統計数理（1999）
第47巻 第1号 105-118

特集「統計基礎理論」
［研究詳解］

確率生成母関数の活用

統計数理研究所　平　野　勝　臣
大阪大学大学院*　安　芸　重　雄

（受付 1998年8月7日；改訂 1998年10月20日）

要　旨

独立ベルヌーイ試行の系列をマルコフ従属系列に拡張したとき，成功連に関する離散分布の問題は確率生成母関数（pgf）を用いて解析できることを示す．成功連に関する待ち時間問題の例では，条件付き pgf についての線形方程式を漸次構成し，これらを解くことによって解析する方法を示す．また pgf を求める一つの方法として，この母関数の母関数を解析することによって，もとの母関数を求める方法（蝦蟇の油法）の具体例，pgf が有理関数で与えられたとき，この母関数を持つ分布の確率を計算するための漸化式，成功連に関する離散分布の問題をマルコフチェインへ埋め込む方法，などが報告される．

キーワード：確率生成母関数，連の分布論，マルコフ従属試行，待ち時間問題．

1. はじめに

本報告で，確率生成母関数の離散分布論への有効な利用法を述べる．確率をはじめて学ぶ際，確率生成母関数（pgf）の基本的な性質をすぐに知る．長さ n の独立ベルヌーイ試行で成功の確率 p の2項分布の pgf は，例えば2項定理を使うことによって，$\varphi_n(t)=(pt+1-p)^n$ であることがわかる．そこで系列を従属にし，例えば $\{0,1\}$-値マルコフ連鎖にしたとき，2項分布に対応する分布，すなわち，長さ n の系列の中に1が起こる回数の分布を考えてみよう．この簡単な一般化を行おうとして，まずわかることは，組合せの方法を使うためには系列の独立性の仮定がいかに重要であったかということである．長さ n の系列の中に1が同じ数だけ起こる場合でも，系列の独立性の仮定が成り立たなければ，0と1の並び方によって，その確率は同じであるとは限らない．ではこの極めて基本的な問題に対してどのように取り組めばよいのだろうか．また，系列が独立の場合に使われる種々の方法の中で，このような一般化にも対応できる道具はあるのだろうか．本報告の目的は，このような問題に対して pgf が便利で有力な道具であることを示すことである．

成功連に関する離散分布の問題が1983年頃から再び議論されはじめた．当時は，系列が独立で，同じ分布（iid）に従う場合，組合せの方法を巧みに使うことによって確率の計算や pgf を調べることができた．1993年頃からこれらを一般化する一つの方向として，iid 系列をマルコフ系列とした場合の議論がなされるようになった．しかしながら組合せ法の適用は，場合分けが複

* 基礎工学研究科 情報数理系専攻 数理科学分野：〒560-8531 大阪府豊中市待兼山町 1-3.

雑になり，困難になってきた．これに対し pgf を用いる方法は，iid 系列の場合と同様にマルコフ系列の場合も解析が可能であることが示されている．系列が iid からマルコフに一般化されたとき，成功連に関する分布論は pgf を用いて解析することができることを具体例で述べる．本報告の目的は，これらのことを通して pgf が有力な道具であることを示すことであり，従って，本文で述べる問題に対してはその pgf を示すことにとどめる．結果については引用文献を参照されたい．

本報告の構成は以下の通りである．2 章で約束と記号を与え，3 章で条件付き pgf を漸次構成し，線形方程式を解くことによって連の分布論を解析する方法を示す．4 章では，確率母関数を求めるために，この母関数の母関数の解析から求める方法（蝦蟇の油法）を扱う．確率母関数が与えられたとき，この母関数を持つ分布の確率は母関数を展開すれば求まるが，5 章では母関数が有理母関数のとき確率を与える漸化式を紹介する．これまでの議論は条件付き pgf の方程式を解くことによって解析する方法であるが，これとは異なった方法として，最近，マルコフチェインへの埋め込みのテクニックが議論されている．この方法も有力であり，pgf に関した部分について 6 章で述べ，本報告を終えたい．

2. 約束と記号

本報告では，0 か 1 の値をとる確率変数の系列 $\{X_i, i=1,2,\ldots\}$ を扱い，これが iid（独立ベルヌーイ試行），マルコフチェイン，高次マルコフチェインに従う場合について考える．慣例に従って適宜，$\{X_i\}$ を i 番目の試行，X_i が値 1 をとることを成功，値 0 をとることを失敗とも言うことにする．

系列が iid であると仮定したときは $P(X_i=1)=p, P(X_i=0)=q(=1-p), i=1,2,\ldots,$ $0<p<1$ とする．

系列が iid であるとき，連続した k 個の成功が起こるまでの試行数の分布をオーダー k の幾何分布（記号で $G_k(p)$ とかく）という．$G_k(p)$ の pgf は Feller ((1968), p. 323) や Johnson et al. ((1992), p. 426) を参照されたい．$G_k(p)$ の台は $\{k, k+1,\ldots\}$ であるが，台を $\{a, a+1,\ldots\}$ にシフトしたときの分布を記号で $G_k(p, a)$ とかく．

系列がマルコフ系列であると仮定したときは，$P(X_0=1)=p$, $P(X_0=0)=q\,(=1-p)$, $0<p<1$ で $i=1,2,\ldots$ に対し

$$\begin{aligned} &P(X_i=1\,|\,X_{i-1}=1)=p_{11}, \qquad P(X_i=0\,|\,X_{i-1}=1)=p_{10}=1-p_{11},\\ &P(X_i=1\,|\,X_{i-1}=0)=p_{01}, \qquad P(X_i=0\,|\,X_{i-1}=0)=p_{00}=1-p_{01}, \end{aligned}$$

とする．

系列がm次マルコフ系列であると仮定したときは，$m\le k$ で，系列 $X_{-m+1}, X_{-m+2},\ldots, X_0, X_1,$ $X_2,\ldots$ が $\{0,1\}$-値m次マルコフ系列である．すなわち $x_1,\ldots,x_m=0,1$ と $i=1,2,\ldots$ に対して

$$\begin{aligned} \pi_{x_1,\ldots,x_m} &= P(X_{-m+1}=x_1, X_{-m+2}=x_2,\ldots, X_0=x_m)\\ p_{x_1,\ldots,x_m} &= P(X_i=1\,|\,X_{i-m}=x_1, X_{i-m+1}=x_2,\ldots, X_{i-1}=x_m)\\ &= 1-q_{x_1,\ldots,x_m} \end{aligned}$$

であるとする．ここで任意の $x_1,\ldots,x_m=0,1$ に対して $0<p_{x_1,\ldots,x_m}, q_{x_1,\ldots,x_m}<1$ であると仮定する．

