

エビデミックに於ける age distribution と ポリオの予測について

崎 野 滋 樹

(1963 年 6 月 受付)

On the Age Distribution of the Epidemic Model and the Prediction of Poliomyelitis.

Sigeki SAKINO

Abstract

Let R_t be the number of susceptibles and $d_x S(x, t)$ be the number of infectives in age $(x, x+dx)$ at the time t . Putting $f(x, t)dx = E_t\{d_x S(x, t)\}$, $g(x, t)dx = E_t\{R_t d_x S(x, t)\}$, we derived the differential and integral equations on $f(x, t)$ and $g(x, t)$. And we gave the particular solutions of these equations. And, using the deterministic model of the epidemic, we tried the prediction of the number of infectives of poliomyelitis which were epidemic in Hokkaido in 1960.

Institute of Statistical Mathematics

1. 序 文

[1] では probability generating functional の性質を用いて、エビデミックにおける各 age の同時分布を求める問題を考えた。ここでは少し観点を变えて、各時点 t で、age が $(x, x+dx)$ にある病人数 $d_x S(x, t)$ がポアソンであると仮定してエビデミックにおける age distribution を求める問題を考えた。このことは次節で詳しく述べる。また、3 節では、上の結果を伝染期間が非常に長い場合に適用し、4 節では 3 節のデターミニスチック・モデルを用いて 1960 年北海道に大流行したポリオの各月における病人数を予測した。

2. エビデミックにおける age distribution について

確率変数 R_t , $S(x, t)$ は時刻 t における感受性者数、age $X \leq x$ なる病人数を、また確率変数 $d_x S(x, t)$ は $x < \text{age } X \leq x+dx$ なる病人数をあらわすものとする (ただし、ここで述べる age とは感染からの経過時間をあらわす)。そして、 $f(x, t)dx$ を $d_x S(x, t)$ の期待値、即ち

$$f(x, t)dx = E_t\{d_x S(x, t)\} \quad (1)$$

をあらわすものとする。ただし、 E_t は時刻 t における感受性者数 R_t , 病人数 $d_x S(x, t)$ に関する期待値をあらわす。そこで、時刻 t で age が $(x, x+dx)$ にある病人数 n がポアソン分布

$$\begin{aligned} P_1 &= f(x, t)dx + o(dx) \\ P_0 &= 1 - f(x, t)dx + o(dx) \\ P_n &= o(dx) \quad n > 1 \end{aligned}$$

に従うとしよう。

さらに; $g(x, t)$ がまた

$$E_t\{R_t d_x S(x, t)\} = g(x, t) dx \quad (2)$$

なる性質をもち、上と同じようにポアソンに従うものとする。

いま、 $\lambda(x)$ を age x なる1人の病人が1人の感受性者に単位時間 $(t, t+1)$ で感染させる確率とすれば、病人数 $d_x S(x, t)$ による単位時間における新患者の発生数は $\lambda(x) R_t d_x S(x, t)$ で与えられる。従って、(2) から平均新患者数は

$$E_t\{\lambda(x) R_t d_x S(x, t)\} = \lambda(x) g(x, t) dx \quad (3)$$

で与えられる。次に、回復についても上と同じように、age x の1人の病人が時間 $(t, t+dt)$ で回復する確率を $\mu(x) dt$ とすると、(1), (2), (3) の関係から

$$f(x, t) = \int_0^\infty f(u) g(u, t-x) e^{-\int_0^x \mu(v) dv} du \quad (4)$$

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial t} = -\mu(x) f(x, t) \quad (5)$$

なる関係式を容易に導くことができる。偏微分方程式 (5) の一般解は

$$A(t-x) e^{-\int_0^x \mu(v) dv} \quad (6)$$

で与えられる。ただし、 $A(t-x)$ は整函数をあらわす。従って、(4), (6) から

$$A(t-x) = \int_0^\infty \lambda(u) g(u, t-x) du \quad (7)$$

が得られる。ところで、(7) の特解は $\lim_{u \rightarrow \infty} \int_0^u \lambda(g) dv = \infty$ を仮定すると、

$$g(x, t) = A(t) e^{-\int_0^x \lambda(u) du} \quad (8)$$

なることがわかる。この $g(x, t)$ から感受性者数 R_t と病人数 $d_x S(x, t)$ の共分散 $g_1(x, t)$ は

$$g_1(x, t) = A(t) e^{-\int_0^x \lambda(u) du} - \left(R_0 - \int_0^t dv \int_0^t f(u, v) du \right) f(x, t) \quad (9)$$

