数理解析研究所講究録
第 853 巻 1993 年 88-98

林構造の数え上げと分枝過程

慶應義塾大学　理工学部　渋谷政昭　(Masaaki Sibuya)

要約とまえがき

ガルトン-ワトソン過程が停止したときの、節点総数を固定した条件の下での林構造を調べると、枝の数が幾何分布に従うときに、可能な順序ありラベル無し林構造が等確率で現れる（定理４）。ポアソン分布のときも類似の成果となる（定理５）。

証明の準備のために、林構造の数え上げ（定理１、２）と、ある種の組合せ論的恒等式（定理３）について述べる。

本稿の一部は、渋谷、他 (1993) 、Sibuya, et al. (1993) 、で発表したが、ここでは組合せ論的な内容を改めて系統的に述べる。

§1　順序あり、ラベル無し、林構造の数え上げ

整数 t, s にたいして、η を

$$
(1) \qquad \eta(t,s) = \begin{cases} \dfrac{t-s}{t}\dbinom{t+s-1}{s}, & 0 \le s < t \text{ のとき}, \\ 0, & \text{他の場合}, \end{cases}
$$

で定義する。$\eta(s+1,s) = C_s$ はカタラン数で、η はその１つの拡張である。（注１）

<u>予備定理１</u>

$$
(2) \qquad \eta(t+1,s) = \sum_{k=0}^{s} \eta(t,k), \qquad t = 1,2,\cdots,
$$

逆に

$$
\eta(1,s) = \begin{cases} 1, & s = 0, \\ 0, & s \neq 0, \end{cases}
$$

を初期値とする (2) の解は (1) である。

注. (2) は t の多項恒等式でもある。

<u>証明</u>　$\eta(t+1,s)$ の (1) 式の表現で、2 項係数の上引数を減少し分解する。それを変形して

$$\eta(t+1,s)=\eta(t,s)+\eta(t+1,s-1) \tag{3}$$

を得る。右辺末項に (3) を反覆適用する。□

<u>定理 1</u>　r 本のラベル無し順序木から成る林構造の数は、木にも順序があり、節点の総数が t であるとして

$$\eta(t,s), \quad s=t-r,$$

である。s は根以外の節点の数である。

<u>証明</u>　$r+1=t+1-s$ 本の、節点の総数 $t+1$ 個の林構造で、第 1 の木の根から出る枝の数を k とする $(0\le k\le s)$ 。第 1 の木の根を除くと、

$$r+k=t-s+k \text{ 本}, \quad \text{節点総数 } t$$

の林構造となり、その総数は $\eta(t,s-k)$ である。$k=0,1,\cdots,s$ について加えれば、予備定理 1 より

$$\sum_{k=0}^{s}\eta(t,s-k)=\sum_{k=0}^{s}\eta(t,k)=\eta(t+1,s)$$

が得られる。□

§2　順序無し、ラベルあり、林構造の数え上げ

整数 t,s にたいして、ζ を

$$\zeta(t,s)=\begin{cases}(t-s)t^{s-1}, & 0\le s<t \text{ のとき},\\ 0, & \text{他の場合},\end{cases} \tag{4}$$

で定義する。（注 2 ）

<u>予備定理 2</u>

$$\zeta(t+1,s)=\sum_{k=0}^{s}\binom{s}{k}\zeta(t,k), \qquad t=1,2,\cdots, \tag{5}$$

逆に

$$\zeta(1,s)=\begin{cases}1, & s=0,\\ 0, & s\neq 0,\end{cases}$$

を初期値とする (5) の解は (4) である。

注. (5) は t の多項恒等式でもある。

<u>証明</u>

$$\begin{aligned}右辺 &= \sum_k \binom{s}{k} t^k - \sum_k \binom{s}{k} k t^{k-1}\\ &= (1+t)^s - s\sum_k \binom{s-1}{k-1} t^{k-1}\\ &= (1+t-s)(1+t)^{s-1}\end{aligned}$$

