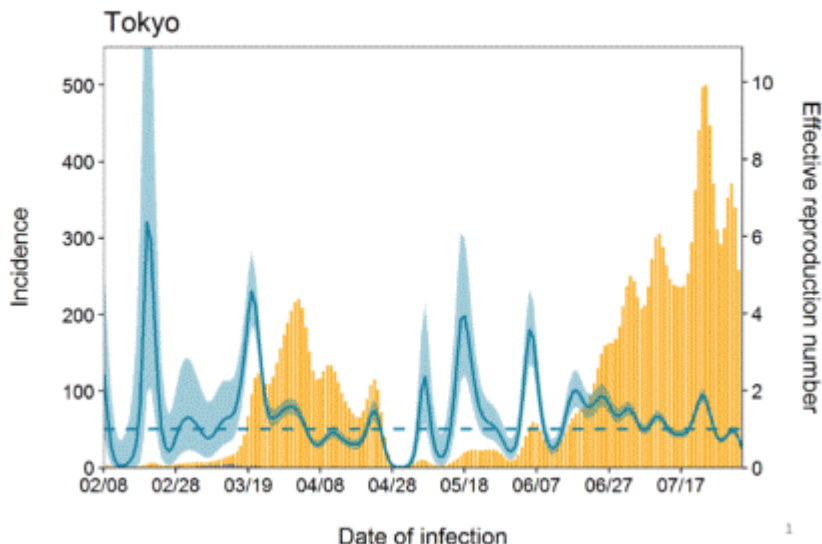
この二次感染者発生率と生残率を掛けたものの分布を Φ と置いてやると、(14) のような新規感染者の再生過程がモデル化でき、このとき、 B が実測できれば、そこから $R(t)$ (このスライド

では $R(t)$ と書いていますが)は計算できます(15)。実は、このようなことを統計的にやって実効再生産数を算出しています。

西浦京都大学大学院医学研究科教授提出資料

資料 1

東京都：実効再生産数（8月19日推定）
使用データと方法：確定日別データから感染日を逆計算



新型コロナウイルス感染症対策分科会（第6回）

COVID-19 の推定では、データとしてでてくる B は本当の意味での新規感染者ではない、実際、ここは発症者で代替していると思うのです。発症しないとそもそも観測できないですから。そして、発症者が発症者を再生産するというような考え方で、日々の実効再生産数を計算して、それが 1 より小さく抑えられていれば、対策は、ある程度うまくいっている、やがて収束が期待できる、という指標として使っているわけです。このスライドは、西浦さんたちが計算している実効再生産数で、新型コロナウイルス感染症対策分科会で発表されたものです。

～休憩 & 質疑応答～

質問 小さいレベルのコミュニティではよく適用するのに、国レベルでは適用しない要因には、何が考えられるのでしょうか？地域ごとの交通の差でしょうか？

稲葉 もちろん、それもあります。やはり、非常にヘテロな集団を考えているわけです。都市と人口希薄な地帯という差もありますし、更に難しく言うと、immunity というものが、そもそも個体によって違うことがあります。今、いろいろなことが言われていて、ファクター X などと言われますけれども、そもそも個体の感受性や感染性も違うわけです。その極端な例、BCG 仮説などがありますが、そのような個体の疫学的な特性も違います。感染力は平均的な接触数と 1 回の接触当たりの感染確率などに依存していますから、社会環境で異なる接触数が違うとずいぶん異なってくるでしょう。そのようなことがいろいろあります。

その意味で、あのような単純な数理モデルを、いきなりナショナル・レベルの人口に適用していかは、極めて疑問ではあるわけです。だから、3月、4月の時点で、そのようなモデルが出てきたときに、直ちに厳しい批判が多く出ました。そこで、SIR モデルとは異なるモデルで「もっと早く収束して、小さな山で済むでしょう」という予測を出してきた方もいます。だけれども、それも結局はあまり当たらなかった。

なぜかと言うと、非常に強く収束するような、山が小さく済むようなモデルは、やはり環境容量のようなものを先決しているわけです。「この辺りでいい」というようなことを考えている。あるいは、S I Rモデルでも、Nが小さいと思えば小さな山になってしまうのです。

例えば、第1波を再現するカーブは、S I Rでも作れるわけです。しかし、そうすると、最初のときのサイズNが、かなり小さかったということになります。それは、もしかすると正しいかもしれないが、なぜそれでいいか分かりません。「この程度だ」と決めて、そのようなカーブを作ることにはできると思うのですが、現実にはそういう法則性自体を生成するルールや有効なホスト人口サイズがどのようにきまるかがわからない。

だから、 β と感染率が下がったから、あのような山で済んだのか、元々、関係する人口が小さかったから小さかったのか、それは分かりません。両方とも可能性はあり得るのです。また、小さかった関係人口も、だんだん時間と共に大きくなっているかもしれないので、そのようなものも含め、いろいろなファクターがあると思っています。

それは、後半でお話したいことです。





- 感染症数理モデルは、「流行」という生命・社会現象を定性的に理解するためには有効であるし、生物学的プロセスの定量化にも効果がある。
- しかし、既存のデータ分析手段としてはともかく、大規模人口集団における流行ダイナミクスを**予測し、制御する手段**とみなした場合、その現象把握はどこまで有効だろうか？モデルと現実のギャップはどのように考えるべきか？
- 既存（過去）のデータから情報を引き出すための数理モデルの有効性は明らかだが、一方、予測方法としてみると期待外れの側面もある。これは経済学や社会学などの**社会科学系の数理モデル**に典型的に見られる状況。

