数列のめがね

対話 9

感染症流行の SIR 数列モデル（V）

矢崎成俊

登場人物：
須宇賀 空（すうが・そら），フリーの実験数楽者
堂仁加 菜怜（どうにか・なれ），高校 2 年生
堂仁加 素琉（どうにか・する），高校 1 年生

ここは日本のどこか．数楽で遊ぶ家．

堂仁加素琉（どうにかする）と姉の菜怜（なれ）が，こんにちは！ と元気に入ってきた．家主の須宇賀空（すうがそら）は，今日の話を次のように導入した．

「こんにちは．前回の最後に，シミュレーションの結果，約 7 割の人が感染して流行が終息したことがわかりました．その流行の結末を，初期時刻の段階で予測すること今日の目標です．今回が感染症流行の SIR 数列モデルの最終回になります」

SIR 数列モデル（おさらい）

次の SIR 数列モデルと条件を考えていた．

$$S_{n+1} = S_n - \sigma S_n I_n \tau \qquad (1)$$

$$I_{n+1} = I_n + \sigma S_n I_n \tau - \rho I_n \tau \qquad (2)$$

$$S_n + I_n + R_n = N \ \ (n = 0, 1, 2, \cdots) \qquad (3)$$

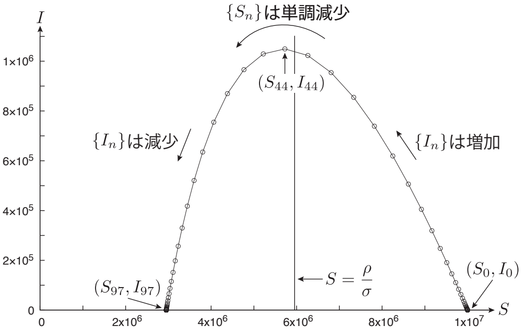
$$S_0 \gg I_0 > 0,\ R_0 = 0 \ \text{（初期条件）} \qquad (4)$$

$$0 < \rho\tau < 1 \ \text{（パラメータの条件 1）} \qquad (5)$$

$$0 < \sigma N \tau < 1 \ \text{（パラメータの条件 2）} \qquad (6)$$

ここで，N は封鎖人口で，感受性者と感染者のそれぞれのサイズ S_n と I_n がわかれば，回復者のサイズ R_n が (3) より算出できる．

下の図は，(1)(2) の解の組 (S_n, I_n) を，SI 座標平面上の点の座標として，$n = 0$ から順次，$I_n < 1$ となる $n = 97$ まで解いた点列 $\{(S_n, I_n)\}$ である．見やすいように点列○を線分でつなげた軌道が描いてある．

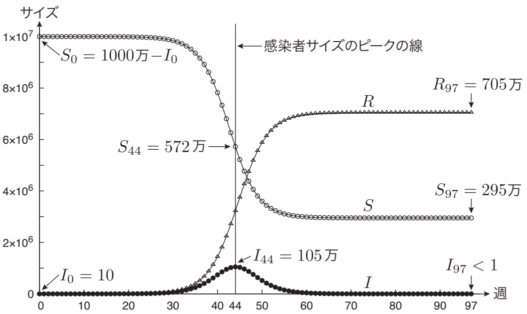


パラメータは以下で与えた．

$$N = 1000\text{万},\ S_0 = N - I_0,\ I_0 = 10,\ R_0 = 0$$

$$\rho = 0.5,\ \tau = 1,\ \sigma = \frac{70}{N} \times 1.2\% \qquad (7)$$

下の図は上の軌道に対応した，同じパラメータ (7) を用いて SIR 数列モデル (1)(2)(3) を数値計算した図である．$I_n < 1$ となった $n = 97$ までの解点列 $\{S_n\}$, $\{I_n\}$, $\{R_n\}$ を線分で繋いでプロットした．基本再生産数が $\mathscr{R}_0 = \dfrac{\sigma S_0}{\rho} = 1.68 > 1$ で，感染者は増殖し，ピークは第 44 週で $I_{44} = 105$ 万となった（$S_{44} = 572$ 万）．



「このシミュレーションでは，約 1000 万の S_0 から $S_{97} = 295$ 万に減りました．$I_{97} < 1$ となった第 97 週は実質最終週で，そのときまでに