で左辺が得られる。□

<u>定理 2</u> (Takács, 1990)　r 本のラベルつき、順序無し木の林構造の数は、根のラベルを $1,2,\cdots,r$ と固定し、木の順序も無いとし、節点の総数を t とすると

$$\zeta(t,s),\quad s=t-r,$$

である。s は根以外の節点の数である。

<u>証明</u>　根のラベルが $0,1,2,\cdots,r$ である r 本の、節点総数 $t+1$ の林構造で、ラベル 0 の根から k 本の枝が出ているとする。ラベル 0 の根を除き、その枝に $r+1,\cdots,r+k$ のラベルをつけ、残りの $t-r-k$ 個の節点に可能なラベルの配列を与えると、その数は$\zeta(t,s-k)$ である。元のラベル 0 の根を戻すと、その枝の節点につけるラベルの可能性が $\binom{t-r}{k}$ 通りある。したがって予備定理 2 より

$$\sum_{k=0}^{s}\binom{s}{k}\zeta(t,s-k)=\sum_{k=0}^{s}\binom{s}{k}\zeta(t,k)=\zeta(t+1,s)$$

が $r+1$ 本の節点総数 $t+1$ の順序無しラベルあり林構造の数である。□

注. Lovász (1979) は別の数え上げ法を述べている。

§3 ガルトン-ワトソン過程の節点総数

ガルトン-ワトソン過程は、分枝過程で、各節点での枝の数（粒子が分裂して生じる新しい粒子の数、生物体の子供の数）が独立で同一分布に従う、もっとも単純な場合である。Watson and Galton (1874) が、家系の断絶のモデルとして導入した。

各節点での枝の数 $X_1, X_2, \cdots$ の確率関数 (probability function, pf) を

$$g(k) = P(X_1 = k), \qquad k = 0, 1, 2, \cdots,$$

とし、r 個の節点（始祖）から出発して、過程が停止する（家系が途絶える）までの、始祖を含めた節点総数（家系の規模、確率変数）を Y_r とする。

$$E(X_1) > 1 \quad \text{ならば} \quad P(Y_r = \infty) > 0$$

となり、Y_r は improper な確率分布をもつ。$E(X_1) = 1$ のときには $P(Y_r = \infty) = 0$ だが $E(Y_r) = \infty$ で、$E(X_1) < 1$ が $P(Y_r = \infty) = 0, E(Y_r) < \infty$ の必要十分条件である。Feller (1968), Harris(1963) 参照。

Y_r の pf を

$$h^{(r)}(k) = P(Y_r = k), \qquad k = 0, 1, 2, \cdots,$$

とする。ただし $h^{(r)}(0) = \cdots = h^{(r)}(r-1) = 0$ である。X_1, Y_r の確率母関数についての議論から

(6) $$h^{(r)}(r+y) = \frac{r}{r+y} g^{*(r+y)}(y), \qquad y = 0, 1, 2, \cdots,$$

が求まる。ただし $g^{*j}(y)$ は $g(y)$ の j 回合成積である。ある種の g についてはこれが容易に計算できる。たとえば、

(7) $$g(y) = e^{-\theta}\theta^y / y! \text{のとき}$$
$$h^{(r)}(r+y) = \frac{r(r+y)^{y-1}\theta^y}{y!} e^{-(r+y)\theta}, \quad 0 < \theta,$$

$$g(y) = \binom{m}{y} p^y q^{m-y} \text{のとき}$$

(8)

$$h^{(r)}(r+y) = \frac{r}{r+y} \binom{m(r+y)}{y} p^y q^{m(r+y)-y},$$

$$g(y) = \binom{m+y-1}{y} p^m q^y \text{のとき}$$

(9)

$$h^{(r)}(r+y) = \frac{r}{r+y} \binom{m(r+y)+y-1}{y} p^{m(r+y)} q^y,$$

などである（いずれも $y = 0, 1, 2, \cdots,\quad 0 < p < 1,\quad q = 1 - p$ ）。

詳しくは、たとえば Consul and Shenton (1972,1973)。

§4 節点総数の直接確率計算

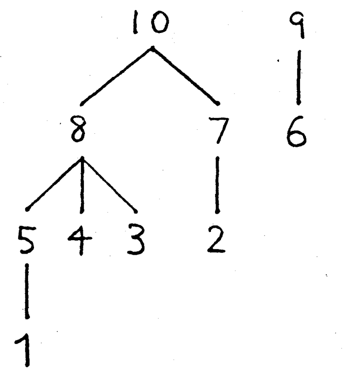

図1 節点番号 j

r 本の木から成る林構造で

節点の数 ＝ 枝の数 ＋ r

である。実際すべての枝と、その先の節点とを対応させると、根に対応する枝がない。

分枝過程一般で、第0水準（世代）から出発して、節に $1, 2, \cdots$ の番号をつける。同じ世代の中での順番は任意である。分枝過程が停止したときの節点総数を t とし、上記の節点番号 i を $j = t + 1 - i$ に変える（図1参照)。節点 j から出る枝の数を x_j とする。節点 $1, 2, \cdots, j$ とそれから出る枝は部分林構造をなしているから、(節点の数) > (枝の数) 、つまり

$$x_1 + \cdots + x_j < j, \qquad j = 1, 2, \cdots, t-1,$$

である。そこで

$$\begin{aligned} \mathcal{K}(t,s) = \{(x_1, \cdots, x_t) :\ & x_i \geq 0,\ i = 1, \cdots, t; \\ & x_1 + \cdots + x_j < j,\ j = 1, \cdots, t-1; \\ & x_1 + \cdots + x_t = s\}, \quad 0 \leq s \leq t, \end{aligned} \tag{10}$$

