感染モデルの消滅確率

Extinction Probability of Infection Model

平場　誠示

2020 年 5 月 7 日（木）

講義ノート『確率過程の基礎』の
家系の存続モデル「ゴルトン-ワトソン過程」を抜粋し,
感染モデルに読み替えをしたものである.

理解するには, **数学 III** の知識 $+\alpha$ (連続と微分; **中間値の定理, ロルの定理, 平均値の定理** + 無限級数論)
と**確率論の基礎**ぐらいは必要だが, そこは, そういうものなのだと, 読み飛ばして貰えば良いかも $\cdots$!?

1　ゴルトン-ワトソン過程 (Galton-Watson Process)

Galton-Watson 過程 $\{Z_n\}_{n\geq 0}$ とは, 家系の世代数を表すモデルで, 家系が存続するかどうかを確率論的に議論するものであるが, 見方を変えれば, ウイルス感染の消滅モデルであるとも見れる.

これを用いて, どうであれば, 感染が消滅するのかについて数学的な結果を与えよう.

最初は 1 人の感染者がいたとして, その人が次に, Z_1 人に感染させるとする. 第 n 次感染者数を Z_n とする. $Z_0 = 1$ である. 各感染者は独立に Y 人に感染させるとするここで Y は $\mathbf{Z}_+$ に値をとる確率変数で分布 $(p_k)_{k\geq 0}$ をもつとする, i.e., $P(Y = k) = p_k$ $(k \geq 0)$ (勿論, 人毎に感染させる数は異なるので j 番目の人が, Y_j 人に感染させると区別するが取る確率は全て同じである, つまり, 分布は Y と同じ; $Y_j \overset{(\mathrm{d})}{=} Y$ と表す).

我々は感染数の分布 (p_k) に対し, その平均が存在することを仮定する.

$$m := \sum_{k=1}^{\infty} kp_k \in (0, \infty) : \quad \textbf{平均感染数} = \textbf{実効再生産数}.$$

今, q を 1 感染者から出発した GW 過程が消滅する確率とすると, $Z_1 \overset{(\mathrm{d})}{=} Y$ で, $\{Z_1 = k\}$ で条件付けると

$$\begin{aligned} \textbf{消滅確率}: \ q &= P(\text{消滅} \mid Z_0 = 1) = P({}^{\exists} n \geq 1; Z_n = 0 | \ Z_0 = 1) \\ &= \sum_{k\geq 0} P(\text{消滅} \mid Z_1 = k) P(Y = k) = \sum_{k\geq 0} q^k p_k. \end{aligned}$$

を得る．このとき $q=1$ は常にこの方程式の解となるが, 他に $q\in[0,1)$ となる解もあるかも知れない．そうならないための条件は何であろうか？この問に答えるために次の母関数 f を導入する．
q は方程式 $s=f(s)$ $(0\le s\le 1)$ の解となる;

$$f(s)=E[s^Y]=\sum_{k=0}^{\infty}p_k s^k \quad (|s|\le 1).$$

この級数は $|s|\le 1$ で絶対収束し, 従って $|s|<1$ で無限回項別微分可能である．また

$$f(0)=p_0,\quad f(1)=1,\quad f'(1)=\sum_{k\ge 1}kp_k=m.$$

定理 1.1　GW 過程 $\{Z_n\}$ は次をみたす: $P_1(\cdot)=P(\cdot|\ Z_0=1)$ と表すと,

$$\begin{aligned} m<1 \text{ or } [m=1,p_0>0] &\implies P_1(^\forall n\ge 1, Z_n\ge 1)=0,\quad \text{i.e., } q=1\\ m>1 &\implies P_1(^\forall n\ge 1, Z_n\ge 1)>0,\quad \text{i.e., } q<1\end{aligned}$$

$m>1$ のとき, 消滅確率 q は方程式 $f(s)=s$ の $[0,1)$ での一意解として与えられる．

注)　$p_0=0$ なら確実に感染者が残るので消滅確率 $q=0$ (このとき $m\ge 1$). 特に $p_1=1$ なら $m=1$ で $q=0$.