感染した人たちのサイズの初期値 S_0 に対する相対的な規模は $\dfrac{S_0-S_{97}}{S_0}=70.5\%$ です．これは，およそ7割の人が感染して流行が終息したことを意味します．この7割という数値を，**SIR 数列モデルを解かないで，基本再生産数 $\mathscr{R}_0$ だけから求めること**が最終目標です．そのために数学的な道具を準備しましょう」と空がいった．

相加平均・相乗平均の不等式

「次の不等式はとても有用です」と空はいって，命題1を掲げた．

> **命題 1**　$a_1, a_2, \cdots, a_n$ を n 個の正数とする．このとき，次の不等式が成り立つ．
> $$a_1a_2\cdots a_n \leqq \left(\frac{a_1+a_2+\cdots+a_n}{n}\right)^n \tag{8}$$

「$n=2$ のとき，$\dfrac{a_1+a_2}{2}\geqq\sqrt{a_1a_2}$ という相加相乗平均の不等式を知ってます」と菜怜がいった．

「はい．有名ですよね．命題1の不等式はその一般化です．(8) を菜怜さんの不等式の形式に合わせると，

$$\frac{a_1+a_2+\cdots+a_n}{n}\geqq(a_1a_2\cdots a_n)^{\frac{1}{n}}$$

となります．左辺を**相加平均（算術平均）**，右辺を**相乗平均（幾何平均）**と呼びます．証明にチャレンジしましょう」と空はいって，証明を始めた．

命題 1 の証明）　$n=2$ のとき，(8) の右辺から左辺を引くと $\dfrac{(a_1-a_2)^2}{4}\geqq 0$ なので，不等式が成り立つ．

次に（$n=3$ を飛ばして）$n=4$ のとき，$n=2$ の結果を2回使って，

$$a_1a_2a_3a_4=(a_1a_2)(a_3a_4)\leqq\Bigg[\underbrace{\frac{a_1+a_2}{2}}_{A}\ \underbrace{\frac{a_3+a_4}{2}}_{B}\Bigg]^2$$

$$\leqq\left[\left(\frac{A+B}{2}\right)^2\right]^2=\left(\frac{a_1+a_2+a_3+a_4}{4}\right)^4$$

を得る．

この方針で進めると，$n=2,4,8,\cdots$ のとき，つまり $n=2^k\,(k=1,2,3,\cdots)$ のとき順次 (8) を示すことができる．

$n=4$ の結果で a_4 を A と書くと，

$$a_1a_2a_3A\leqq\left(\frac{a_1+a_2+a_3+A}{4}\right)^4$$

である．ここで，$A=\dfrac{a_1+a_2+a_3}{3}$ とおくと，

$$a_1a_2a_3\leqq\frac{1}{A}\left(\frac{3A+A}{4}\right)^4=A^3$$

となって，$n=3$ のとき (8) が示される．

一般に n が2のべき乗でないとき，$2^k<n<2^{k+1}$ を満たす k がただ一つ定まる（例．$n=3$ のとき $k=1$）．$N=2^{k+1}$，$m=N-n$ とおき，$b_1, b_2, \cdots, b_m$ を m 個の任意の正数とする．このとき，N に対して (8) は成り立つから，

$$a_1a_2\cdots a_nb_1b_2\cdots b_m$$
$$\leqq\left(\frac{a_1+a_2+\cdots+a_n+b_1+b_2+\cdots+b_m}{N}\right)^N$$

ここで，$A=\dfrac{a_1+a_2+\cdots+a_n}{n}$，$b_1=b_2=\cdots=b_m=A$ とおくと，$N=n+m$ だから，

$$a_1a_2\cdots a_n\leqq\frac{1}{A^m}\left(\frac{nA+mA}{N}\right)^N=A^n$$

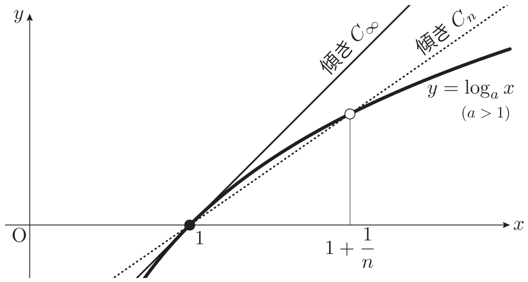
となって，n が2のべき乗でないときも (8) が示される．■

「こんな証明は思いつきません」と菜怜がいった．素琉も同調してお手上げポーズをしている．

「$n=4$ のときはどうにか思いつくとしても，$n=3$ のときに $n=4$ の結果を使うなんて，なかなかオシャレですよね．これはコーシーによる証明で，相加平均・相乗平均の不等式の証明の中で最古といわれています」と空はこたえた．