で与えられる。ただし、 R_0 は $t=0$ における感受性者数をあらわす。

3. 伝染病患者数がロジスチックで発展するときの age distribution

$t=0$ で age 0 の病人1人と感受性者が n 人いたとする。 $f(x, t) dx$, $g(x, t) dx$ は何れも前節において定義されたごとくである。この節では、感染からの病気の持続時間が充分長い場合について、age distribution, $g(x, t) dx = E_t\{R_t d_x S(x, t)\}$ について述べることにしよう。

いま、age x の1人の病人が他の1人の感受性者に単位時間 $(t, t+1)$ で感染させる確率を $\lambda(x)$ に等しいとし、かつ時刻 t における感受性者数を $R_t = n - S(t, t)$ としよう。つまり、これは $\mu(x) = 0$ 即ち、一度感染した病人は回復したいと仮定したことになる。この仮定は伝染期間が非常に長い伝染病に対しては許される。従って時間 $(t, t+dt)$ における新患者の発生数の期待値は

$$\begin{aligned} \lambda(x) E_t\{R_t d_x S(x, t)\} &= \lambda(x) E_t\{(n - S(t, t)) d_x S(x, t)\} \\ &= \lambda(x) g(x, t) dx \end{aligned} \quad (10)$$

で与えられる。従って、(10) から、 $g(x, t)$ は

$$\begin{aligned} g(x, t) &= n f(x, t) - E_t\{d_u S(u, t) d_x S(x, t)\} \\ &= n f(x, t) - \int_0^t \gamma(u, x, t) du \end{aligned} \quad (11)$$

を満足する。ただし、 $\gamma(u, x; t) = \gamma(x, u; t) = E_t\{d_u S(x, t) d_x S(x, t)\}$ である。ところで、 $\gamma(u, x; t)$ は論文 [1] の結果から、

$$\frac{\partial \gamma}{\partial t} + \frac{\partial \gamma}{\partial x} + \frac{\partial \gamma}{\partial u} = 0 \quad (12)$$

を満足することがわかる。この一般解は

$$B(t-u, t-x) \text{ (ただし, } B \text{ は整函数をあらわす)}$$

で与えられる。また, (4), (5) なる関係から

$$A(t)e^{-\int_0^x \lambda(u)du} = nA(t-x) - \int_0^t B(t-u, t-x)du \quad (13)$$

を導くことができる。即ち (11) あるいは (13) 式を満足するように $A(t)$, $B(t-x, t-y)$ を決めれば, age distribution $f(x, t)$ や age x, y の共分散 $E_t\{d_x S(x, t)d_y S(y, t)\} = r(x, y, t)dxdy$ を決めることができる。そこで, 境界条件として $r(0, x; t) = r(x, 0; t)$ なるように A, B を決めてやればよい。が併し, 整函数の位数についての知識が現象から与えられれば問題は容易に解決される。

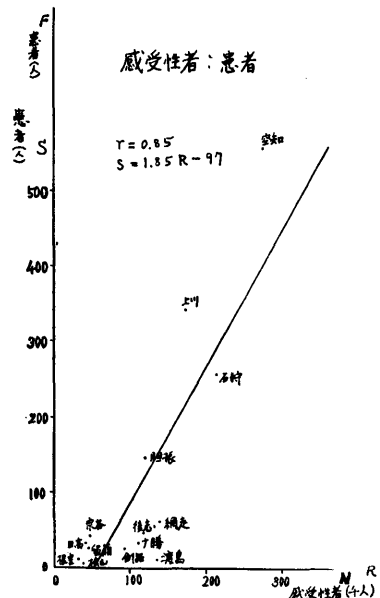
以上, エビデミックにおける時刻 t の感受性者数 R_t と病人数 $d_x S(x, t)$ の積の期待値と $d_x S(x, t)$ の期待値の関係について述べた。が併し, これを現実に適用するには, なおさらに解析的な問題の解決乃至生物学的現象の追求が必要であるし, またこのような観測資料を得ることは現在の段階ではなかなか容易なことではない。そのようなわけで以下に述べるポリオの予測においても, age distribution を考えることはできなかつた。