この結果は,
平均感染数 m が 1 より小なら, 確率 1 で, 感染者はいなくなり, たとえ $m=1$ であっても, $p_0>0$, 即ち, **感染者が感染させない確率が正でありさえすれば**, やはり, 感染者はいなくなる !!　ということである．

しかし, $m>1$ であれば, 感染者がいなくなる確率は 1 にはならない !!
つまり, 感染者が (即ちウイルスが) **永久に残り続ける可能性**があるということである．

ここで, **感染者が感染させない確率が正** $p_0>0$ とは, 感染した人が, 他へ感染させる前に治るか, 少なくとも入院して隔離されるか (勿論, その前に亡くなってしまうこともあるだろうが), そういう人が一人でもいれば良いということになる．逆に言えば, 周りが感染させられない様に, 常に注意することが必要であるということ．

ちなみに $p_0=0$ とは, 感染者は必ず, 1 人以上の人へ感染させるということで, 当然, 平均感染数は $m\ge 1$ となってしまい, 上の後半の残念な結果の方になってしまう．

まずいくつか必要な命題を証明する．母関数 f は $0\le s\le 1$ 上, $f(0)=p_0\ge 0$ から単調増加に $f(1)=1$ までの値をとるので, その合成を考えることができる．$f_1=f$, $f_{n+1}=f\circ f_n$ $(n\ge 1)$ と定義する．

命題 1.1　$n\ge 1$ に対し, $Z_0=1$ の条件のもと, Z_n の母関数は f_n となる, i.e., $E_1[s^{Z_n}]=f_n(s)$.

証明　$P_1=P, E_1=E$ と書く．$g_n(s)=E[s^{Z_n}]=\sum_{k=0}^{\infty}s^kP(Z_n=k)$ とおく．$n=1$ のとき $\{Z_0=1\}$ のもと Z_1 と Y は同分布なので, 明らかに $g_1(s)=E[s^Y]=f(s)$. $n\ge 1$ に対し, $g_n=f_n$ と仮定する．$\{Z_n=k\}$ のもと Z_{n+1} の分布は $\sum_{i=1}^{k}Y_i$ と同じで, $\{Y_i\}$ は独立で Y と同分布であることから

$$E[s^{Z_{n+1}}|\ Z_n=k]=E[\prod_{i=1}^{k}s^{Y_i}|\ Z_n=k]=\prod_{i=1}^{k}E[s^{Y_i}]=f(s)^k.$$

よって

$$g_{n+1}(s) = \sum_{k=0}^{\infty} E[s^{Z_{n+1}} |\ Z_n = k] P(Z_n = k) = \sum_{k=0}^{\infty} f(s)^k P(Z_n = k) = g_n(f(s)).$$

帰納法の仮定より, $g_{n+1}(s) = g_n(f(s)) = f_n(f(s)) = f_{n+1}(s)$. ■

命題 1.2 $E_1[Z_n] = m^n$ $(n \geq 0)$.

証明 $P_1 = P, E_1 = E$ と書く. $m = E[Y] = E[Z_1]$ と $E[Z_n |\ Z_{n-1} = k] = E[\sum_{i=1}^{k} Y_i] = km$ に注意して,

$$E[Z_n] = \sum_{k \geq 1} E[Z_n |\ Z_{n-1} = k] P(Z_{n-1} = k) = \sum_{k \geq 1} km P(Z_{n-1} = k) = mE[Z_{n-1}].$$

これを繰り返して $E[Z_n] = m^{n-1} E[Z_1] = m^n$ をえる. ■

[定理 1.1 の証明]

P_1 のもと Z_n の母関数が f_n であったことから, $P_1(Z_n = 0) = f_n(0)$ が成り立つ. $\{Z_n = 0\} \uparrow$ に注意すれば,

$$q = P_1(^{\exists} n \geq 1; Z_n = 0) = P_1(\bigcup_{n \geq 1} \{Z_n = 0\}) = \lim_{n \to \infty} f_n(0).$$

ここで $f_{n+1}(0) = f(f_n(0))$ より, $n \to \infty$ とすれば, f の連続性から, $q = f(q)$ をみたす.