桁数と対数関数の増加の様子

「話は変わって，底が a の対数関数 $y=\log_a x$ について考えます．$a>1$ のとき，グラフは次のようになります」と空はいって図をみせた．



対数関数 $y=\log_a x$ は，与えられた x に対して $x=a^y$ を満たす y は何？ という関係を表したものだ．底が $a=10$ の対数関数，つまり常用対数の例で考えてみると，$x=10000$ ならば $y=4$ で，1 を足せば x の桁数だから，常用対数は桁数引く 1 を表している．x が実数の場合にも同様のことがいえる．実際，x が整数部分が k 桁の正数だったら $10^{k-1}\leqq x<10^k$ であるから，辺々の常用対数をとると，$k-1\leqq\log_{10}x<k$ となって，$[\log_{10}x]=k-1$ を得る（$[a]$ はガウス記号で a を超えない最大の整数）．$x=a^y$ の両辺の常用対数をとると，$\log_{10}x=y\log_{10}a$ より，$b=\dfrac{1}{\log_{10}a}$ とおくと，$y=b\log_{10}x$ となる．

「大雑把にいうと，$y=\log_a x$ は $[x]$ の桁数引く 1 を何倍かしたものだから，増え方が緩慢です．$[x]$ の桁が 1 つ上がるごとに y の値が b だけあがるのですから」と空がいった．

「緩慢ってこういうことですか」と菜怜はいって，ノートに整理した表をみせた．

$[x]$ の桁が 1 つ増える		y は b 増加
$x:10\to100$	$(+90)$	$y:b\to2b$
$x:100\to1000$	$(+900)$	$y:2b\to3b$
$x:1000\to10000$	$(+9000)$	$y:3b\to4b$

「まさしくそうです．x が大きくなればなるほど，1 桁あげるのが大変になって，y を b だけ増やすのに時間がかかるから，$y=\log_a x$ は単調増加なのだけど，グラフがだんだん横に寝ていくのです」と空は曲線をなぞりながらいった．

自然対数

「$a=10$ の他にもう一つとても重要な底があります．それを導出しましょう」と空は続けた．

先の図に●と○がある．●は $y=\log_a x$ の x 切片で $(x,y)=(1,0)$ の点，○は $y=\log_a x$ と直線 $x=1+\dfrac{1}{n}$ の交点である．●と○を通る直線は，曲線 $y=\log_a x$ を貫く．このように曲線上の 2 点を通る直線を**割線**という．この割線の傾きを C_n とすると，C_n は次のように表される．

$$C_n=\frac{\log_a\left(1+\dfrac{1}{n}\right)}{1/n}=\log_a\left(1+\frac{1}{n}\right)^n \tag{9}$$

○を動かして●に限りなく近づけることを考える．この○の場合，$n\to\infty$ の極限がそれに相当する．一般に○→●の極限をとり，割線がある直線に収束したとき，その直線を**接線**と呼ぶ．この接線の傾きを C_∞ とすると，

$$e_n=\left(1+\frac{1}{n}\right)^n$$

とおいて，

$$C_\infty=\lim_{n\to\infty}\log_a e_n=\log_a\left(\lim_{n\to\infty}e_n\right)$$

となる．だから，接線の傾き C_∞ を求めるには数列 $\{e_n\}$ の極限値を求める必要がある．

「$\{e_n\}$ は収束するのですか？」と菜怜が聞いた．

「とても鋭い質問です．次の命題を覚えていますか？」と空は 2 人に聞いた．

> **命題 2**　単調増加な数列 $\{a_n\}$ のすべての項がある定数 C 以下であるとする．このとき，$\{a_n\}$ の $n\to\infty$ のときの極限が存在する．

「覚えています．だから，数列 $\{e_n\}$ は単調増加で，すべての項がある定数以下であることを示すのですね」と素琉が嬉しそうにこたえた．

「いいですね！ まず，相加平均・相乗平均の不等式（命題 1）を使って，$e_n\leqq e_{n+1}$ を示しましょう」と空がいった．

$e_n\leqq e_{n+1}$ の証明）$A=1+\dfrac{1}{n}$ とし，

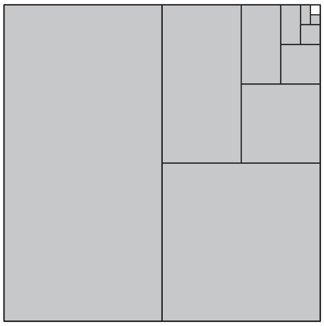
$$a_1=a_2=\cdots=a_n=A,\ a_{n+1}=1$$

とおいて，命題1の不等式 (8) の n を $n+1$ の場合に適用すると，$\{e_n\}$ の単調増加性

$$a_1 a_2 \cdots a_{n+1} \leqq \left(\frac{nA+1}{n+1}\right)^{n+1} = e_{n+1}$$

を得る． ■

「この半分半分の和の図を覚えていますか？」と空が2人に聞いた．



2人とも頷いたので,「この図と2項定理

$$(a+b)^n = \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} a^{n-k} b^k \qquad (11)$$