4. ポリオの伝播の予測について

1960 年北海道に大流行したポリオの伝播資料の始めの一部を用いて, 伝播の予測を試みるのであるが, 確かに, 得られた観測資料は診断等の誤差について多くの問題を含んでいる。これについては後で詳しく述べることにしよう。これまで, 多くの人達がエビデミック・モデルについて論議して来たが, これらのモデルの効用並びにその面白味は何といてもエビデミックのピークの時点や患者数の予測にある。勿論, 今日のようにポリオの予防法が完成された今日では, ポリオの予測そのものはそれほど大きな意味をもたない。が併し, 一般に Hamer (1906) 以来発展したモデルが伝染病の伝播の予測にかなり効力を持つものであるであろうことは以下の解析から容易に察することができる。もし, ポリオの age distribution の観測資料が得られるならば, エビデミック曲線の上昇, 下降の意味が容易に把握されるであろうが, 現実には技術的な困難が伴うようである。また生物学的要求がそこまで到達していないというのが現状ではあるまいか。

以下に於いて, 乳幼時, 学童期にあるポリオの患者の予測を試みるのであるが, これは次のような理由による。即ち, 1960 年 11 月末におけるその年の北海道全域の総り患数は 1548 名であつて, その中約 98% が学童期並にそれ以下の乳幼時期のものに発生している。このような理由から, よくいわれるように, 学童期並にそれ以下のもの対策が特に重要であるように思われる。そこで, これらの総り患数を予測するためには, できるだけ homogeneous と考えられる地域に北海道を幾つかに分けて, その各々の中で予測を試みねばならない。そのような理由から, 支庁別に予測を試みることにした。

先ず, 11 月末における支庁別総り患数 S とその年の 1 月始めの 13 歳以下の感受性者数 R との関係を見たところ, 第 1 図に示されたごとく, 相関係数は 0.85 であつて, かなり高い値を示している。特に, 感受性者数の多い空知, 石狩, 上川などはよくその特徴があらわれ



第 1 図

ている。

また、同じように支庁別に総り患数と感受性者密度との関係を調べたところ、相関係数は 0.53 となり上の値に較べてかなり小さい。

さて、ポリオのり患数の予測にうつる前に、資料の問題点について述べよう。われわれが用いた資料における難点は資料の精度が明確でない点にあることは勿論であるが、次にあげられるべき点は先に述べたように age distribution 即ちり患から回復までの時間の分布の要求が満されていないということである。しかもなお、流行時においてはその道の専門家達は何等かの方法で予測を試み、それに対する対策をせまられているわけである。当時、北海道では5月頃から本格的流行が始まり、専門家達はおよそ9月頃終結するであろうと予想を立てたのであるが、9月になっても依然として流行は止まらなかった。如何なる観点に立脚して予想が立てられたかについては全く不明ではあるが、予想が完全にはずれたという事実は見逃せない。

また、上に述べたように資料の精度について述べねばならない。われわれの資料は高度の専門的精密検査に頼らずしてほとんどが開業医の外的所見から診断されたものようである。従って多くの他の病気が混入している疑いが充分ある。が併し、現在の単純な外的所見による診断に基いた報告を解析して、さらに次の月には何の位の同種の報告があるかを予測するという立場に立てばこの問題は解決されよう。さらに、正確なポリオの実数が必要ならば sampling によって精密検査を施してポリオの実数を推定すればよい。

データミステックなポリオの予測モデルとしては、3節の結果を用いたかったのであるが、先にも述べたように、このような資料は技術的に得られないので、以下に述べるような結果になってしまった。特に伝染病では患者の実数とともに不顕性感染の実数を把握することが重要である。いま、患者1人に対する不顕性感染の割合を $(K-1)$ 人とする。また、 i -支庁において、 t 時点で age が $(x, x+dx)$ にある平均患者数を $f_i(x, t)dx$ とすると、時刻 t における平患均者数は

$$F_i(t) = \int_0^t f_i(x, t) dx \quad (14)$$

で与えられる。従って i -支庁の t 時点における新患者の発生割合は

$$h_i(t) = \frac{dF_i(t)}{dt} = \int_0^t \lambda(x)(n_i - KF_i(t))f_i(x, t) dx \quad (15)$$