($m < 1$ の場合) $P_1(Z_n \geq 1) \leq E_1[Z_n] = m^n$ より, $\{Z_n \geq 1\} \downarrow$ に注意して,

$$0 = \lim_{n \to \infty} P_1(Z_n \geq 1) = P_1(\bigcap_{n \geq 1} \{Z_n \geq 1\}) = P_1(^{\forall} n \geq 1, Z_n \geq 1) \quad \text{i.e., } q = 1.$$

($m = 1$ の場合) $p_0 > 0$ なら $p_0 + p_1 < 1$ (実際, もし $p_0 + p_1 = 1$ とすると $p_1 = 1 - p_0 < 1$ となり矛盾する). よって, $^{\exists} k \geq 2; p_k > 0$ に注意して,

$$f'(s) = \sum_{k \geq 1} k p_k s^{k-1} < f'(1) = \sum_{k \geq 1} k p_k = m = 1 \quad (0 < s < 1).$$

平均値の定理から $s \in (0,1)$ に対し, $^{\exists} c \in (s,1); f(1) - f(s) = f'(c)(1-s) < 1 - s$. $f(1) = 1$ より, 結局, $f(s) > s$ $(0 < s < 1)$ をえる. さらに $f(0) = p_0 > 0$ なので, $f(s) = s$ をみたす解は $[0,1]$ では $s = 1$ のみとなる. 故に $q = 1$.

($m > 1$ の場合) $p_0 + p_1 < 1$ に注意する (もし $p_0 + p_1 = 1$ なら $m = p_1 \leq 1$ より明らか). $f'(1) = m > 1$ で f' の連続性から

$$^{\exists} \eta > 0; 1 - \eta < {}^{\forall} s < 1, 1 < f'(s) < f'(1) = m$$

(ここでは最後の不等式までは必要としないが, 理由は上と同様である). 従って $1 - \eta < s < 1$ なら $f(s) < s$. また $f(0) = p_0 \geq 0$ なので, $g(s) = f(s) - s$ に対し中間値の定理を用いることにより, $^{\exists} s_1 \in [0,1); f(s_1) = s_1$. この解の一意性を示す. もし $^{\exists} s_2 \in [0,1); s_1 < s_2, f(s_2) = s_2$ とすると, $g(s_i) = 0$ で, また $f(1) = 1$ より, $g(1) = 0$. ロルの定理より, $0 \leq s_1 < {}^{\exists} \xi_1 < s_2 < {}^{\exists} \xi_2 < 1; g'(\xi_1) = g'(\xi_2) = 0$, i.e., $f'(\xi_1) = f'(\xi_2) = 1$. 一方, $p_0 + p_1 < 1$ より,

$$s \in (0,1) \quad \Longrightarrow \quad f''(s) = \sum_{k \geq 2} k(k-1) p_k s^{k-2} > 0.$$

これから $f'(s)$ は $s \in (0,1)$ で狭義単調増加となるが, 上の $f'(\xi_1) = f'(\xi_2) = 1$ に矛盾する. 故に $f(s) = s$ の解は $q = s_1$ or $q = 1$ のみとなる. さらに $q = 1$ とすると $1 = q = \lim_{n\to\infty} f_n(0)$ より, $n \gg 1$ (十分大) なら, $f_n(0) > 1 - \eta$. 上で示したことから $f_{n+1}(0) = f(f_n(0)) < f_n(0)$ となり, f_n が (n に関して) 単調増加であることに反する. よって $q = s_1 \in [0,1)$. ■

例 1.1 $p_0 = p_2 = 1/2$ のとき, 即ち, 感染者が誰にもうつさないか, または, 2 人にうつす場合しかなく, それらの確率が半々のとき, 平均感染数は $m = 1$ だが, ウイルスはいつか消滅する.

例 1.2 (Lotka (1939) アメリカの家系)

$$\text{幾何分布}: P(Y=0) = \frac{1}{2}, \quad P(Y=k) = \frac{1}{5}\left(\frac{3}{5}\right)^{k-1} \ (k \geq 1).$$

このとき

$$m = \frac{1}{5}\sum_{k\geq 1} k\left(\frac{3}{5}\right)^{k-1} = \frac{5}{4} > 1.$$

なので, 定理より, 消滅確率 q は次の方程式の 1 より小さい方の解:

$$s = f(s) = \frac{1}{2} + \frac{1}{5}\sum_{k\geq 1}\left(\frac{3}{5}\right)^{k-1} s^k, \quad \text{即ち}, \quad \frac{3}{5}s^2 - \frac{11}{10}s + \frac{1}{2} = 0$$

これを解くと $s = 5/6, 1$ となり, $q = 5/6$ を得る. 従って, ウイルスが永遠に残り続ける確率は 1/6 となる.

問 1.1 上の例で, 平均 $m = 5/4$ と方程式 $s = f(s)$ の解 $s = 5/6, 1$ を導く計算を確かめよ.