を使って，$\{e_n\}$ のすべての項が3未満になることを示しましょう」と空はいった．

$e_n < 3$ **の証明**）（11）で $a=1$, $b=\dfrac{1}{n}$ とすると，

$$\begin{aligned} e_n &= \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} \frac{1}{n^k} \\ &= 2 + \sum_{k=2}^{n} \frac{1}{k!} \underbrace{\frac{n}{n} \frac{n-1}{n} \cdots \frac{n-(k-1)}{n}}_{k\text{項}} \\ &< 2 + \sum_{k=2}^{n} \frac{1}{k!} \end{aligned}$$

がいえる．ここで，$k \geqq 2$ のとき，

$$\begin{aligned} \frac{1}{k!} &= \frac{1}{\underbrace{k(k-1)(k-2)\cdots 3\cdot 2}_{(k-1)\text{項で全部2以上}}\cdot 1} \\ &\leqq \frac{1}{\underbrace{2\cdot 2\cdot \cdots \cdot 2}_{(k-1)\text{項}}\cdot 1} = \frac{1}{2^{k-1}} \end{aligned}$$

が成り立つ．これより，半分半分の和の図から $\displaystyle\sum_{k=2}^{n} \frac{1}{2^{k-1}}$ は1未満であることを使って，

$$e_n < 2 + \sum_{k=2}^{n} \frac{1}{k!} \leqq 2 + \sum_{k=2}^{n} \frac{1}{2^{k-1}} < 3$$

を得る． ■

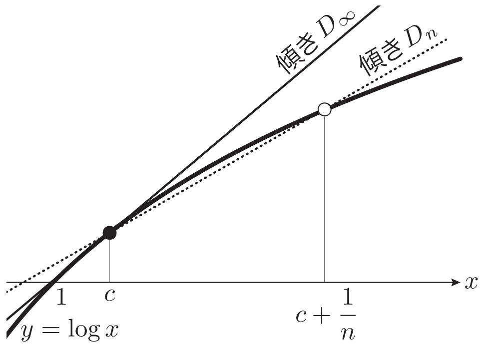
こうして，命題2から数列 $\{e_n\}$ の $n \to \infty$ のときの極限が存在することがわかった．極限値は，無理数2.7182818284590…であることがわかっていて，これを記号 e で表す．すなわち，

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

である．これより，接線の傾き (10) が，$C_\infty = \log_a e$ となることがわかった．特に，$C_\infty = 1$ となる底 $a=e$ のときの対数関数 $\log_e x$ を**自然対数** (natural logarithm) といい，底 e を省略して単に $\log x$ と表す．

「何が自然なのですか？」と素琉が聞いた．

「常用対数や他の底の対数よりも，底 e の対数が数学的に自然という意味でしょう．例えば，角度は度数法よりも弧度法の方が数学的には自然である，といったならば，底 e の対数の自然さはこれと同じで，数学的に便利とか有難いという気持ちも含まれていると考えています」と空が持論でこたえ，次の図の自然対数 $y=\log x$ の●における接線の傾き D_∞ を求めると自然さが実感できるかもしれません，と続けた．



この図は $c>1$ ときの概念図になっているが，以下は一般に $c>1$ で成り立つ話である．図の●は $y=\log x$ と直線 $x=c$ の交点，○は $y=\log x$ と直線 $x=c+\dfrac{1}{n}$ の交点である．●

と○を通る割線の傾きを D_n とすると，

$$D_n=\frac{\log\left(c+\frac{1}{n}\right)-\log c}{1/n}=\log\left(1+\frac{1}{nc}\right)^n$$