3. 確率生成母関数

3.1 iid 系列

Feller ((1968), pp. 349-350) は古典的な破産問題で，破産するまでの賭けの回数の分布の pgf を与える方法を述べている．1回の賭けで勝つ確率を p，負ける確率を q $(=1-p)$ とする．1回の賭け金が1で，勝つと2になってもどり，負けると賭け金を失う．はじめの所持金は z で，何回かの賭けの後（試行数），所持金が a $(0<z<a)$ になるか，0になるかでこのゲームが終わる．このときの試行数の分布の pgf を求める．$u_{z,n}$ を n 回の試行でこのゲームが終わる確率とする．ここで，a と z は整数とする．このとき

$$u_{z,n+1}=pu_{z+1,n}+qu_{z-1,n}\,,\qquad 1<z<a-1,\quad n\geq 1$$

である．ここでこの問題に対する境界条件を

$$\begin{aligned} u_{0,n}&=u_{a,n}=0\,, & n&\geq 1\\ u_{a,0}&=u_{0,0}=1\,, & u_{z,0}&=0\,,\quad z>0 \end{aligned}$$

とおく．このとき上の式は $0<z<a$ を満たすすべての z と，$n\geq 0$ を満たすすべての n に対して成り立つ．pgf $U_z(s)=\sum_{n=0}^{\infty}u_{z,n}s^n$ とすると

$$U_z(s)=psU_{z+1}(s)+qsU_{z-1}(s)$$

を得る．

Ebneshahrashoob and Sobel (1990) は独立ベルヌーイ試行において，長さ s の成功の連かまたは長さ r の失敗の連のいずれかが起こるまでの試行数 (W_{sooner} とかく) の分布と，これらの連のいずれもが起こるまでの試行数 (W_{later} とかく) の分布を求めている．前者を sooner problem，後者を later problem ということにする．

はじめに sooner problem を解くための方程式を与える．成功の確率を p (>0)，失敗の確率を q $(=1-p)$，$\phi^{(s,r)}(t)$ を W_{sooner} の分布の pgf，長さ i の成功連を S_i，長さ j の失敗連を F_j と表す．また，現在 S_i が起こっているという条件の下で，今から，長さ s の成功の連かまたは長さ r の失敗の連のいずれかが起こるまでの試行数の条件付き分布の pgf を $\phi_i(t)$，現在 F_j が起こっている条件の下で，今から，いずれかの事象が起こるまでの試行数の条件付き分布の pgf を $\phi^{(j)}(t)$ とする．このとき次の方程式が成り立つ．

$$\begin{cases} \phi^{(s,r)}(t)=pt\phi_1(t)+qt\phi^{(1)}(t)\\ \phi_1(t)\quad\;=pt\phi_2(t)+qt\phi^{(1)}(t)\\ \qquad\cdots\\ \phi_{s-2}(t)\;=pt\phi_{s-1}(t)+qt\phi^{(1)}(t)\\ \phi_{s-1}(t)\;=pt\times 1+qt\phi^{(1)}(t) \end{cases}\qquad \begin{cases} \phi^{(1)}(t)\quad\;=pt\phi_1(t)+qt\phi^{(2)}(t)\\ \phi^{(2)}(t)\quad\;=pt\phi_1(t)+qt\phi^{(3)}(t)\\ \qquad\cdots\\ \phi^{(r-2)}(t)=pt\phi_1(t)+qt\phi^{(r-1)}(t)\\ \phi^{(r-1)}(t)=pt\phi_1(t)+qt\times 1 \end{cases}$$

この方程式を解いて $\phi^{(s,r)}(t)$ が得られる．

次に later problem を解くための方程式を与える．$\psi^{(s,r)}(t)$ を W_{later} の分布の pgf とする．まず，現在 S_i が起こっていて，sooner event はまだ起こっていないという条件の下で，今から later event が起こるまでの試行数の条件付き分布の pgf を $\psi_i(t)$ とする．また，現在 F_j が起こっていて，sooner event はまだ起こっていないという条件の下で，今から later event が起こるまでの試行数の条件付き分布の pgf を $\psi^{(j)}(t)$ とする．さらに，今丁度 sooner event が起こったところで，それが S_s であったという条件の下で，今から later event が起こるまでの試行数の条件付き分布の pgf を $\psi_s(t)$ とする．同様に，今丁度 sooner event が起こったところで，そ

れが F_r であったという条件の下で，今から later event が起こるまでの試行数の条件付き分布の pgf を $\psi^{(r)}(t)$ とする．sooner event が起こった後，later event が起こるまでに，sooner event と同種の連が起こったとしても，そのときの状態は，sooner event がはじめて起こったときの状態と，later event を待つことに関してはまったく同じであることに注意する．また，sooner event が S_s であり，S_s はすでに起こっていて，現在 F_j を観測中であるという条件の下で，今から，later event（この場合は sooner event が S_s であったから later event は F_r である）が起こるまでの試行数の条件付き分布の pgf を $\psi_s^{(j)}(t)$ と表す．また，sooner event が F_r であり，F_r はすでに起こっていて，現在 S_i を観測中であるという条件の下で，今から，later event（この場合は sooner event が F_r であったから later event は S_s である）が起こるまでの試行数の条件付き分布の pgf を $\psi_i^{(r)}(t)$ と表す．このとき，各状態から1ステップ先の状態を考慮することによって，次の方程式が成り立つ．

$$\begin{cases} \psi^{(s,r)}(t) = pt\psi_1(t) + qt\psi^{(1)}(t) \\ \psi_1(t) \quad = pt\psi_2(t) + qt\psi^{(1)}(t) \\ \quad \cdots \\ \psi_{s-2}(t) \quad = pt\psi_{s-1}(t) + qt\psi^{(1)}(t) \\ \psi_{s-1}(t) \quad = pt\psi_s(t) + qt\psi^{(1)}(t) \\ \psi_s(t) \quad = pt\psi_s(t) + qt\psi_s^{(1)}(t) \end{cases} \qquad \begin{cases} \psi^{(1)}(t) \quad = pt\psi_1(t) + qt\psi^{(2)}(t) \\ \psi^{(2)}(t) \quad = pt\psi_1(t) + qt\psi^{(3)}(t) \\ \quad \cdots \\ \psi^{(r-2)}(t) = pt\psi_1(t) + qt\psi^{(r-1)}(t) \\ \psi^{(r-1)}(t) = pt\psi_1(t) + qt\psi^{(r)}(t) \\ \psi^{(r)}(t) \quad = pt\psi_1^{(r)}(t) + qt\psi^{(r)}(t) \end{cases}$$

$$\begin{cases} \psi_s^{(1)}(t) \quad = pt\psi_s(t) + qt\psi_s^{(2)}(t) \\ \quad \cdots \\ \psi_s^{(r-2)}(t) = pt\psi_s(t) + qt\psi_s^{(r-1)}(t) \\ \psi_s^{(r-1)}(t) = pt\psi_s(t) + qt \times 1 \end{cases} \qquad \begin{cases} \psi_1^{(r)}(t) \quad = pt\psi_2^{(r)}(t) + qt\psi^{(r)}(t) \\ \quad \cdots \\ \psi_{s-2}^{(r)}(t) \quad = pt\psi_{s-1}^{(r)}(t) + qt\psi^{(r)}(t) \\ \psi_{s-1}^{(r)}(t) \quad = pt \times 1 + qt\psi^{(r)}(t) \end{cases}$$