と定義すると、$(x_1,\cdots,x_t)\in\mathcal{K}(t,s)$ は r 本の節点から出発して節点総数 t のときに停止するすべての分枝過程の枝の数を表しているから、

(11)
$$h^{(r)}(t)=\sum_{(x_1,\cdots,x_t)\in\mathcal{K}(t,s)}\left(\prod_{i=1}^{t}g(x_i)\right)$$

が得られる。

(6) 式でなく (11) 式から $h^{(r)}(t)$ を計算することもできるが、逆に (6) から得られる (7)–(9) を (11) の左辺に入れることにより次のような等式が得られる。

<u>定理 3</u>

(11)
$$\sum_{(x_1,\cdots,x_t)\in\mathcal{K}(t,s)}\binom{s}{x_1,\cdots,x_t}=(t-s)^{s-1}=\zeta(t,s).$$

(12)
$$\sum_{(x_1,\cdots,x_t)\in\mathcal{K}(t,s)}\left\{\prod_{i=1}^{t}\binom{m}{x_i}\right\}=\frac{t-s}{t}\binom{mt}{s}.$$

(13)
$$\sum_{(x_1,\cdots,x_t)\in\mathcal{K}(t,s)}\left\{\prod_{i=1}^{t}\binom{m+x_i-1}{x_i}\right\}=\frac{t-s}{t}\binom{mt+s-1}{s}.$$

最後の式で $m=1$ とすれば

(14)
$$|\mathcal{K}(t,s)|=\frac{t-s}{t}\binom{t+s-1}{s}=\eta(t,s).$$

注. 最初の式は Takács (1989) の得た式の変形である。

§5 ガルトン-ワトソン過程の林構造の等確率性

はじめに述べた主目的に入る。ガルトン-ワトソン過程で、枝の数の分布が幾何分布またはポアソン分布であるとき、それぞれ幾何分布ガルトン-ワトソン過程、ポアソン分布ガルトン-ワトソン過程と呼ぶ。

<u>定理４</u>　r 個の節点から出発する幾何分布のガルトン-ワトソン過程が、節点総数 t で停止した条件の下で、$\eta(t,s)$, $s=t-r$, 種類の順序あり、ラベル無し、林構造が等確率で現れる。

<u>証明</u>　X_j が幾何分布に従うとき、(11) の右辺の積は

$$\prod_{j=1}^{t} pq^{x_j} = p^t q^s, \quad s=t-r,$$

となって、r,t だけで定まり、$(x_1,\cdots,x_t)$ にはよらない。数列 $(x_1,\cdots,x_t)$ が異なれば、それに対応する順序林構造が異なる（(14) 式はそれを裏付けている)。つまり $\eta(t,s)$ 種類の順序あり、ラベル無し、林構造が等確率で現れる。□

<u>定理５</u>　r 個の節点から出発するポアソン分布ガルトン-ワトソン過程が、節点総数 t で停止したとき、根以外の $s=t-r$ 個の節点に、可能な異なるラベルを等確率で与える。このとき、$\zeta(t,s)$ 種類の順序無し、ラベル付き、林構造が等確率で現れる。

<u>証明</u>

第１段．　同じラベル無し順序無し林構造に対応するラベル無し順序あり林構造の数え上げ:　　まず１本の順序木の根元で考える。根から出る x 本の枝の先の節点を根とする部分木で、順序無し木としては同一のものを類別する。A の型が α_1 本、B の型が α_2 本、$\cdots$ あるとする。$x=\alpha_1+\cdots+\alpha_k$(図２の場合には $x=7$, A 型 $\alpha_1=2$, B 型 $\alpha_2=2$, C 型 $\alpha_3=3$)。

$A,B,\cdots$ 等の型が順序木として多様であることを無視して、与えられた部分木の順列を考えると

$$\binom{x}{\alpha_1,\alpha_2,\cdots,\alpha_k} = \frac{x!}{\prod_{i=1}^{k}\alpha_i!} \tag{15}$$