となる。ただし、 $\lambda(x)$ は age x の新患者の発生割合、 n_i は $t=0$ における i -支庁の感受性者数をあらわす。そこで、 $\lambda(x) = \lambda$ (常数) とすると、 t 時点における i -支庁の患者数 F_i は

$$F_i = \frac{n_i a_i}{K a_i + (n_i - K n_i) e^{-u_i t}} \quad (16)$$

ただし、 a_i は $t=0$ における i -支庁の患者数をあらわす。従って支庁間で互いに emigration, immigration が無いとし、互いに独立に発展すると仮定すると、北海道全域における t 時点の患者数は $\sum_i F_i(t)$ で支えられる。観測資料としては各時点(単位は日)で各支庁別新患者数 $h_i(t)$ が与えられている。これらの観測資料を用いて λ, K を推定しよう。新患者の発生間隔を $\bar{t}(m)$ とすると、 i -支庁の始めの単位時間 $(0, 1)$ における λ_i の推定値として

$$\frac{1}{\lambda_i(1)} = \frac{1}{h_i(1)} \sum_{m=0}^{h_i-1} (n_i - m) a_i \bar{t}_i(m) \quad (17)$$

を得る。 n_i は充分大きいとすると、(17) は

$$\frac{1}{\lambda_i(1)} = \frac{n_i a_i}{h_i(1)} \quad (18)$$

続く単位時間で K の推定値として

$$\frac{1}{K_i(1)} = \frac{F_i(1)}{\left(n_i - \frac{n_i a_i}{h_i(1)} \frac{h_i(2)}{F_i(1)} \right)} \quad (19)$$

続いて、 $K_i(1)$ が得られた条件の下で単位時間 (2, 3) では

$$\frac{1}{\lambda_i(2)} = \bar{t}_3(n_i - \bar{K}_1 F_i(2)) F_i(2) = \frac{(n_i - \bar{K}_1 F_i(2)) F_i(2)}{h_i(3)} \quad (20)$$

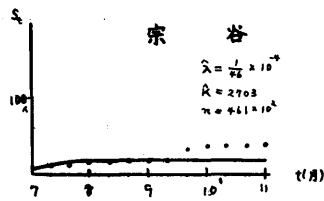
さらに、単位時間 (3, 4) では

$$\frac{1}{K_i(2)} = \frac{F_i(3)}{\left(n_i - \frac{\lambda_i(2) F_i(3)}{h_i(4)} \right)} \quad (21)$$

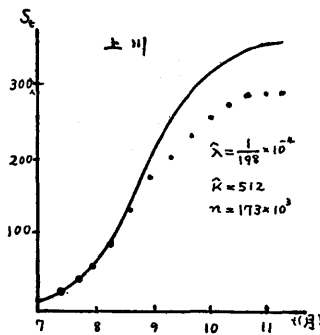
を導くことができる。つまり、上の λ, K は過去の process に関係した条件付平均値と考えればよい。このようにして λ_i, K_i の推定値 $1/\hat{\lambda}_i = \frac{1}{p} \sum_{j=1}^p 1/\lambda_i(p)$, $1/\hat{K}_i = \frac{1}{p} \sum_{j=1}^p 1/K_i(p)$ を求めた。そして各支庁別に $\hat{\lambda}_i, \hat{K}_i$ を計算した結果、安定した値を得るのに 40~50 日位の観測資料を必要とした。これらの推定値を用いて、各支庁別に予測を試みた結果は第 2 図に示された如くである。北海道全体で予測値は 1350 名であり、これに対して観測値は 1507 名であるから、推定誤差を

$$\text{推定誤差} = \frac{|\text{観測値} - \text{予測値}|}{\text{予測値}}$$

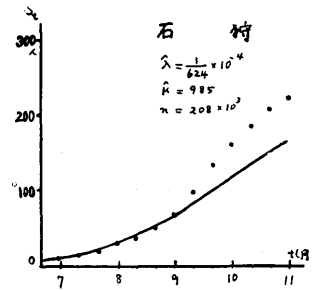
とすれば、各支庁別では例えば胆振のように相当くい違っているところはあるが、北海道全体での推定誤差は 15% となって、かなりよいことがわかる。