この方程式から $\psi^{(r)}(t)$, $\psi_s(t)$, $\psi^{(1)}(t)$, $\psi_1(t)$ について解いて，$\psi^{(s,r)}(t)$ が得られる．

3.2 マルコフ系列

Aki and Hirano (1993) は，iid 系列をマルコフ系列に拡張して，sooner problem と later problem を解いた．ここでの pgf の利用法について述べる．iid 系列をマルコフ系列のような場合に拡張するとき，組合わせ法の解析に比べて pgf を用いる方法は極めて有用であることがわかる．

マルコフ系列 $\{X_i, i=1,2,\ldots\}$ において長さ r の0の連を表す事象 E_0 と長さ k の1の連を表す事象 E_1 を考え，いずれか早く起こる方の事象 (sooner event) が起こるまでの試行数 (W_{sooner} とかく）と後で起こる方の事象 (later event) が起こるまでの試行数 (W_{later} とかく）の分布に注目する．

まず sooner problem について述べる．$\phi(t)$ を W_{sooner} の分布の pgf，$\xi(t)$ を $X_0=0$ が与えられたときの W_{sooner} の条件付き分布の pgf とし，$\psi(t)$ を $X_0=1$ が与えられたときの W_{sooner} の条件付き分布の pgf とする．$\phi_i(t)$ を，長さ i の1の連が起こっているという条件の下で，そのときから sooner event が起こるまでの試行数の条件付き分布の pgf とする．また，$\phi^{(j)}(t)$ を，長さ j の0の連が起こっているという条件の下で，そのときから sooner event が起こるまでの試行数の条件付き分布の pgf とする．このとき，独立ベルヌーイ試行の場合と同様に考えて，$\phi(t), \xi(t), \psi(t), \phi_i(t), (i=1,2,\ldots,k-1)$，$\phi^{(j)}(t), (j=1,2,\ldots,r-1)$ は次の方程式を満たすことがわかる．

$$\phi(t) = p_0\xi(t) + p_1\psi(t) \tag{3.1}$$

$$\xi(t) = p_{01}t\phi_1(t) + p_{00}t\phi^{(1)}(t) \tag{3.2}$$

$$\psi(t) = p_{11}t\phi_1(t) + p_{10}t\phi^{(1)}(t) \tag{3.3}$$

$$\begin{cases} \phi_1(t) = p_{11}t\phi_2(t) + p_{10}t\phi^{(1)}(t) \\ \quad\cdots \\ \phi_{k-2}(t) = p_{11}t\phi_{k-1}(t) + p_{10}t\phi^{(1)}(t) \\ \phi_{k-1}(t) = p_{11}t + p_{10}t\phi^{(1)}(t) \end{cases} \tag{3.4}$$

$$\begin{cases} \phi^{(1)}(t) = p_{01}t\phi_1(t) + p_{00}t\phi^{(2)}(t) \\ \quad\cdots \\ \phi^{(r-2)}(t) = p_{01}t\phi_1(t) + p_{00}t\phi^{(r-1)}(t) \\ \phi^{(r-1)}(t) = p_{01}t\phi_1(t) + p_{00}t \end{cases} \tag{3.5}$$

(3.4) と (3.5) を解くと

$$\phi_1(t) = (p_{11}t)^{k-1} + B(t)p_{10}t\phi^{(1)}(t), \tag{3.6}$$

$$\phi^{(1)}(t) = (p_{00}t)^{r-1} + A(t)p_{01}t\phi_1(t), \tag{3.7}$$

を得る．ここに

$$A(t) = \frac{1-(p_{00}t)^{r-1}}{1-p_{00}t}, \qquad B(t) = \frac{1-(p_{11}t)^{k-1}}{1-p_{11}t},$$

である．(3.6) と (3.7) から pgf $\phi(t), \xi(t), \psi(t)$ は

$$\phi(t) = p_0\xi(t) + p_1\psi(t),$$

$$\xi(t) = p_{01}t\phi_1(t) + p_{00}t\phi^{(1)}(t),$$

$$\psi(t) = p_{11}t\phi_1(t) + p_{10}t\phi^{(1)}(t),$$

で与えられる．ここに

$$\phi_1(t) = \frac{(p_{11}t)^{k-1} + B(t)(p_{10}t)(p_{00}t)^{r-1}}{1-A(t)B(t)(p_{01}t)(p_{10}t)},$$

$$\phi^{(1)}(t) = \frac{A(t)(p_{11}t)^{k-1}(p_{01}t) + (p_{00}t)^{r-1}}{1-A(t)B(t)(p_{01}t)(p_{10}t)},$$

である．

ここで $\lim_{r\to\infty}\phi(t)$ を求めると，これは，系列を iid からマルコフに一般化したとき，オーダー k の幾何分布に対応する分布の pgf を与える．すなわち，マルコフ系列において長さ k の 1 の連が起こるまでの試行数の分布の pgf は

$$\frac{(p_0p_{01}+p_1p_{11})p_{11}{}^{k-1}t^k + p_1(p_{01}p_{10}-p_{00}p_{11})p_{11}{}^{k-1}t^{k+1}}{1-(p_{00}t + p_{01}p_{10}t^2 + p_{01}p_{10}p_{11}t^3 + \cdots + p_{01}p_{10}p_{11}{}^{k-2}t^k)}$$

である．また，これを展開すれば，この分布の確率関数は

$$P(X=x)=\begin{cases} p_0p_{01}p_{11}{}^{k-1}+p_1p_{11}{}^k & \text{if } x=k\,, \\ p_0\cdot\displaystyle\sum_{x_1+2x_2+\cdots+kx_k=x-k}\binom{x_1+\cdots+x_k}{x_1,\ldots,x_k}p_{00}{}^{x_1}p_{01}{}^{x_2+\cdots+x_k+1} & \\ \quad\times p_{10}{}^{x_2+\cdots+x_k}p_{11}{}^{x_3+2x_4+\cdots+(k-2)x_k+k-1} & \\ \quad+p_1\cdot\displaystyle\sum_{j=1}^{k}p_{11}{}^{j-1}p_{10}\cdot\sum_{x_1+2x_2+\cdots+kx_k=x-k-j}\binom{x_1+\cdots+x_k}{x_1,\ldots,x_k} & \\ \quad\times p_{00}{}^{x_1}p_{01}{}^{x_2+\cdots+x_k+1}p_{10}{}^{x_2+\cdots+x_k} & \\ \quad\times p_{11}{}^{x_3+2x_4+\cdots+(k-2)x_k+k-1} & \text{if } x\geq k+1 \end{cases}$$