この感染症数理モデルは、「流行」という生命・社会現象、この定性的な理解に対して非常に有効でした。生物学的な側面では、定量化にも効果があります。例えば 100 年前のスペイン・インフルエンザの時代は Kermack and McKendrick の前ですから、どのようなメカニズムで流行するかが分かっていなかったわけです。それに比べると、いろいろ不確実性はあっても、基本的なメカニズムが分かってきている、非線形相互作用だと分かっていることは、大きなメリットです。

一方で、先ほど申しましたように、予測や制御する手段として見ると、現象把握がどこまで有効かは、いろいろ問題があろうかと思えます。

実際に既に起こってしまった過去のデータから情報を引き出すための数理モデル、そのような側面での有効性は極めて明らかですが、予測として見ると非常にプアではないかということはあるのです。とはいえ、社会現象に関わる数理モデルは皆、そのような運命を持っていると思います。それは、経済学などのモデルも同じです。過去のデータには極めてよく合うけれども、それで予測して将来が合うかと言うと、それは、ほとんど期待できない。今のところの社会現象に関する数理モデルをフォアキャストとして使おうと思うと、そのようなものではないかと思えます。

① R_0 の理論と現実

- 各国において、観測されたCOVID-19症例の初期成長率から推定される R_0 は多様。2~6程度のばらつきがある。その結果大きな差になる。
- R_0 の値は生物学的パラメータと社会的パラメータからなり、かつモデル依存である。モデルがホスト人口の多様性をとらえているか？
- R_0 をきめてしまうと、多状態モデルによってホストの多様性を導入しても、ランダムな接触を仮定している以上、微分方程式モデルの推計結果はそれほどかわらないだろう。
- 大規模な感染人口の長期推計も、集団免疫閾値の推定もそう信頼できるものではない。仮定の含意をひきだしてみせた試算である。

始めに「基本再生産数がキー概念である」と言いましたが、ばらつきがかなり大きい、状況依存であるという現実があります。これも、やはり、生物学的なパラメータ部分は、かなりきちんとした決まった値だけれども、社会的なパラメータ部分が大きなばらつきを生み出すことがあるわけです。

一方で、ホストの免疫性や感受性の分布という意味で言うと、individual heterogeneity（個々の異質性）のような大きい要素を捉えられていないわけです。一旦、 R_0 を決めてしまうと、数理モデルを多少複雑にしても、挙動は、大体もう決まってしまうので、単純なモデルをそのまま信頼することはなかなか難しいと思っています。

それから、集団免疫に関しても盛んに議論されました。実は、「先ほどのモデルなどから出てくる60%などという集団免疫値よりも、ずっと低いところに集団免疫閾値があるのではないか」という議論が幾つか出てきています。

②人口の異質性の効果

- 多状態モデルによれば、臨界免疫化割合は、状態別に免疫化するのであれば様々であり得る。現実の流行過程での免疫獲得においては、**再流行不可能性**という意味での**集団免疫閾値**が、古典的な（一様性を仮定した）臨界免疫化割合から乖離する可能性はおおきい。
- もし全人口が一様に $\varepsilon \in (0,1)$ だけ免疫化されるのであれば、次世代行列 K のスペクトル半径は $(1-\varepsilon)$ 倍に変わるだけであるから、単状態の SIR モデルと同じ閾値 $1-1/R_0$ が得られる。
- もし第1状態のみを免疫化すれば、**次世代行列**は

$$K_\varepsilon = \begin{pmatrix} (1-\varepsilon)K_{11} & (1-\varepsilon)K_{12} \\ K_{21} & K_{22} \end{pmatrix}$$

となるから、その正固有値と R_0 の関係は単純ではない。

それはなぜかと言うと、要するに、先ほどは一様な人口を考えていたわけですが、感染率などで幾つものグループに分かれているような状況を考えたとしましょう。例えば、ここに行列が書いてあります。 2×2 の行列です。これは、人口を二つに分けて、1状態と2状態では再生産数が異なると仮定します。例えば、 K_{11} は1状態の人が1状態の人をどれだけ生み出すかという数字ですが、このようなものを並べて行列を作りますと、その行列の正固有値、スペクトル半径が、全体の基本再生産数になります。このような内部状態において、部分的に免疫化することを考えます。

例えば、この第1状態の人だけを免疫化すると、この第1行目だけに $1-\varepsilon$ がかかるという行列ができます。これは次世代行列 (next generation matrix) といって、ある世代の感染者が次の世代の感染者のベクトルをどのように作ったかということを示しています。このときは、閾値となる ε を求める式は、先ほどのような簡単な1次式になりません。固有値を計算すると $1-\varepsilon$ が入った2次方程式になってしまいます。 2×2 行列であれば具体的に計算できますが、 R_0 との関係は自明ではないわけで、そのように部分的に免疫化が進む場合には、全体の $1-1/R_0$ という、一様な場合の集団免疫閾値とは違う数字が、当然、出てくるわけです。

人口の異質性と集団免疫閾値

- COVID-19における獲得免疫による集団免疫化に関しては、古典的な集団免疫閾値よりもずっと低いところに閾値があるのではないかという議論がいくつかあらわれている。
- 一般に流行を途中で制御・終息させた場合、そこまでに感染した人口割合が免疫化されている。残された感受性人口集団において、第二波の流行が起きないような最小の免疫化割合が集団免疫閾値と考えられる。
- 感受性集団の異質性によって、高い感受性集団から選択的に流行が発生すれば、制御による流行中断によってより低い感受性集団が残されていくから、一様な集団よりも低い閾値が期待されることが想像できる。