通りの順列が可能である（図２の場合、$\frac{7!}{2!2!3!}$)。

A の型の順序木が ν_1 通り、B の型の順序木が ν_2 通り、$\cdots$ 可能であるとする。完全に対称な木で、各節点で枝を並べ換えても不変な木は $\nu=1$ である（図２の場合、$\nu_1=4,\nu_2=3,\nu_3=1$)。

これらの対称性を考えると、同じ順序無し木を導く順序木の数は

$$\binom{x}{\alpha_1,\alpha_2,\cdots,\alpha_k}\prod_{i=1}^{k}\nu_i^{\alpha_i}=x!\prod_{i=1}^{k}\frac{\nu_i^{\alpha_i}}{\alpha_i!} \tag{16}$$

である（図2の場合は、$\frac{7!}{2!2!3!}\cdot 4^2 3^2 1^3$）。これを少し違った形で表そう。

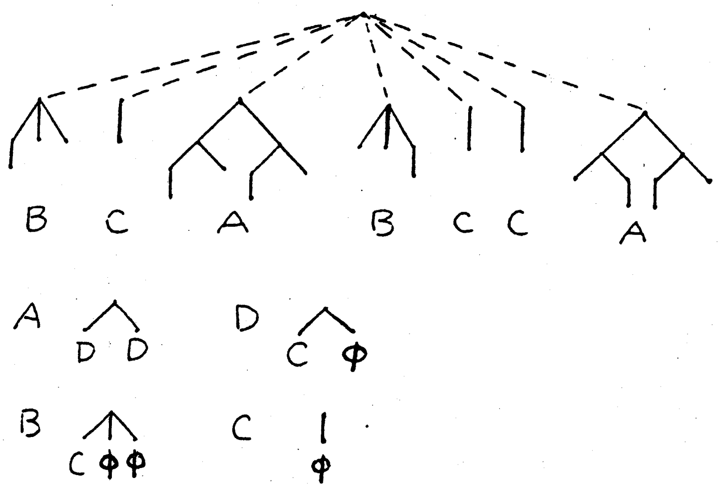

図2　順序木と順序無し木

上記の部分木の構造について、再び上記の考察が成り立つ。i 番目の型の1本の順序木を取り上げる。x_i 本の枝の先が $\alpha_{i1},\cdots,\alpha_{ik_i}$ 本ずつの順序無し木として同値なものに分かれ、それぞれが $\nu_{i1},\cdots,\nu_{ik_i}$ 通りの順序木をもつものとすると、

$$\nu_i=\binom{x_i}{\alpha_{i1},\cdots,\alpha_{ik_i}}\prod_{j=1}^{k_i}\nu_{ij}^{\alpha_{ij}}$$

となる（図2の場合、A は D,D の形で $x_1=2,k_1=1,\alpha_{11}=2,\nu_{11}=2,\nu_1=\binom{2}{2}2^2=4$; B は C,ϕ,ϕ の形で $x_2=3,k_2=2,\alpha_{21}=1,\nu_{21}=1,\alpha_{22}=2,\nu_{22}=1$, $\nu_2=\binom{3}{2,1}1^2 1=3$, である）。(16) での $\prod\nu_i^{\alpha_i}$ は、ν_i を根の先のすべての節点（順序木）についての乗算であることを注意しておく。

この $\nu_i,\nu_{ij},\cdots$ の計算を反復していくと、最終的には末端が空となり、$\nu_{ij}=1, j=1,2,\cdots$ の場合で終わる（図2では A が DD, D がさらに $C\phi$ と表される）。

そこで根も含めてすべての節点に $j=1,2,\cdots,t$ と番号を付け、j 番目の節点に (16) の構造を考えると、1本の木に関する計算結果は

$$\prod_{j=1}^{t}\binom{x_j}{\alpha_{j1},\cdots,\alpha_{jk_j}}$$

となる（図2の場合は $\binom{7}{2,2,3}\binom{2}{2}\binom{2}{2}\binom{3}{1,2}\binom{3}{1,2}\binom{2}{1,1}\binom{2}{1,1}\binom{2}{1,1}\binom{2}{1,1}$）。