で与えられることがわかる．Rajarshi (1974) も renewal theory (Feller (1968) 参照) を利用して，特別な初期条件の下でこの分布を扱っている．

次に later problem について述べる．

事象 E_0 と E_1 のうち後で起こる方の事象 (later event) が起こるまでの試行数を W_{later} と表す．$\phi(t)$ を W_{later} の分布の pgf, $\xi(t)$ を $X_0=0$ が与えられたときの W_{later} の条件付き分布の pgf, $\psi(t)$ を $X_0=1$ が与えられたときの W_{later} の条件付き分布の pgf とする．$\phi_i(t)$ を長さ i の 1 の連が起こっているという条件の下で，そのときから later event が起こるまでの試行数の条件付き分布の pgf とする．また，$\phi^{(j)}(t)$ を長さ j の 0 の連が起こっているという条件の下で，そのときから later event が起こるまでの試行数の条件付き分布の pgf とする．$i=0,1,\ldots,k-1$ に対して，$\psi_i(t)$ を，E_1 が E_0 よりも先に起こり，かつ，E_1 はすでに起こり，現在長さ i の 1 の連が起こっているとき，今から later event (この場合は E_0) が起こるまでの試行数の条件付き分布の pgf とする．また，$j=0,1,\ldots,r-1$ に対して，$\psi^{(j)}(t)$ を，E_1 が E_0 よりも先に起こり，かつ，E_1 はすでに起こり，現在長さ j の 0 の連が起こっているとき，今から later event (この場合は E_0) が起こるまでの試行数の条件付き分布の pgf とする．$i=0,1,\ldots,k-1$ に対して，$\xi_i(t)$ を，E_0 が E_1 よりも先に起こり，かつ，E_0 はすでに起こり，現在長さ i の 1 の連が起こっているとき，今から later event (この場合は E_1) が起こるまでの試行数の条件付き分布の pgf とする．また，$j=0,1,\ldots,r-1$ に対して，$\xi^{(j)}(t)$ を，E_0 が E_1 よりも先に起こり，かつ，E_0 はすでに起こり，現在長さ j の 0 の連が起こっているとき，今から later event (この場合は E_1) が起こるまでの試行数の条件付き分布の pgf とする．

このとき

$$\phi_0(t)=\phi^{(0)}(t)\,,\qquad \psi_0(t)=\psi^{(0)}(t)\,,\qquad \xi_0(t)=\xi^{(0)}(t)$$

である．さらに，$\phi_0,\ldots,\phi_{k-1},\phi^{(0)},\ldots,\phi^{(r-1)},\psi_0,\ldots,\psi_{k-1},\psi^{(0)},\ldots,\psi^{(r-1)},\xi_0,\ldots,\xi_{k-1},\xi^{(0)},\ldots,\xi^{(r-1)}$ はつぎの方程式を満たす．

(3.8) $$\phi(t)=p_0\xi(t)+p_1\psi(t)$$

(3.9) $$\xi(t)=p_{01}t\phi_1(t)+p_{00}t\phi^{(1)}(t)$$

(3.10) $$\psi(t)=p_{11}t\phi_1(t)+p_{10}t\phi^{(1)}(t)$$

(3.11) $$\begin{cases}\phi_1(t)=p_{11}t\phi_2(t)+p_{10}t\phi^{(1)}(t)\\ \quad\cdots\\ \phi_{k-2}(t)=p_{11}t\phi_{k-1}(t)+p_{10}t\phi^{(1)}(t)\\ \phi_{k-1}(t)=p_{11}t\psi_0(t)+p_{10}t\phi^{(1)}(t)\end{cases}$$

$$\text{(3.12)}\quad \begin{cases} \phi^{(1)}(t) = p_{01}t\phi_1(t) + p_{00}t\phi^{(2)}(t) \\ \cdots \\ \phi^{(r-2)}(t) = p_{01}t\phi_1(t) + p_{00}t\phi^{(r-1)}(t) \\ \phi^{(r-1)}(t) = p_{01}t\phi_1(t) + p_{00}t\xi_0(t) \end{cases}$$

$$\text{(3.13)}\quad \begin{cases} \psi_0(t) = p_{11}t\psi_1(t) + p_{10}t\psi^{(1)}(t) \\ \psi_1(t) = p_{11}t\psi_2(t) + p_{10}t\psi^{(1)}(t) \\ \cdots \\ \psi_{k-2}(t) = p_{11}t\psi_{k-1}(t) + p_{10}t\psi^{(1)}(t) \\ \psi_{k-1}(t) = p_{11}t\psi_0(t) + p_{10}t\psi^{(1)}(t) \end{cases}$$

$$\text{(3.14)}\quad \begin{cases} \psi^{(1)}(t) = p_{01}t\psi_1(t) + p_{00}t\psi^{(2)}(t) \\ \cdots \\ \psi^{(r-2)}(t) = p_{01}t\psi_1(t) + p_{00}t\psi^{(r-1)}(t) \\ \psi^{(r-1)}(t) = p_{01}t\psi_1(t) + p_{00}t \end{cases}$$

$$\text{(3.15)}\quad \begin{cases} \xi_0(t) = p_{01}t\xi_1(t) + p_{00}t\xi^{(1)}(t) \\ \xi_1(t) = p_{11}t\xi_2(t) + p_{10}t\xi^{(1)}(t) \\ \cdots \\ \xi_{k-2}(t) = p_{11}t\xi_{k-1}(t) + p_{10}t\xi^{(1)}(t) \\ \xi_{k-1}(t) = p_{11}t + p_{10}t\xi^{(1)}(t) \end{cases}$$

$$\text{(3.16)}\quad \begin{cases} \xi^{(1)}(t) = p_{01}t\xi_1(t) + p_{00}t\xi^{(2)}(t) \\ \cdots \\ \xi^{(r-2)}(t) = p_{01}t\xi_1(t) + p_{00}t\xi^{(r-1)}(t) \\ \xi^{(r-1)}(t) = p_{01}t\xi_1(t) + p_{00}t\xi^{(0)}(t) \end{cases}$$

(3.13) と (3.14) を解くと

$$\text{(3.17)}\quad \psi_0(t) = \frac{B(t)(p_{10}t)(p_{00}t)^{r-1}(p_{11}t) + (p_{10}t)(p_{00}t)^{r-1}}{1 - A(t)B(t)(p_{01}t)(p_{10}t) - (p_{11}t)^k - A(t)(p_{11}t)^{k-1}(p_{01}t)(p_{10}t)}$$