例えば、非常に感受性の高いグループから低いグループまであって、高いグループで流行が始まったとします。ただし、感受性の高いグループは、全体の中での数は少ないとします。そうすると、その人たちはあっという間に皆かかってしまって、そこが免疫化されます。ところが、そのような集団が先にかかって脱落していくと、全体としての感受性は非常に下がってしまうのです。残された人は、それほど感受性は高くない人たちですから。そのように考えると、集団の免疫の閾値は、実は、もっとずっと低いことも十分にあり得るわけです。

実際に、エイズなどの性的感染症の場合、性的アクティビティーはひとによって非常に違うので、分布がベキ分布になるのですが、性的にアクティブな人はエピセンターになるわけですが、人数は非常に少ない。ロングテールの先の方の、そのようなところを例えばカットしてしまうと、全体としては、感染がそれほど広がらないことが十分考えられるということで、そのような異質性を考慮することで免疫閾値も変わってしまうわけです。

③緊急事態宣言の効果 (Kuniya 2020: SOEの後に β を一様に削減)

$$\begin{cases} \frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N}, \\ \frac{dE(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \varepsilon E(t), \\ \frac{dI(t)}{dt} = \varepsilon E(t) - \gamma I(t), \\ \frac{dR(t)}{dt} = \gamma I(t). \end{cases}$$

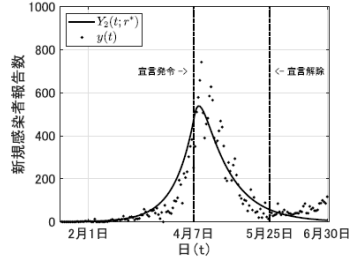
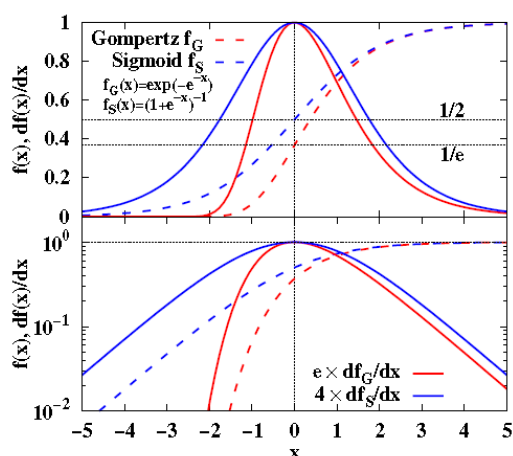


図 10.6: 2020 年 1 月 15 日から 6 月 30 日までの国内の COVID-19 の新規感染者報告数に対する流行曲線 $Y_2(t; r^*)$ (実線) と実際のデータ $y(t)$ (黒点)。4 月 7 日と 5 月 25 日がそれぞれ緊急事態宣言の発令日と解除日。

次の問題は、「緊急事態宣言 (SOE) は、うまくいったのか？」という議論についてです。これは國谷さんの研究で、SEIRモデルによって、実際の症例数に合うようにパラメータを変化させて、このようなエピカーブを作ったのです。それを見ると、このようなカーブを作るには、やはりある程度 β が削減される必要があります。つまり、緊急事態宣言によって、感染率は実際に下がったのではないか。感染率は、結局は 1 人当たりの平均的な出会いの数です。当時「8割削減」などとありましたけれども、「そのようなことは、ほぼ実現されたのではないか」という結論を出しております。

エピカーブの解釈は一意的ではない！

- 全人口を対象としたSEIRモデルによれば、第一波においてSOEによって β が低下したと考えれば、**接触率8割削減**は達成できたといえる。
- しかしながら、たとえば第一波に関係した人口は限られていた（例えば数万のオーダー）と考えれば、 β が低下せずとも、似たようなエピカーブを作ることにはできる。その場合、SOEの効果は関係する感受性人口の限定化・縮小という効果があったのかもしれないが、あるいはSOEがなくとも有効な感受性人口サイズは限定的であったかもしれない。
- パラメータ変動を仮定しないでエピカーブを再現しようとする、飽和型成長モデル（ロジスティック曲線やゴンベルツ曲線）による曲線当てはめは、民間の推計で色々試みられているが、飽和値の妥当性の根拠がみえない欠点があるが、そうした「感受性サイズの限定性」を考慮したものとも言える。



$$f_G(x) = \exp(-e^{-x}).$$

$$f_S(x) = \frac{1}{1 + e^{-x}} = \frac{1}{2} [1 + \tanh(x/2)]$$

しかし、先ほど言いましたように、そのようなカーブは全体のNを小さくしてしまえば、 β を変えなくても実現はできるのです。ですから、最初から関わった人数は少なかったのだと思うと、「それは β のせいではない」という理屈も言えなくはない。では、「なぜ、Nが小さかったのだ、Nは変わらないのか？」と言うと、そうではない。SOEがリフトアップされた後、また流行が始まったわけですから、やはり潜在的な R_0 は、それなりに大きいものがあって、関係する人たちも時間で一定なわけではありません。やはり、だんだん広がってきているというイメージではないかと思います。だから、これも、実際に何が起こったのかという解釈は、一義的ではない。 β の変化もあったであろうし、やはり関係人口の変化もあったのではないかと、そのように考えられるわけです。

最後に書いてあることは、そのようなNを小さくするよりは、ある程度のところでカーブをフィットさせるために、飽和値を決めてしまおうという人たちがいました。大阪の方で有名になったK値理論などもそうですけれども、ゴンペルツ・カーブという飽和曲線を累積症例数に当てはめています。