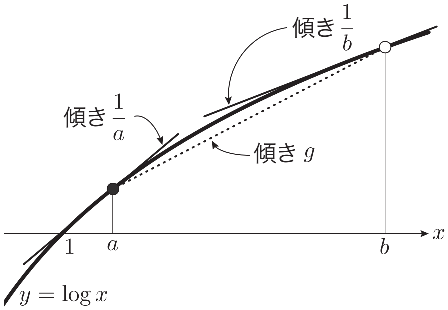
である．ここで，$t=nc$ とおくと，$n=t\times\frac{1}{c}$ であるから，$D_n=\frac{1}{c}\log\left(1+\frac{1}{t}\right)^t$ となる．$n\to\infty$ のとき，$t\to\infty$ で，$\left(1+\frac{1}{t}\right)^t\to e$ がわかるので，自然対数 $y=\log x$ の点 $(c,\log c)$ における接線の傾き $D_\infty=\frac{1}{c}$ を得る．

「$a\neq e$ として，もし底が a の対数だったら，接線の傾きは $\frac{\log_a e}{c}$ となって，いつも $\log_a e$ という係数が登場して煩わしい．この煩わしさは $a=e$ のときのみ解消されます」と空がいった．

「係数が 1 になるという意味で自然なのですね」と菜怜が感心した．

「そう考えてよいです．さて，自然対数の話の最後に，ある不等式を紹介します．この不等式を用いて，SIR 数列モデルの感受性者数の極限サイズが $S_\infty>0$ となることを示します」

空はそういって，次の図をみせた．



この図は $1<a<b$ ときの概念図になっているが，以下は一般に $0<a<b$ で成り立つ話である．図の●は $y=\log x$ と直線 $x=a$ の交点，○は $y=\log x$ と直線 $x=b$ の交点で，それぞれの点における接線の傾きは $\frac{1}{a}$ と $\frac{1}{b}$ である．このとき，●と○を通る割線の傾きを g とすると，次の不等式が成り立つ．

$$g=\frac{\log b-\log a}{b-a},\ \frac{1}{b}<g<\frac{1}{a} \qquad (12)$$

「もっと極端に，a を 1 よりもどんどん小さくして 0 に近づけていくと，●における接線は立っていきます．一方，b をどんどん大きくしていくと，○における接線は寝そべっていきます．だから，両端点●○を結んだ右肩上がりの割線の傾きは，$\frac{1}{a}$ と $\frac{1}{b}$ の間にあるということです」と空が補足した．

「そういわれると納得です」と 2 人はいった．

最終規模の見積り

「いよいよフィナーレです」と空はいった．第 n 週までに感染した人たちのサイズの S_0 に対する相対的な大きさ（第 n 週の規模）を

$$F_n=\frac{S_0-S_n}{S_0}$$

とおく．最終的な感染者割合，すなわち**最終規模**

$$F_\infty=\frac{S_0-S_n}{S_0}$$

を $\mathscr{R}_0$ だけから求めることが目標である．各 n について，$0<S_n<S_0$ だから $0<F_n<1$ である．一方，$0\leqq S_\infty<S_0$ だから $0<F_\infty\leqq 1$ であり，$S_\infty>0$ と $F_\infty<1$ は同値である．

まず，(1) (2) を辺々足して，S_n, I_n を左辺に移項し，$n=j-1$ を代入すると，

$$S_j-S_{j-1}+I_j-I_{j-1}=-\rho I_{j-1}\tau$$

を得る．辺々を $j=1,2,\cdots,n$ について足すと，

$$S_n-S_0+I_n-I_0=-\rho\sum_{j=1}^{n}I_{j-1}\tau \qquad (13)$$

となる．

次に，(1) で $n=j-1$ とおくと，

$$I_{j-1}=\frac{S_{j-1}-S_j}{\sigma S_{j-1}\tau}$$

がわかる．これを (13) の右辺の和の中に代入し，両辺を $-S_0$ で割ると，

$$A_n = \sum_{j=1}^{n} \frac{S_{j-1} - S_j}{S_{j-1}}$$

とおいて，

$$\frac{S_0 - S_n}{S_0} + \frac{I_0 - I_n}{S_0} = \frac{\rho}{\sigma S_0} A_n$$

を得る．よって，基本再生産数 $\mathscr{R}_0 = \dfrac{\sigma S_0}{\rho}$ を用いて整理すると，

$$\mathscr{R}_0 \left(F_n + \frac{I_0 - I_n}{S_0} \right) = A_n$$

となる．左辺は $n \to \infty$ のとき $\mathscr{R}_0 \left(F_\infty + \dfrac{I_0}{S_0} \right)$ に収束するので，右辺 A_n も同じ値に収束し，