を得る．また同様に (3.15) と (3.16) を解くと

$$\text{(3.18)}\quad \xi_0(t) = \frac{A(t)(p_{00}t)(p_{01}t)(p_{11}t)^{k-1} + (p_{01}t)(p_{11}t)^{k-1}}{1 - A(t)B(t)(p_{01}t)(p_{10}t) - B(t)(p_{01}t)(p_{10}t)(p_{00}t)^{r-1} - (p_{00}t)^r}$$

を得る．結局，W_{later} の pgf は

$$\phi(t) = (p_0p_{01}t + p_1p_{11}t)C(t) + (p_0p_{00}t + p_1p_{10}t)D(t)$$

で与えられる．ここに

$$C(t) = \frac{(p_{11}t)^{k-1}\psi_0(t) + B(t)(p_{10}t)(p_{00}t)^{r-1}\xi_0(t),}{1 - A(t)B(t)(p_{01}t)(p_{10}t)},$$

$$D(t) = \frac{(p_{00}t)^{r-1}\xi_0(t) + A(t)(p_{11}t)^{k-1}(p_{01}t)\psi_0(t)}{1 - A(t)B(t)(p_{01}t)(p_{10}t)}$$

で，$\psi_0(t)$ と $\xi_0(t)$ は (3.17) と (3.18) でそれぞれ与えられている．

なお，Uchida and Aki (1995) はこの問題を ℓ 個の E_0 と m 個の E_1 についての sooner and later problem として一般化し，pgf を用いて解析している．

3.3 高次マルコフ系列

Hirano et al. (1997) は，m 次マルコフ系列で，長さ k の1の連がはじめて起こる（までの試行数を τ とかく）までに長さ ℓ $(m \le \ell)$ の1の連の起こる回数（オーバーラップして数える）の分布について pgf を用いて調べている．さらにそこでは，長さ ℓ の成功連の数え方を non-overlapping，丁度長さ ℓ，長さ ℓ 以上，の3通りについても，やはり pgf を用いて調べている．この節では，この4通りの数え方で，pgf を用いて調べる方法について述べる．

はじめに，長さ $\ell(m \le \ell)$ の1の連の数え方をオーバーラップして数えるとき，m 次マルコフ系列で，τ までに長さ ℓ の1の連の起こる回数の分布を pgf を用いて調べる．

$x_i = 0, 1$ $(i = 1, 2, \ldots, m)$ に対して，$\phi^{(x_1,\ldots,x_m)}(t)$ を $X_{-m+1} = x_1, \ldots, X_0 = x_m$ のとき，はじめて長さ k の1の連が起こるまでに長さ ℓ の1の連が起こる回数 (overlap) の条件付き分布の pgf とする．$\phi_i(t)$ を直前の結果が長さ i の1の連であるとき，これから先，長さ ℓ の1の連の起こる回数 (overlap) の条件付き分布の pgf とする．このとき，問題の性質上，どのような初期条件から出発しても，いつか必ず（長さ k の1の連が起こってしまう前に）長さ ℓ の1の連をはじめて観測することがあるのだから，

$$\phi^{(x_1,\ldots,x_m)} = t\phi_\ell, \qquad \text{for each initial state } (x_1, x_2, \ldots, x_m)$$

である．これはすべての条件付き分布は等しく，初期条件に依存していないことを示している．そこで $\phi = t\phi_\ell$ とおく．ところで ϕ_ℓ は，その定義から，その時点（例えば i 番目の試行としよう）の直前までに，長さ ℓ の1の連を観測中である．このとき，その時点（i 番目）の試行を行うと1または0が必ず観測される．それが1であれば，（これは条件付き確率 $p_{1\cdots1}$ で起こり）オーバーラップして1の連を数えているので，長さ ℓ の1の連が1回起こり，そして次の時点（$i+1$ 番目の試行）の直前で長さ $\ell+1$ の1の連を観測中という状態になる．i 番目の試行の結果が0であれば（これは条件付き確率 $1-p_{1\cdots1}$ で起こる）1の連はここで途絶えてしまう．そしてさらに試行を続けて行くと必ずその次の長さ ℓ の1の連を観測する．従って，

$$\phi_\ell = p_{1\cdots1}t\phi_{\ell+1} + (1-p_{1\cdots1})\phi.$$

同様にして

$$\phi_{\ell+1} = p_{1\cdots1}t\phi_{\ell+2} + (1-p_{1\cdots1})\phi$$
$$\cdots$$
$$\phi_{k-2} = p_{1\cdots1}t\phi_{k-1} + (1-p_{1\cdots1})\phi.$$

さらに ϕ_{k-1} では，つぎの観測で1が起これば長さ k の1の連が完成し，そのときの条件付き分布の pgf は1である．また，つぎの観測で1が起こらなければ，前の説明のように，長さ k の1の連を待っているので，必ず長さ ℓ の1の連がその前に起こる．従って

$$\phi_{k-1} = p_{1\cdots1}t \times 1 + (1-p_{1\cdots1})\phi$$

となる．以上から $\phi^{(x_1,\ldots,x_m)}$ について解けば，この pgf は $G_{k-\ell}(p_{1\cdots1}, k-\ell+1)$ の pgf と一致することが分かる．

次に non-overlapping の数え方で条件付き pgf の求め方を与える．$m \le \ell < k$ で $k = \nu\ell + \mu$ とする．ここに ν と μ は非負整数で $0 \le \mu < \ell$ である．$x_i = 0, 1$ $(i = 1, 2, \ldots, m)$ に対して，$\phi^{(x_1,\ldots,x_m)}(t)$ を $X_{-m+1} = x_1, \ldots, X_0 = x_m$ のとき，はじめて長さ k の1の連が起こるまでに長さ ℓ の1の連が起こる回数 (non-overlap) の条件付き分布の pgf とする．$\phi_i(t)$ を直前の結果が長さ i の1の連であるとき，これから先，長さ ℓ の1の連の起こる回数 (non-overlap) の条件付き分布の pgf とする．このとき，先程と同様に，

$$\psi^{(x_1,\dots,x_m)} = t\psi_\ell \qquad \text{for each initial state}\,(x_1, x_2,\dots, x_m)$$