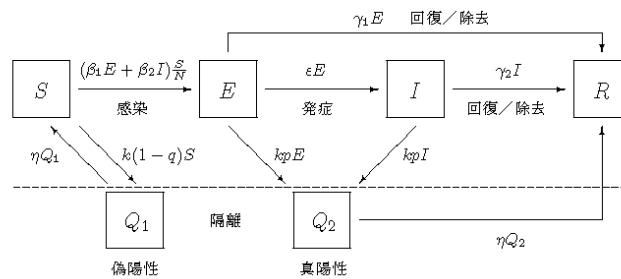
京大基礎物理研の人たちが面白い論文を書いていて、このグラフはそこから取って来ました。これは、ゴンペルツ・カーブとシグモイドを累積症例数のカーブに当てはめています。エピカーブは、このような一山の形に出てくるので、これが症例数の発生数に当てはまるようにパラメータを決めるわけです。もちろん、このようなことは、後からならばできます。結局、どこにピークが来ると決めることができれば、このような当てはめはできるのですが、それを「この症例数が立ち上がってきた最初の段階で、こんなことできるんですか？」と言うと、できない、全部外れています。

この問題点は、「では、どのようなメカニズムでゴンペルツ法則は発生するのですか？」という点です。シグモイドはSISモデルなどで再現できますが、要するに、力学系的なメカニズムが分からない中で、パラメータをある程度先決するようなやり方で予測できるかと言うと、

ごく短期的にはある程度合うと思いますけれども、やはり、それをそのまま信用するわけにはいかないと思います。そもそも、そのような飽和値自体が変わるでしょうし、それはなぜなのかも分かりません。曲線当てはめという考え方もあってよいとは思いますが、予測としてはプアだと思います。

④検査と隔離の効果

$$\begin{cases} \frac{dS(t)}{dt} = -(\beta_1 E(t) + \beta_2 I(t)) \frac{S(t)}{N} - k(1-q)S(t) + \eta Q_1(t), \\ \frac{dE(t)}{dt} = (\beta_1 E(t) + \beta_2 I(t)) \frac{S(t)}{N} - (\varepsilon + \gamma_1 + kp) E(t), \\ \frac{dI(t)}{dt} = \varepsilon E(t) - (\gamma_2 + kp) I(t), \\ \frac{dR(t)}{dt} = \gamma_1 E(t) + \gamma_2 I(t) + \eta Q_2(t), \\ \frac{dQ_1(t)}{dt} = k(1-q)S(t) - \eta Q_1(t), \\ \frac{dQ_2(t)}{dt} = kp(E(t) + I(t)) - \eta Q_2(t). \end{cases}$$



次は、検査や隔離をどのように考えるかです。これは、私と國谷さんがやった話から取っています。S E I Rモデルですが、潜伏期間クラスは asymptomatic であって部分的に感染性を持っていると仮定しています。感染者はS、E、I、Rと動いていくのですが、それとは別に Quarantine (隔離) クラスが Q 1、Q 2 とあります。

SにもEにも皆、一定の率で検査を行います。症状がなくても検査するわけです。今の日本では、症状がある者にある程度絞っていますけれども、諸外国ではユニバーサルに検査をやりまます。検査で陽性であれば隔離するというプロセスを、常時回すことにします。すると、SのところからQ1に隔離される人は、偽陽性で隔離されてしまった人です。Eの人で真陽性の人、やはりQ2のところ隔離されてきます。もちろん、症候性感染者の人(I)も同じです。S E I Rクラスに加えて Quarantine クラスのQ1、Q2のダイナミクスも、ここに付けてあって、さて、これで、どのようなことが言えますかということ。

実効再生産数 R_e の計算方法

- 線形化システム

$$\begin{aligned}\frac{dE(t)}{dt} &= \kappa(\beta_1 E(t) + \beta_2 I(t)) - (\varepsilon + \gamma_1 + kp) E(t), \\ \frac{dI(t)}{dt} &= \varepsilon E(t) - (\gamma_2 + kp) I(t)\end{aligned}\quad \kappa = \frac{\eta}{\eta + k(1 - q)}$$

- 係数行列の分解と次世代行列の構成

$$M = \begin{pmatrix} \kappa\beta_1 & \kappa\beta_2 \\ 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} -(\varepsilon + \gamma_1 + kp) & 0 \\ \varepsilon & -(\gamma_2 + kp) \end{pmatrix} \quad K = -MQ^{-1}$$

- 実効再生産数は次世代行列のスペクトル半径（正固有値）

$$R_e = \frac{\kappa\beta_1}{\varepsilon + \gamma_1 + kp} + \frac{\kappa\beta_2\varepsilon}{(\varepsilon + \gamma_1 + kp)(\gamma_2 + kp)}$$

先ほどのように、感染者がいないホスト集団のところに潜伏期と症候性感染が少数発生したときの方程式を書いてやります。今度は連立方程式ですから係数行列があるのですが、その係数行列は、この M と Q を足したものですが、この $M \times (-Q)^{-1}$ が、実は、次世代行列 K になります。 $-Q$ の逆行列は、結局のところ、各状態での滞在時間を計算しているのです。 M は、その状態での2次感染者の発生数ですから、要するに $M \times (-Q)^{-1}$ は滞在時間 \times 2次感染者発生数で、各状態から各状態の2次感染者が、どれだけ発生しましたかという数字が、 2×2 の行列で出てきます。その正固有値が基本再生産数です。ここでは隔離や検査などの介入をしていますので、自然状態での R_0 ではなく、実効再生産数 R_e になっています。その計算がこのように具体的にできます。

$$R_e = \frac{\kappa\beta_1}{\varepsilon + \gamma_1 + kp} + \frac{\kappa\beta_2\varepsilon}{(\varepsilon + \gamma_1 + kp)(\gamma_2 + kp)}$$