$$A_\infty = \mathscr{R}_0 \left(F_\infty + \frac{I_0}{S_0} \right) \tag{14}$$

和 A_n の意味を理解するために，(12) の不等式の辺々に $b-a$ をかけた不等式

$$\frac{b-a}{b} < \log \frac{b}{a} < \frac{b-a}{a} \tag{15}$$

を用いる $(0 < a < b)$．

数列 $\{S_n\}$ は単調減少で，すべての n について $S_n > 0$ なので，$a = S_j$, $b = S_{j-1}$ とおいて，A_n の $\sum$ の中身に (15) の左の不等式を適用して，

$$A_n < \sum_{j=1}^{n} \log \frac{S_{j-1}}{S_j} = \log \frac{S_0}{S_n} = \log \frac{1}{1-F_n} \tag{16}$$

がわかる．同様に (15) の右の不等式を適用して，

$$\log \frac{1}{1-F_n} < \sum_{j=1}^{n} \frac{S_{j-1} - S_j}{S_{j-1}} \, \frac{S_{j-1}}{S_j} \tag{17}$$

を得る．

(1) で $n = j-1$ とおき，両辺を S_{j-1} で割ると，$I_{\max}$ を感染者のサイズの最大値として，

$$\frac{S_{j-1}}{S_j} = \frac{1}{1 - \sigma I_{j-1} \tau} < \frac{1}{1 - \sigma I_{\max} \tau}$$

となる．よって，$I_{\max} < N$ と (17) から，

$$\log \frac{1}{1-F_n} < \frac{A_n}{1 - \sigma I_{\max} \tau} < \frac{A_n}{1 - \sigma N \tau}$$

を得る．$n \to \infty$ のとき，(14) と $F_\infty \leqq 1$ から，

$$\log \frac{1}{1-F_\infty} < \frac{A_\infty}{1 - \sigma N \tau} \leqq \frac{\mathscr{R}_0}{1 - \sigma - N \tau} \left(1 + \frac{I_0}{S_0} \right)$$

となる．ここで，もし $F_\infty = 1$ だったら，最左辺は無限大に発散するが，最右辺はパラメータと初期値だけで決まる定数であるので，$F_\infty < 1$ でなければならない．よって，$S_\infty > 0$ を得る．

(16) で $n \to \infty$ とし，上の結果と合わせると，

$$A_\infty < \log \frac{1}{1-F_\infty} < \frac{A_\infty}{1 - \sigma I_{\max} \tau} \tag{18}$$

を得る．パラメータを (7) とした場合の結果から，$I_{\max}$ は第 44 週で達成され $I_{44} = 105$ 万だった．この例だと，$\sigma I_{\max} \tau = 0.09$ 程度であるので，大雑把に $1 - \sigma I_{\max} \tau \approx 1$ として，不等式 (18) を $\log \dfrac{1}{1-F_\infty} = A_\infty$ と近似すると，最終規模 F_∞ の満たす方程式

$$\log \frac{1}{1-F_\infty} = \mathscr{R}_0 \left(F_\infty + \frac{I_0}{S_0} \right) \tag{19}$$

を得る．さらに，初期条件 (4) のもとで $\dfrac{I_0}{S_0} \approx 0$ と仮定すれば，(19) からもっと簡単な近似式

$$\log \frac{1}{1-F_\infty} = \mathscr{R}_0 F_\infty \tag{20}$$

を得る．(19) や (20) を**最終規模方程式**と呼ぶ．特に，(20) は基本再生産数 $\mathscr{R}_0$ だけがわかれば最終規模がわかることを意味している．

上述したように，パラメータ (7) のもとで，基本再生産数は $\mathscr{R}_0 = 1.68$ で，実質最終週の第 97 週の規模は $F_{97} = 70.5\%$ だった．同じパラメータで (19) や (20) を F_∞ について解くと，いずれも最終規模 $F_\infty = 68.2\%$ を得る．これは SIR 数列モデルを解かなくても，初期値とパラメータだけから，特に $I_0 \ll S_0$ の場合は基本再生産数だけから，およそ 7 割の人が感染して流行が終息することがわかることを意味している．

「盛り沢山でヤバイです．でも最初の値だけから，結末の感染者の割合がわかるということですよね」と素琉がまとめたら，菜怜も「凄みを感じました」といった．「一度に全部わからなくてもよいですから．じっくりと」と空はこたえて，今日の対話を終了した．

（やざき　しげとし／明治大学）