林構造に移ると、r 本の順序木が与えられた状態は、木で根の下の部分木が与えられた状況と同等であるから、仮想的な番号 0 の根を考え、r 本の順序木をその枝先の部分木とみなし、

$$\nu_0=\binom{x_0}{\alpha_{01},\alpha_{01},\cdots},\quad x_0=r,$$

を乗ずればよいが、次のラベル付けのときに固定したラベルを付けるので、木の順序を考えることになり、ν_0は無視する。

第2段.　ラベル付け:　上のように数えた順序無し林構造にラベルを付ける。各節点に x_j 個のラベルを与える与え方は

$$\binom{s}{x_1,\cdots,x_t}=\frac{s!}{\prod_{j=1}^{t}x_j!}$$

通りある。木の根の番号は予め指定・固定するので、仮想的な根 0 は考慮しない。各節点で、同じ型の順序木の根の番号の順序木は無意味であるから、結局

$$\prod_{j=1}^{t}\binom{x_k}{\alpha_{j1},\cdots,\alpha_{jk_j}}$$

通りのラベル付けが可能である。これらが等確率となるように、ランダムなラベル付けを行う。

第3段.　等確率性:　以上の準備により、1つのラベル付き順序無し林構造の確率は

$$e^{-t\lambda}\lambda^{s}\frac{1}{\prod_{j=1}^{t}x_j!}\prod_{j=1}^{t}\binom{x_j}{\alpha_{j1},\cdots,\alpha_{jk_j}}\Big/\frac{s!}{\prod_{j=1}^{t}x_j!}\prod_{j=1}^{t}\binom{x_j}{\alpha_{j1},\cdots,\alpha_{jk_j}}=e^{-t\lambda}\lambda^{s}/s!$$

で一定である。□

（注１）　$\eta(t,s)$ はいろいろの形に書ける。

$$\eta(t,s)=\frac{(t-s)(t+1)(t+2)\cdots(t+s-1)}{s!}=\binom{t+s-1}{s}-\binom{t+s-1}{s-1}$$
$$=\frac{r}{s+r}\binom{2s+r-1}{s}=\frac{r}{2s+r}\binom{2s+r}{s},$$

ただし $r=t-s,\ 1\leq r\leq t$ 。その特別な場合は、

$$\eta(t,0)=1,\quad t=1,2,\cdots,$$
$$\eta(t,1)=t-1,\quad t=1,2,\cdots,$$
$$\eta(t,2)=\frac{(t-2)(t+1)}{2},\quad t=2,3,\cdots,$$
$$\eta(s+1,s-1)=\eta(s+1,s)=\frac{1}{s+1}\binom{2s}{s},\quad s=1,2,\cdots,$$

などである。

（注２）

$$\zeta(t,s)=r(r+s)^{s-1},\quad r=t-s,\ 1\leq r\leq t,$$

である。ζ の特別な場合は、

$$\zeta(t,0)=1,\quad t=1,2,\cdots,$$
$$\zeta(t,1)=t-1,\quad t=1,2,\cdots,$$
$$\zeta(t,2)=(t-2)t,\quad t=2,3,\cdots,$$
$$\zeta(t,t-1)=t^{t-2},\quad t=1,2,\cdots,$$

などである。

文献

Consul, P.C. and Shenton, L.R. (1972), "Use of Lagrange expansion for generating discrete generalized probability distribution," *SIAM J. Appl. Math.*, 23, 239-248.

Consul, P.C. and Shenton, L.R. (1973), "Some interesting properties of Lagrangian distributions," *Comm. Statist. A - Theory and Meth.*, 2, 263-272.

Feller, W. (1968), *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., John Wiley, New York, N.Y.（特に Chap. XII, Section 5）.

Harris, T.E. (1963), *The Theory of Branchimg Processes*, Reprint 1989, Dover, N.Y.（特に Chap. 1）.

Lovász, L. (1979), *Combinatorial Problems and Exercises*, Akadémiai Kiadó, Budapest.（成嶋弘、土屋守正訳　東海大出版, 1988）.

渋谷政昭、宮脇典彦、住田潮 (1993), "分枝過程、待ち行列、酔歩、確率グラフ," 応用統計学会 *1993* 年度年会講演予稿集, 1-6.

Sibuya, M., Miyawaki, N. and Sumita, U. (1993), "Aspects of the Lagrangian probability distributions," *to appear in J. App. Prob.*

Takács, L. (1989), "Ballots, queues and random graphs," *J. Appl. Prob.*, 26, 103-112.

Takács, L. (1990), "On Cayley's formula for counting forests," *J. Combinatorial Theory A*, 53, 321-323.

Watson, H.W. and Galton, F. (1874), "On the probability of the extinction of families," *J. Anthropol. Inst. Great Britain and Ireland*, 4, 138-144.