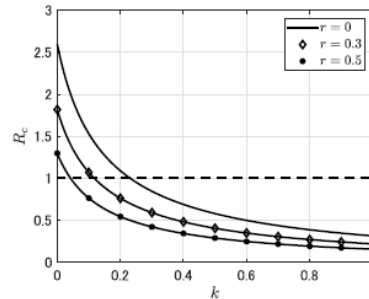


図 10.10: 異なる社会距離拡大政策 ($r = 0, 0.3, 0.5$) のもとでの検査隔離の検査率 k に対する実効再生産数 R_e 。

それで見てもみますと、横軸は検査率、縦軸は R_e にとったこのようなグラフが出てきます。これは、 r は社会的距離拡大政策による感染率の削減割合です。この形を見ると下に凸な曲線になっているのです。つまり、検査をやらなところから少しでも検査をやっていくと、かなりシャープに R_e が落ちてくるのです。だから、効果が割と大きいわけです。ちょうど R_e が 1 を切るところの水準の検査率も出ます。この検査率は、例えば、 $r = 0.3$ の場合は 10% くらいですが、まだかなり大きい値です。「毎日毎日、人口の 10% は検査しなさい」というのは、なかなかできませんが、ローカルな集団に対しては極めて効果的だと、よく分かります。

これも、まだまだ議論があるところですがけれども、経済学者の人たちは「国家レベルで検査を徹底的にやりなさい」と、そのようなことをよく言っています。しかし一方で、日本の感染症疫学系の人たちには普遍的な検査政策は非常に評判がよろしくなくて、今も、分科会などでも、そのような方向には動いていません。理由としては、偽陽性の問題などを非常に重く見ているようですけれども、中国やオーストラリアなどでの例を見ますと、偽陽性率は極めて小さいです。何十万、何百万とやっても、偽陽性がそれほど大きな問題にはなっていないようですから、その辺り真の理由は私も、ちょっとよく分からないですけれども、理論的には明らかに検査と隔離は効果があります。もちろん、この検査率があまり高いと実用的ではありませんが、社会的距離拡大政策をあわせておこなえば、必要な検査率はもっと低く済みますから、そのようなポリシー・ミックスを考えていけば、ある程度、実用的ではないかとは思っています。

⑤再感染による定在化の可能性(endemicity)

The origin of SIR modeling:
-2nd Kermack and McKendrick model-

- Kermack-McKendrick II (1933): Endemic model
→ virgin-sick-recovered model: recovered population can be reinfected (with partial immunity)

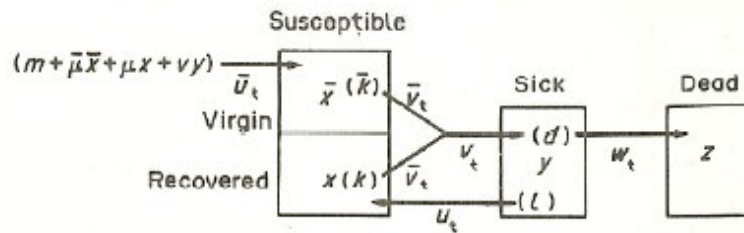


Figure 1.

最後に、再感染による定在化の可能性 (endemicity) です。歴史的に言いますと、Kermack and McKendrick の論文は、五つほどのシリーズ・ペーパーだったのです。最初に示したものは 1927 年の論文でしたけれども、30 年代の論文は、この endemicity の問題を取り扱っていたのです。

彼らが、どのようなモデルを考えたかと言うと、ここに論文から取った古い図があります。これは susceptible class が Virgin と Recovered の二つに分かれていて、Virgin は、そもそも 1 度もかかったことがない人、Recovered は、1 度かかったことがある人です。つまり、Recover しても感受性がまだある、つまり再感染するというモデルです。

ケルマック-マッケンドリックのエンデミックモデルは
偏微分方程式の境界値問題 (あるいは連立再生方程式)
(Inaba 2001, 2016)

The Kermack-McKendrick reinfection model is then formulated as

$$\begin{aligned} \frac{\partial s(t, \tau)}{\partial t} + \frac{\partial s(t, \tau)}{\partial \tau} &= -\mu s(t, \tau) - \lambda(t) s(t, \tau), \\ \frac{\partial i(t, \tau)}{\partial t} + \frac{\partial i(t, \tau)}{\partial \tau} &= -(\mu + \gamma(\tau)) i(t, \tau), \\ \frac{\partial r(t, \tau)}{\partial t} + \frac{\partial r(t, \tau)}{\partial \tau} &= -\mu r(t, \tau) - \theta(\tau) \lambda(t) r(t, \tau), \\ s(t, 0) &= m \int_0^\infty (s(t, \tau) + i(t, \tau) + r(t, \tau)) d\tau, \\ i(t, 0) &= \lambda(t) \int_0^\infty (s(t, \tau) + \theta(\tau) r(t, \tau)) d\tau, \\ r(t, 0) &= \int_0^\infty \gamma(\tau) i(t, \tau) d\tau, \quad \lambda(t) = \int_0^\infty \beta(\sigma) i(t, \sigma) d\sigma, \end{aligned}$$

with initial data

$$s(0, \tau) = s_0(\tau), \quad i(0, \tau) = i_0(\tau), \quad r(0, \tau) = r_0(\tau).$$

このモデルは、現代的に書けば、このような偏微分方程式の問題になります。なぜ偏微分かというと、それぞれの状態での滞在時間を全部考慮しているからです。つまり、一般の時間と各状態でのローカル・タイムが両方あるから偏微分方程式になっているのです。

再感染を許容する単純なSIRモデル

- Katrial のモデル
(後期ケルマック-マッケンドリックモデルの特殊なケース)