である．これはすべての条件付き分布は等しく，初期条件に依存していないことを示している．そこで $\psi = t\psi_\ell$ とおく．ところで ψ_ℓ は，その定義から，その時点（例えば i 番目の試行としよう）の直前で長さ ℓ の1の連を観測している．i 番目の試行の結果，1が起これば（これは確率 $p_{1\cdots1}$ で起こる），オーバーラップせずに1の連を数えているので，$\ell>1$ であれば長さ ℓ の1の連はまだ起こらず，そして次の $(i+1)$ 番目の試行の直前では長さ $\ell+1$ の1の連を観測中という状態になる．i 番目の試行の結果，0が起これば（これは確率 $q_{1\cdots1}$ で起こる），ここで1の連は一旦途絶えるが，さらに試行を続けていくと必ずその次の，長さ ℓ の1の連を観測する．従って，

$$\psi_\ell = p_{1\cdots1}\,\psi_{\ell+1} + q_{1\cdots1}\psi$$

同様に，$\ell>2$ であれば，

$$\psi_{\ell+1} = p_{1\cdots1}\,\psi_{\ell+2} + q_{1\cdots1}\psi$$

が成り立つ．

オーバーラップせずに長さ ℓ の1の連を数えるので，長さ 2ℓ の1の連を観測するときにも長さ ℓ の1の連が起こることになる．従って，

$$\psi_{2\ell-1} = p_{1\cdots1}t\psi_{2\ell} + q_{1\cdots1}\psi\,.$$

以下，まったく同様にして

$$\text{(3.19)}\qquad \begin{cases} \psi_{2\ell} = p_{1\cdots1}\psi_{2\ell+1} + q_{1\cdots1}\,\psi \\ \psi_{2\ell+1} = p_{1\cdots1}\psi_{2\ell+2} + q_{1\cdots1}\,\psi \\ \cdots \\ \psi_{2\ell+\ell-2} = p_{1\cdots1}\psi_{2\ell+\ell-1} + q_{1\cdots1}\,\psi \\ \psi_{2\ell+\ell-1} = p_{1\cdots1}t\,\psi_{3\ell} + q_{1\cdots1}\,\psi \end{cases}$$

$$\cdots$$

$$\text{(3.20)}\qquad \begin{cases} \psi_{(\nu-1)\ell} = p_{1\cdots1}\psi_{(\nu-1)\ell+1} + q_{1\cdots1}\,\psi \\ \psi_{(\nu-1)\ell+1} = p_{1\cdots1}\psi_{(\nu-1)\ell+2} + q_{1\cdots1}\,\psi \\ \cdots \\ \psi_{(\nu-1)\ell+\ell-2} = p_{1\cdots1}\psi_{(\nu-1)\ell+\ell-1} + q_{1\cdots1}\,\psi \\ \psi_{\nu\ell-1} = p_{1\cdots1}t\psi_{\nu\ell} + q_{1\cdots1}\,\psi \end{cases}$$

$$\text{(3.21)}\qquad \begin{cases} \psi_{\nu\ell} = p_{1\cdots1}\psi_{\nu\ell+1} + q_{1\cdots1}\,\psi \\ \psi_{\nu\ell+1} = p_{1\cdots1}\psi_{\nu\ell+2} + q_{1\cdots1}\,\psi \\ \cdots \\ \psi_{\nu\ell+\mu-2} = p_{1\cdots1}\psi_{\nu\ell+\mu-1} + q_{1\cdots1}\,\psi \\ \psi_{\nu\ell+\mu-1} = p_{1\cdots1}\times 1 + q_{1\cdots1}\,\psi \end{cases}$$

そこで (3.19)，(3.20)，(3.21) を解いて ψ_ℓ を得，$t\psi_\ell=\psi$ より pgf を求めることができる．

次に丁度長さが ℓ の数え方で考察したときの条件付き pgf の求め方を述べる．$x_i=0,1$ $(i=1,2,\dots,m)$ に対して，$\varphi^{(x_1,\dots,x_m)}(t)$ を $X_{-m+1}=x_1,\dots, X_0=x_m$ のとき，はじめて長さ k の1の連が起こるまでに丁度長さ ℓ の1の連が起こる回数の条件付き分布の pgf とする．$\varphi_i(t)$ を直前の結果が長さ i の1の連であるとき，これから先，丁度長さ ℓ の1の連の起こる回数の条件付き分布の pgf とする．今長さ ℓ の1の連が起こっているとき，丁度長さ ℓ の1の連は次の結果が0で

なければ起こらない．すなわち overlap と non-overlap とは異なり，この時点では丁度長さ ℓ の 1 の連はまだ起こっていないので，$\varphi^{(x_1,\dots,x_m)}(t)$ は

$$\varphi^{(x_1,\dots,x_m)} = \varphi_\ell \qquad \text{for each initial state}\,(x_1, x_2,\dots, x_m)$$

とかける．これはすべての条件付き分布は等しく，初期条件に依存していないことを示している．そこで $\varphi^{(x_1,\dots,x_m)} = \varphi$ とおく．今丁度長さ ℓ の 1 の連が起こっているとする．次に 1 が起これば，丁度長さ ℓ の 1 の連は起こらないし，0 が起これば丁度長さ ℓ の 1 の連が起こるので

$$\varphi_\ell = p_{1\cdots1}\varphi_{\ell+1} + q_{1\cdots1}t\varphi$$

を得る．また今丁度長さ $\ell+1$ の 1 の連が起こっているとする．次に 1 が起こっても，0 が起こっても，丁度長さ ℓ の 1 の連は起こらなくて，1 が起こっていれば，丁度長さ $\ell+1$ の 1 の連は長さ $\ell+2$ の連となり，0 が起こっていれば，長さ k の 1 の連を待っているので必ず長さ ℓ の 1 の連が起こるので

$$\varphi_{\ell+1} = p_{1\cdots1}\varphi_{\ell+2} + q_{1\cdots1}\varphi$$

を得る．以下全く同様にして順次 $\varphi_{k-2} = p_{1\cdots1}\varphi_{k-1} + q_{1\cdots1}\varphi$ まで得る．φ_{k-1} については，仮定から今丁度長さ $k-1$ の 1 の連が起こっているので，次に 1 が起これば，長さ k の 1 の連が起こり，この試行は終わりで，このときの pgf は 1 である．また 0 が起これば，長さ k の 1 の連を待っているので必ず長さ ℓ の 1 の連が起こるので

$$\varphi_{k-1} = p_{1\cdots1} \times 1 + q_{1\cdots1}\varphi$$

を得る．これらの方程式を解くと，

$$\varphi = \frac{p_{1\cdots1}{}^{k-\ell}}{p_{1\cdots1}{}^{k-\ell} + q_{1\cdots1} - q_{1\cdots1}t}$$

であるから，求める分布は幾何分布 $G_1(\alpha, 0)$, $\alpha = p_{1\cdots1}{}^{k-\ell}/(p_{1\cdots1}{}^{k-\ell} + q_{1\cdots1})$ であることがわかる．