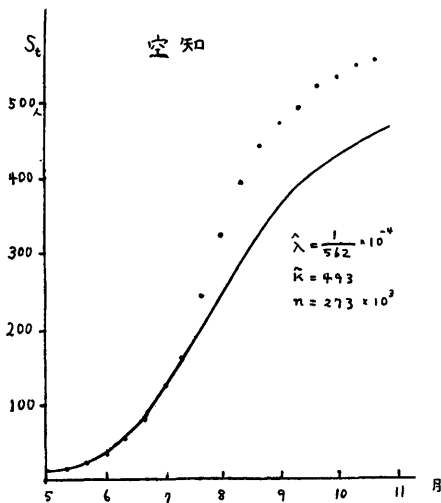
第 2 図の 1



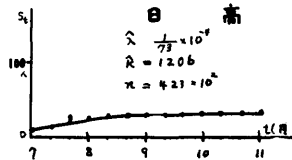
第 2 図の 2



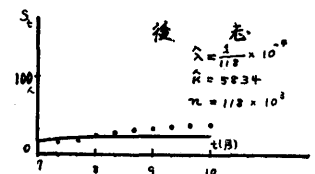
第 2 図の 3



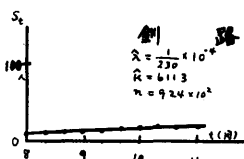
第 2 図の 4



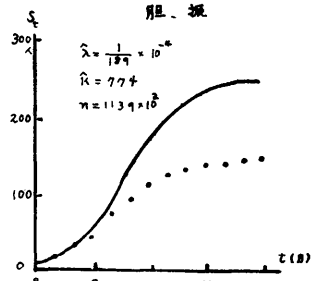
第 2 図の 5



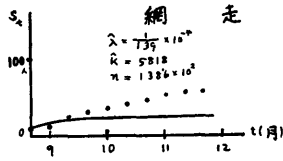
第 2 図の 6



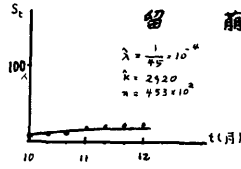
第 2 図の 7



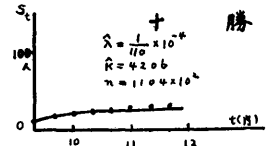
第 2 図の 8



第2図の9



第2図の10



第2図の11

5. 結 論

2節では時刻 t における感受性者数と age が $(x, x+dx)$ にある病人数 $d_x S(x, t)$ との積の期待値 $g(x, t)dx$ に関する微積分方程式並びにその特殊解, 3節ではその特別な場合として病気の特続時間が非常に長い場合の age distribution について述べた. さらに4節では 1960年北海道に大流行したポリオの資料にデターミニスチックなモデルを適用し, 予測を試みた. が併し, このポリオの伝播モデルの安定したパラメーターの推定値を得るのに 40~60日間の観測資料を必要とした. これでは, 早期に伝播を予測することはほど遠い. これには, homogeneous と考えられる地域の区分に1つは問題があると考えられる. が併し, 気象条件, 交通等を考えて厳格に地域を区分することはなかなか困難な問題であり, やはり現在ではいろいろな観点から行政的区分を用いた方が便利であるように思う.

最後にポリオの資料を提供していただいた北海道庁衛生部長小林医学博士並びに係長兵藤医学博士に心から謝意をあらわしたい.

この研究は文部省科学研究費による研究の1部である.

(統計数理研究所)

文 献

- [1] 崎野: エピデミックにおける age distribution について; 統計数理研究所彙報, 10巻, 1号, 1962.
- [2] // : エピデミック・モデルの歴史的発展; 統計数理研究所彙報, 9巻, 2号, 1962.