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

G. Katrial (2010), Epidemics with partial immunity to reinfection,
Math. Biosci. 228, 153-159.

それをもっと簡単にしたモデルがこれです。これは、イスラエルの Katrial という人が 10 年くらい前に出してきたモデルですが、Kermack and McKendrick の後期のモデルの特殊なケースに相当します。具体的には、滞在時間依存性を無視して考えたケースです。単純な S I R モデルと比べると、I の方程式が変わっています。R が感受性を持っているので、 $S + \sigma R$ が感受性のサイズになっているのです。 σ は R の S に対する相対的な感受性の度合いです。 σ が 1 より大きければ R は S よりも感受性が高い、1 より小さければ感受性が低いということで、一般には、1 度かかった場合の方が感受性は下がるでしょうから、 σ は 1 より小さいと考えるほうが自然かもしれませんが、いずれにしろそのようなモデルになっています。

この微分方程式をじっくりと見ていると次のようなことが分かります。まず、S という人は単調に減少していきます。ずっと将来の方を考えますと、外からの補充がないので、S はいなくなってしまう可能性があります。ところが、I と R だけを見ても、S がいなくなっても、I と R で流行は続きうることが分かります。このモデルでは、R は感受性があるので、I と R というサブシステムが自立して動けるわけです。だから、S という次元はいずれ消えてしまっても、今度は I R システムでエンデミックな定常解がありうる。

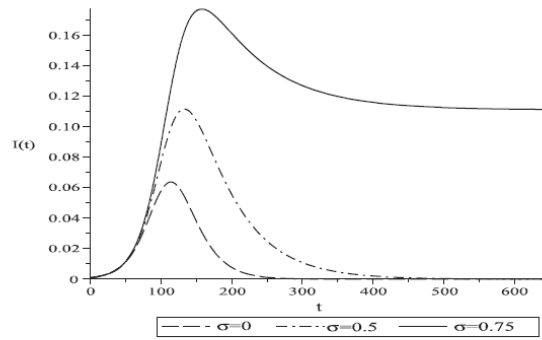


Fig. 1. Graphs $I(t)$ for $\gamma = 0.1, \beta = 0.15 (R_0 = 1.5)$, for $\sigma = 0, 0.5, 0.75, S(0) = 0.999, I(0) = 0.001, R(0) = 0$.

R_0 が再感染閾値 $1/\sigma$ を超えると、流行はエンデミックになる: $R_0=2.5$ ならば $\sigma=0.4$

これをグラフにしますと、この σ が0やある程度小さいときには、エピカーブは、このような山型です。再感染の相対的な割合が小さければ、古典的なS I Rと定性的に変わらないということです。だけれども、 $\sigma = 0.75$ の場合、一番上のグラフですが、もうIが消えていきません。つまり、定常状態に行ってしまいます。つまり、それは、I Rシステムの中にエンデミックな定常解があって、それが大域安定になってしまう。感染者がいつでも人口の中に定在しているという状態があるわけです。

このとき、 $\sigma \times R_0 = 1$ が閾値になっていて、 R_0 が $1/\sigma$ を超えると、このようなエンデミックな定常状態が出てきます。例えば、 R_0 が2.5であれば σ が0.4で、Rが元のSよりも4割ぐらいの感受性を持っているというところで閾値があって、それを超えてしまうと、もう病気がなくなってくれない、自然には消えない、ということになります。ですから、外側から新しい感受性人口の補充がなくても、流行がエンデミックになってしまうわけです。

- 再感染がおきると、感受性人口の補充がなされて、流行はエンデミックになる可能性がある。エンデミックな状態は、感染者と、既往感染者（回復者）の共存状態である。
- ワクチンは、感染状態から自然回復した状態にホスト人口を変化させるだけなので、 R_0 が再感染閾値を超えている場合は、流行を防ぐことはできない（ $R_0=2.5$ であれば、回復者が40パーセントの相対感受性を維持していればエンデミックになる）。
- 例えば、再感染の多くが無症候におわれれば、根絶は難しい。そこに人口学的な感受性補充がおきれば、症候性の再流行も予期できる。
- 罹患履歴の効果（個体のimmunity history）を取り入れた構造化モデルが望ましい。

このとき、仮にワクチンができたとしても、ワクチンの効果は自然回復よりは、恐らく大きくはないでしょうから、もしも基本再生産数が再感染閾値を超えているような場合には、ワクチンを使っても流行は止められないことも考えられるわけです。

更に言えば、COVID-19のように再感染が無症候に終わってしまうと、根絶は非常に難しいでしょう。もちろん、無症候で症状がないのであれば流行しても構わないではないかということはある。だけれども、基礎的な疾患を持っている人にとっては感染が致命的になるなどという、ロシアン・ルーレットのようなところもあるわけです。だから、ひそかに全体で広がっていった、脆弱な人が感染すると高い致死率となるということになると、流行してもいいとは言えないのではないかと思います。

要するに、個体のimmunityといいますか、そのようなことによって再感染したり、感染性が変わったり、そのような個体によって異なるhistoryを、それぞれが持っているのです。本来は、そのような罹患履歴の効果までモデルとして考えていく必要があると思います。しかし、immunity historyまでを取り入れたモデルはまだ理論的な段階で、更に、罹患履歴のようなことを測定する、定量化するといったことは、はるかに難しい問題かと思えます。

感染症数理モデルの実証・実装の問題

- 基本再生産数は感染症数理モデルの**定性的解析における決定的なキー概念**だが、対照的に、実測しようとする、モデル依存性ととも、**生物学的データと社会的データの境界に位置する**という意味で、特定するのが困難である（**定量的な不確定性**）。したがって、推定の変動幅は大きく、単純な数理モデルでの予測の信頼性は期待できない。
- 感染症数理モデルの実証的意義を向上させるためには、**ホスト個体群の社会的、生物学的異質性**をよく知る必要がある。
- 実験系を観察できるウイルス・細胞レベルへの適用は、ずっと信頼できるだろう。
- 社会問題として感染症流行をとらえた場合に必要な、**社会経済変数との相互作用**はほとんど考えられていなかった。