最後に，長さ ℓ 以上の場合について考察したときの条件付き pgf の求め方を述べる．$x_i = 0, 1$ $(i = 1, 2,\dots, m)$ に対して，$\rho^{(x_1,\dots,x_m)}(t)$ を $X_{-m+1} = x_1,\dots, X_0 = x_m$ のとき，はじめて長さ k の 1 の連が起こるまでに長さ ℓ 以上の 1 の連が起こる回数の条件付き分布の pgf とする．$\rho_i(t)$ を直前の結果が長さ i の 1 の連であるとき，これから先，長さ ℓ 以上の 1 の連の起こる回数の条件付き分布の pgf とする．今長さ k の 1 の連を待っているので必ず長さ ℓ の 1 の連は起こり，この時点で，次に 1 が起こっていれば，長さ ℓ 以上の 1 の連はまだ起こらない．従って $\rho^{(x_1,\dots,x_m)}(t)$ は

$$\rho^{(x_1,\dots,x_m)} = \rho_\ell \qquad \text{for each initial state}\,(x_1, x_2,\dots, x_m)$$

とかける．これはすべての条件付き分布は等しく，初期条件に依存していないことを示している．そこで $\rho^{(x_1,\dots,x_m)} = \rho$ とおく．今長さ ℓ の 1 の連が起こっているとする．次に 1 が起これば，長さ ℓ の 1 の連は続いているのでまだ起こらないし，0 が起これば長さ $\ell+1$ の 1 の連が 1 回起こるので

$$\rho_\ell = p_{1\cdots1}\rho_{\ell+1} + q_{1\cdots1}t\rho$$

を得る．また今長さ $\ell+1$ の 1 の連が起こっているとする．次に 1 が起これば，長さ ℓ の 1 の連は続いているので起こらなくて，長さ $\ell+1$ の 1 の連は長さ $\ell+2$ の連となる．また 0 が起こっていれば，長さ ℓ 以上の 1 の連が 1 回起こるので

$$\rho_{\ell+1} = p_{1\cdots1}\rho_{\ell+2} + q_{1\cdots1}t\rho$$

を得る．以下全く同様にして順次 $\rho_{k-2} = p_{1\cdots1}\rho_{k-1} + q_{1\cdots1}t\rho$ を得る．ρ_{k-1} については，仮定から今長さ $k-1$ の 1 の連が起こっているので，次に 1 が起これば，長さ ℓ 以上の 1 の連が 1 回起こり，かつ長さ k の 1 の連が起こり，この試行は終わりで，このときの pgf は 1 である．また 0 が起これば，長さ ℓ 以上の 1 の連が 1 回起こり，かつ長さ k の 1 の連を待っているので必ず長さ ℓ の 1 の連が 1 回起こるので

$$\rho_{k-1} = p_{1\cdots1}t + q_{1\cdots1}t\rho$$

を得る．これらの方程式を解くと，

$$\rho = \frac{{p_{1\cdots1}}^{k-\ell}t}{1 - t + {p_{1\cdots1}}^{k-\ell}t}$$

であるから，求める分布はオーダー 1 の幾何分布 $G_1({p_{1\cdots1}}^{k-\ell})$ であることがわかる．

4. 蝦蟇の油法

今までは，サポートが無限に広がっている分布について見てきたが，サポートが有限な分布に対しては，特有のテクニックがある．簡単のため，$X_1,\ldots,X_n$ を iid とし，$X_1,\ldots,X_n$ の中に起こる長さ k の 1 の連の数（オーバーラップしないように数える）の分布について考える．この分布の pgf を $\varphi(k,n;t)$ と表す．はじめて 0 が現れる位置を $1,2,\ldots,k-1,k,k+1$ 以上，と分類して確率空間を分割すれば，漸化式

$$\begin{aligned}
\varphi(k,n;t) &= 1, & n<k\\
\varphi(k,n;t) &= q\varphi(k,n-1;t) + pq\varphi(k,n-2;t)\\
&\quad + \cdots + p^{k-1}q\varphi(k,n-k;t) + p^k t\varphi(k,n-k;t), & n\ge k
\end{aligned}$$

が得られる．新しい変数 z を導入してこれらの式の両辺に z^n を掛けて加えると，$\varphi(k,n;t)$ の生成母関数 $\Phi(k;t,z) = \sum_{n=0}^{\infty}\varphi(k,n;t)z^n$ に関する簡単な方程式が得られて，それを解いて，

$$\Phi(k;t,z) = \frac{\sum_{m=1}^{k}(pz)^{m-1}}{1-(\sum_{m=1}^{k}p^{m-1}qz^m) - p^k tz^k}$$

を得る．この式を展開して z^n の係数を求めれば，$\varphi(k,n;t)$ が得られる．Aki and Hirano (1988) は，より一般な拡張されたオーダー k の 2 項分布の pgf をこの方法を用いて与えている．

このように，直接 φ を求めるかわりに，φ の生成母関数 Φ を求め，これを基にして φ を求める方法を，Wilf ((1994), p. 118) は蝦蟇の油法（"Snake Oil" method）と言っている．

蝦蟇の油法を用いることで，Aki and Hirano (1993) は，系列がマルコフのとき，オーダー k の 2 項分布に対応する分布（試行数 n で，長さ k の 1 の連の起こる回数の分布）の pgf を陽の形で与えている．また，この拡張として Uchida (1997) は，系列が m 次マルコフのとき，対応する分布の pgf を（蝦蟇の油法を用いることなく）与えている．

5. 有理母関数からの確率の計算

pgf が与えられたとき，この pgf を持つ分布の確率はこの pgf を展開すれば求まる．この章で pgf が有理母関数のとき，この pgf から確率を求めるための漸化式を与える．

$P(x)=\beta_0+\beta_1x+\cdots+\beta_ex^e, Q(x)=1+\alpha_1x+\cdots+\alpha_dx^d$ のとき，有理母関数

$$\frac{P(x)}{Q(x)}=\sum_{n\geq 0}f(n)x^n.$$

の x^n の係数は

$$f(n)=-\alpha_1f(n-1)-\cdots-\alpha_nf(n-d)+\beta_n$$

で与えられる (Stanley (1986), p. 205)．ここに $k<0$ のとき $f(k)=0, k>e$ のとき $\beta_k=0$ とする．

この結果を用いて，pgf が有理母関数で与えられれば，この pgf をもつ分布の確率を早く求めることができる．

6. マルコフチェイン埋め込み法

Fu and Koutras (1994) は連の分布を求めるために，finite Markov chain imbedding approach を与えた．これは長さ n の試行で，t $(1\leq t\leq n)$ 番目の試行までに考察している連 (0 と 1 のパターンとしてもよい．以下同様) の起こった回数と，$t+1$ 番目以降の各試行後でこの連が起こる回数を数え上げることができる適切な情報を状態としてもつマルコフチェインで表現する方法である．このとき推移確率行列は考察している連の構造によって決まり，この行列の積で連の分布の確率関数，更に pgf が与えられる．またこの方法で waiting time problem も扱っている．この方法は Koutras and Alexandrou (1995), Han and Aki (1998) によって，さらに研究され，pgf を求める方法としても有力なものになっているが，ここではこれ以上ふれない．