繰り返しになりますが、感染症数理モデルにおいては、基本再生産数が定性的な解析の決定的なキー概念で、それによって力学系的な挙動は明らかになるわけです。しかし一方で、それを実データに適用しようとする、 R_0 は捉え難い現実があるわけです。 R_0 に不確実性が大きいと、その結果も、定量的には不確定ということになります。ゆえに、感染規模の推定、予測などとなると、変動幅が非常に大きいというようなことになってしまうかと思えます。先ほど申しましたように、実証的な意義を向上させるためには、ホスト個体群の生物学的、社会学的な heterogeneity をよく知る必要があるかと思えます。かなり大変なことですが。

しかしながら、モデルとしてもっとうまくいっているケースもあって、今日はお話ししませんが、感染症の数理モデルは、細胞レベルで非常に活躍しています。細胞レベルでは何が起きているかと言うと、ウイルスが細胞に取り付いて、その中で自分を再生産して、それが更に細胞を壊して血中に出てきて、また次に正常細胞に取り付いて、というサイクルで増えているのです。それは、本当に感染症モデルと同じように定式化できます。このプロセスは、生物学的なプロセスですから、実際に観測してパラメータを取って、更に実験もできます。パラメータも安定的です。そうすると、例えば、「どこで、どのような薬を使うのがいいか」というような研究に非常に役に立つわけです。それは、ちょうど HIV のとき、体内のエイズ・モデルとして盛んに行われるようになって、今回の COVID-19 でも、そのようなことが行われています。

今は、細胞レベル・分子レベルで全部観測できますから、そのような実験系では極めていい理論的なフレームを与えていることは間違いないと思えます。かつては、モデルを作っても、それを実用化できないことがいくらでもあったのですけれども、今は、全部実験して観測することができますから、その点ではとても役に立っています。

それからもう一つ、今回のことで非常に問題になったことは、社会経済変数との相互作用です。ロックダウンなどをしたときに、疫学的な面からは有効かもしれないけれども、経済に大

きなダメージを与える。その結果、間接的に死亡率を上げてしまうこともあり得るわけです。そうすると、もう、単に流行の問題だけではない。だから、全体に、社会システムに、どのように影響を与えるかを考えなければいけなくなって、やはりそれは、感染症数理モデルの今までの考え方を大きく超えるところでした。本当は、そこをやっていなければいけなかったのですけれども、一今は経済学者が一生懸命やっておられると思いますけれども、一今まで、そのような射程を持ってモデルを考えていくことはなかったのです。もはや感染症流行だけの問題ではなく、社会システムへのダメージを、総体としてどこまで下げられるか、というオペレーションズ・リサーチのようなものが必要になってきているということかと思います。

総括

- 「感染」という現象は、生物個体群において、ミクロからマクロまで至る所に見られる普遍的な現象であり、その数学的理論は、非線形力学系として豊かな理論的構造をもっている。その構造を解明していくキー概念が「基本再生産数」であり、「再生方程式」である（人口学との共通性に注意）。
- 感染現象の主に生物学的プロセスの定量化や測定は大きく進歩したし、動的モデルの数理解析も発達したが、そのヒト集団への現実的適用においては困難がある。パラメータは生物的定数ではなく、社会集団のあり方によって大きく変化する。
- 個体群レベルの「流行」はヒトの意志の影響をもふくむ「社会現象」である。人口を構成する個体の多様性は、つねにモデルの射程を超えるものがある。国家レベルの流行制御は公衆衛生学的・疫学的問題をこえる社会経済問題である。

総括です。この「感染」という現象は、生物個体群において、細胞レベル、ミクロレベルからマクロまで至るところに見られる普遍的な現象で、それは、数学的な理屈としては、非線形力学系として非常に豊かな理論的な構造を持っています。それは、先ほど言った基本再生産数や再生方程式で表現されてきているわけです。また、ここに少し書きましたが、デモグラフィーとの関係も、強調はしておきたいと思います。

ところが、研究としては非常に進歩してきたわけですが、それを現実世界に適用しようとする、やはり、社会集団のあり方によってパラメータが大きく変化してしまうということで、社会の多様性といいますか、そのようなものを、きちんと捉えていかないといけません。ベンチマークや試算などは幾らでもできるわけですが、それをうまく受け止めることが必要です。何が分かっていないのか、分かっていない中で、ある程度の計算結果を、ある程度利用していく、そのようなリテラシーのようなものが、必要になってくると思います。そして、更に国レベルの流行制御は、もはや、狭い意味での疫学的な問題を越えた社会経済的な問題だと思います。総括は以上です。

～質疑応答～

質問 モデル上の人の動きは、線形にしか捉えきれない印象がある。一方で、現実では、緊急事態宣言により、摂動項として非線形にきいてくるかと思いますが、これは定量的に評価できるのでしょうか？

稲葉 おっしゃる通りです。今は、いろいろなアプリを使ったり、人の動きを直接観測するなどという技術が出てきましたから、そのような研究も、やがて、だんだん出てくるかな、という気はいたします。一方で、国によっては、なかなか、そのようなものは受け入れられないところもあるかもしれません。データを取ることに、社会、人間の情報を、どこまで与えていいかは、合意が得られるか、よく分かりませんが、そのような行動学的な研究は、非常に大事かと思えます。

質問 COVID-19 において感染症数理モデルは、どのように役立ったか。また、今後、どのように役立つと期待されるか

稲葉 これも賛否があります。3月、4月の段階で出たマクロなシミュレーションに対しては、強烈な反発がありました。実際、第1波が、それほど大きくならなかったものですから、微分方程式モデルで非常に大きな流行を予測したことに對して、大きな批判が出たことは確かだと思います。批判もよく分かるのですけれども、やはりモデルの性格がよく理解されていなかったことがあると思うのです。例えば「40万人死亡説」が出ましたけれども、専門家ほどあまり驚かなかった。「それは、あくまで、そのような単純な仮定の下に、そのような数字が出るのだな」ということで。ただ、そのような勢いのようなものは、 R_0 の中に含意されているわけですから、要するに、パラメータが持っている潜在的な力といえますか能力のようなものを示して見せる、そのようなものだったわけで、事実、それが実現されるかということ、専門家ほど、そうは思わなかったのではないかと思う。けれども、「それを、そのまま出してよかったか？」と言われると、やはり、そこはコミュニケーションの問題が何かあったと思わざるを得ないです。それに対して、いろいろな、もっと自己制約的な成長のモデルも出てきたのですけれども、やはり、それも、必ずしもうまくいかないこともあったかと思えます。

一方で、実効再生産数のようなデータを計算する、実態を明らかにする面では、圧倒的に力があつたことは認めないといけないと思います。予測のようなところは一番難しく、外れると全部だめかと言うと、そのようなことはなくて、細かい統計的な推定や、キャラクタライゼーションのところは、全部、数学があつてこそ初めてできたわけです。あれがなければ、全部当て推量になってしまうわけですから、そこは最大限評価すべきだし、今後も役立つことは間違いない。先ほど言ったミクロの面でも同じです。今、ミクロでも実験系でやっている人たちがいますが、そこで数理モデルは非常に有効に使われています。そこは認めてあげないといけないと思いますが、マクロな予測は、やはりなかなか難しいことは事実だと思います。

質問 インフルエンザとの違いは

稲葉 インフルエンザとの違いは、パラメータの違いも、もちろんありますけれども、インフルエンザは、ドリフトとシフトとあって、遺伝的な特性が毎年少しずつ変わってきているものです。それは、自動的に、治った人も次第に感受性を回復していくプロセスでもあるのです。COVID-19については、まだよく分からないですけれども、インフルエンザについては、

遺伝的な変異が定常的に起きていて、不定期に大きな変異がおきる、ということが最も大事で、そこに制圧の難しさもあることは言うておきたいと思います。COVID-19で同じようなことが起きるのかどうか、わたしにはわかりませんが、変異株の出現は報告されてきていますが、問題を複雑化することはたしかでしょう。

質問 同時に進行する複数の感染症ウイルスの変質モデル化

稲葉 もちろん、そのような多種共存系なども多く研究されています。ウイルスの変異もモデル化しています。ただ、進化は難しいです。ある意味では、進化は新しいファクターが出現することですから、それを予測するという事は、なかなか、まだ、うまくいっていないと思います。ただ多様性は無限ではないので、可能性をしばっていくことは可能かも知れません。

質問 感染症 R_0 、 R_t を推定する際の課題として、無症状、無症候の感染者の存在以外には何かあるか

稲葉 やはり、発症者をうまく捉えているか。無症状というのも非常に難しいです。本当に無症状なのか、単に統計的につかまっていけないだけなのか、ということもあります。無症候の問題は、実は非常に大きいです。なぜなら、先ほどの再生産数の計算方法も、あれはスカラー方程式で考えていますけれども、症候性と無症候性の感染者の連立方程式であれば、推定は上手くいくかどうかわかりません。それ以外の問題は、やはり個体の異質性をどの程度モデルの中に取り入れていけるか、ということかと思います。

質問 エージェント・ベース・モデルの数値シミュレーションによるアプローチは、どのくらいか

稲葉 あまりエージェント・ベースを知らないですけれども、やられていることは多くやられている。ただ、「課題は何か」と言われると、複雑なものを、そのままコンピュータで実装してしまうと、やはり、因果的理解が、なかなかできないということが、難しいところかと思います。実際の都市を再現して、いろいろな政策の効果を考えることは大事ですけれども、そのときの内的構造、因果連鎖というものが、われわれ観測者に追い切れないと、なかなか、それを正しく評価することは難しいかと思います。

質問 なぜ研究が遅れているか

稲葉 それは難しいですが、日本の医学部、公衆衛生などの今までの歴史が全て関わっているので、それからもちろん、数学の方の関わり方、応用数理のあり方、そのようなことが、いろいろ原因としては考えられるかと思います。一般的に言えば、日本における数学のとらえ方が非常に狭くて、純粋数学と物理・工学系のような伝統的応用数理以外の領域、社会現象や生命現象の研究において、数学が基本的な言語になりうるという感覚が乏しかったのではないのでしょうか。欧米に比べても応用数理の活躍する領域が日本では狭くとらえられているということはあると思います。そういう環境の下で、医学や疫学、公衆衛生領域における数学活用ということも後れを取ったという側面はあるかなと思います。より直接的には、90年代のエイズパンデミックのときに、世界的に感染症数理に関しては政策実装や実験系への応用が進んだのに、

そのチャンスを逃した、先見性がなかったということがあるかと思います。

稲葉 すみません、あまり超過してはいけないので、この辺で終わりにしたいと思います。どうもありがとうございます。

進行役 稲葉先生、たくさんの質問にお答えいただきまして、ありがとうございました。

それでは、セッションB-3、こちらで終了させていただきたいと思います。どうも、ご聴講ありがとうございました。