pgf を直接扱うのではないが，指数型分布族の性質を用いて，ラプラス変換を求める Stefanov and Pakes (1997) の方法がある．

長さ k の 1 の連 (0 と 1 のパターンとしてもよい) が起こるまでの試行数を τ とする．長さ k の 1 の連が起こるのを待っているので，いま 1 が i $(i<k)$ 回続けて起こっているとする．この次の試行で長さ $i+1$ の 1 の連か，または長さ 0 の 1 の連のいずれかが起こることになる．そこでこれを一般的に表して，状態 i から状態 j への one-step transition とするようなマルコフチェインを考える．状態の集合 $\{0,1,\ldots,n+1\}$ を持つ time homogeneous なマルコフチェインが，状態 i から状態 j へ推移する確率を $p_{i,j}$, τ をある停止時刻，$N_{i,j}(\tau)$ を τ までに状態 i から状態 j への one-step transition の起こる回数とする．このとき，$N_{i,j}(\tau)$ の起こる確率

$$\prod_{i,j}p_{i,j}^{N_{i,j}(\tau)}=\exp\Bigl(\sum_{i,j}N_{i,j}(\tau)\ln p_{i,j}\Bigr)$$

を得る．ここにはパラメータ間に線形制約があり，一般的には，これは curved exponential family である．absorbing state を 1 つ持つマルコフチェインでは，absorption の時刻はよい性質を持った停止時刻である．即ち，τ を長さ k の 1 の連が起こるまでの試行数とすると，このとき，線形独立な $N_{i,j}(\tau)$ の数はフリーパラメータの数と等しくなり，指数型分布族となる．そこで適当なパラメータ変換を行い，指数型分布族の性質から $N_{i,j}(\tau)$ の同時分布のラプラス変換を求めるというものである．このラプラス変換から pgf を求めればよい．

この方法で Stefanov and Pakes (1997) は，Aki and Hirano (1994, 1995) の結果を確認している．

謝　　辞

本報告は統計数理研究所共同研究 10-共研 A-8 の援助を受けた。

参 考 文 献

Aki, S. and Hirano, K. (1988). Some characteristics of the binomial distribution of order k and related distributions, *Statistical Theory and Data Analysis II* (ed. K. Matusita), 211-222, North-Holland, Amsterdam.

Aki, S. and Hirano, K. (1993). Discrete distributions related to succession events in a two-state Markov chain, *Statistical Science & Data Analysis* (eds. K. Matusita, M. L. Puri and T. Hayakawa), 467-474, VSP Publishers, Utrecht.

Aki, S. and Hirano, K. (1994). Distributions of numbers of failures and successes until the first consecutive k successes, *Ann. Inst. Statist. Math.*, **46**, 193-202.

Aki, S. and Hirano, K. (1995). Joint distributions of numbers of success-runs and failures until the first consecutive k successes, *Ann. Inst. Statist. Math.*, **47**, 225-235.

Ebneshahrashoob, M. and Sobel, M. (1990). Sooner and later waiting time problems for Bernoulli trials: Frequency and run quotas, *Statist. Probab. Lett.*, **9**, 5-11.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Wiley, New York.

Fu, J. C. and Koutras, M. V. (1994). Distribution theory of runs: A Markov chain approach, *J. Amer. Statist. Assoc.*, **89**, 1050-1058.

Han, Q. and Aki, S. (1998). Joint distributions of runs in a sequence of multi-state trials, *Ann. Inst. Statist. Math.* (to appear).

Hirano, K., Aki, S. and Uchida, M. (1997). Distributions of numbers of success-runs until the first consecutive k successes in higher order Markov dependent trials, *Advances in Combinatorial Methods and Applications to Probability and Statistics*, (ed. N. Balakrishnan), 401-410, Birkhäuse, Boston.

Johnson, N. L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete Distributions*, 2nd ed., Wiley, New York.

Koutras, M. V. and Alexandrou, V. A. (1995). Runs, scans and urn model distributions: A unified Markov chain approach, *Ann. Inst. Statist. Math.*, **47**, 743-766.

Rajarshi, M. B. (1974). Success runs in a two-state Markov chain, *J. Appl. Probab.*, **11**, 190-192.

Stanley, R. P. (1986). *Enumerative Combinatorics*, Vol. I, Wadsworth & Brooks/Cole, California.

Stefanov, V. T. and Pakes, A. G. (1997). Explicit distributional results in pattern formation, *Ann. Appl. Probab.*, **7**, 666-678.

Uchida, M. (1998). On number of occurences of success runs of specified length in a higher-order two -state Markov chain, *Ann. Inst. Statist. Math.*, **50**, 587-601

Uchida, M. and Aki, S. (1995). Sooner and later waiting time problems in a two-state Markov chain, *Ann. Inst. Statist. Math.*, **47**, 415-433.

Wilf, H. S. (1994). generatingfunctionology, 2nd ed., Academic Press, California.

Use of Probability Generating Function for Distribution Theory of Runs

Katuomi Hirano

(The Institute of Statistical Mathematics)

Sigeo Aki

(Department of Informatics and Mathematical Science, Osaka University)

Let X_1, X_2,... be a sequence of independent and identically distributed $\{0, 1\}$-valued random variables, a $\{0, 1\}$-valued Markov chain or a $\{0, 1\}$-valued higher order Markov chain. Let E_0 be the event that a run of "0" of length r occurs and let E_1 be the event that a run of "1" of length k occurs in the sequence X_1, X_2,... .

In the case of Markov dependent trials, discrete distributions related to the events E_0 and E_1 are studied. The probability generating functions of the distributions of the waiting times of the sooner and later occurring events are given. To obtain the probability generating function of the distribution of the number of occurrences of E_1 in X_1, X_2,..., X_n, we use the *snake oil method*.

The distributions of numbers of overlapping and non-overlapping occurrences of success-runs of length ℓ, and the distributions of numbers of occurrences of success-runs of exact length ℓ and of length ℓ or more until the first occurrence of success-run of length k in the m-th $(m \leq \ell < k)$ order Markov dependent trials are studied: The distribution of overlapping occurrences is the geometric distribution of order $(k - \ell + 1)$, and the both of exact length ℓ and of length ℓ or more are the geometric distribution of order 1. To show these results we describe how to solve by means of the conditional probability generating function method. A finite-state Markov chain imbedding technique is also illustrated.

Key words: Probability generating function, distribution theory of run, Markov dependent trials, waiting time problems.