数理解析研究所講究録
1193 巻 2001 年 64-78

Generalized Ewens' Sampling Formulas

清水 昭信 (Akinobu Shimizu)
名古屋市立大学 (Nagoya City University)

1 Introduction

$\{Z(t)\}_{t\geq 0}$ は、$\Lambda(dx)$ を Lévy measure とする subordinator とする。すなわち、$\Lambda(dx)$ は

$$\Lambda((1,+\infty))<+\infty, \quad \int_0^1 x\Lambda(dx)<+\infty$$

をみたす1次元開区間 $(0,+\infty)$ 上の測度であり、$\{Z(t)\}_{t\geq 0}$ は

$$\varphi_t(u):=E[e^{-uZ(t)}]=\exp\{t\int_0^{\infty}(e^{-ux}-1)\Lambda(dx)\}. \tag{1.1}$$

をみたす Lévy 過程である。
ここでは、更に

$$\Lambda((0,\infty))=\infty \tag{1.2}$$

を仮定する。つまり、複合ポアソン過程は除外する。
$\{Z(t)\}_{t\geq 0}$ の path は右連続（狭義）増加であり jump のみで増加し $P[Z(1)>0]=1$ だから (Sato [8])

$$W(t)=\frac{Z(t)}{Z(1)}, \quad 0\leq t\leq 1$$

は、path を固定すると区間 $[0,1]$ 上の 分布関数であり、区間 $[0,1]$ 上のランダムな離散的分布を定める。この分布についての sampling formula をあきらかにすることが目的である。

$$\Lambda(dx)=\theta x^{-1}e^{-x}dx$$

の場合、即ち $\{Z(t)\}_{t\geq 0}$ が Γ-process のケースは、集団遺伝学において Ewens のサンプリング公式として良く知られている。この公式のある意味での一般化を目的とする。

subordinator から定まる random discrete distribution について議論したものとして、Perman [5] がある。彼は、jump の 大きいほうから並べたランダムな点列の有限個の結合分布について議論している。

まず自然数 n の分割の定義をする。まず、自然数 n の分割の定義を述べよう。$\boldsymbol{N}$ は非負整数全体とする。数列 $\lambda:=(\lambda_1,\lambda_2,\cdots,\lambda_k)\in\boldsymbol{N}^k$ が $\sum_i\lambda_i=n,\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_k$ をみたすとき、λ を自然数 n の分割という。0 でない λ_i を、λ の部分という。2つの分割が後方の 0 の個数のみ異なっているとき、それらは同一とみなす。$\beta_j=$ the cardinality of $\{i:\lambda_i=j\},(j=1,2,\cdots,n)$ とおく。即ち、β_j は、n の分割に現われる j に等しい部

分の個数である。$\beta_k \geq 0, \sum_k k\beta_k = n$ が 成り立つ。$\beta := (\beta_1, \beta_2, \cdots, \beta_n)$ が自然数 n の一つの分割を定めるから、以後 β を自然数の分割という。
β を自然数の分割とし、$[0,1]^n$ の上で定義される次の関数 f_β を考える。

$$f_\beta(x_1, x_2, \cdots, x_n) = I_{A_\beta}(x_1, x_2, \cdots, x_n)$$

$$A_\beta = \bigcup_{\pi:\ \text{permutations}} B(\pi(1), \pi(2), \cdots, \pi(n))$$

$$B(1, 2, \cdots, n) = \{(x_1, x_2, \cdots, x_n) \in [0,1]^n : x_1 = x_2 = \cdots x_{i_1}, x_{i_1+1} = \cdots = x_{i_1+i_2},$$

$$\cdots, x_{\sum_{k=1}^{d-1} i_k+1} = \cdots = x_{\sum_{k=1}^{d} i_k}, \quad x_1, x_{i_1+1}, \cdots, x_{\sum_{k=1}^{d-1} i_k+1} \text{are} \quad \text{distinct}\},$$

とおく。ここで $i_1 + i_2 + \cdots + i_d = n$, $\#\{k : i_k = l\} = \beta_l$ としている。

自然数の分割は Young 図形と対応していることは、良く知られている。関数 f_β は, 座標 $x_1, x_2, \cdots, x_n$ が $\beta = (\beta_1, \beta_2, \cdots \beta_n)$ で定まる Young 図形をつくる集合の定義関数である。
$d = d(\beta) = \sum_i \beta_i$ を Young 図形の長さということにする。つまり、相異なる点（遺伝学では、the number of alleles）の個数である。

$\Pi_W(d\mu)$ を上の $W(t)$ が導く $\mathcal{P}([0,1])$ 上の確率測度とするとき、次の期待値を計算することがこのノートの第一の目的である。

$$I_{n,\beta} = \int < f_\beta, \mu^n > \Pi_W(d\mu)$$

ここで、

$$< f_\beta, \mu^n > = \int \cdots \int f_\beta(x_1, x_2, \cdots, x_n)\mu(dx_1)\mu(dx_2)\cdots\mu(dx_n).$$

つまり、ランダムな分布 μ を一つ固定して、n 個独立にサンプルした点が β タイプの自然数の分割 (Young 図形) を形成する確率を考え、その期待値を問題とする。

結果は次のようになる。

Theorem 1.

$$I_{n,\beta} = \frac{n!}{\prod_{j=1}^n (j!)^{\beta_j} \beta_j!} \frac{1}{\Gamma(n)} \times$$

$$\int_0^\infty x^{n-1} dx \prod_{k=1}^n (\int_0^\infty z^k e^{-xz} \Lambda(dz))^{\beta_k} \exp\{\int_0^\infty (e^{-xz} - 1)\Lambda(dz)\}.$$

(証明は Section 3).

Remark 1.1 *Appendix* の合成関数の高階導関数の結果を用いると、

$$\begin{aligned}\sum_{\beta} I_{n,\beta} &= \frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}dx(-1)^n\frac{d^n}{dx^n}\exp\{\int_0^\infty(e^{-xz}-1)\Lambda(dz)\}\\ &= \frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}dxE[Z(1)^n e^{-xZ(1)}]\\ &= 1\end{aligned}$$

となることが分かる。この計算は、定理1の右辺が有限値であることも示している。

確率変数 $< f_\beta, \mu^n >$ の高次モーメントを求めることは、一般に困難であり、新たなアイディアを必要とすると思われるが、$\beta_n = 1$ のときは、定理1と同様の計算で求めることが出来る。その結果が次の命題である。

Theorem 2.
$\beta = (0, 0, \cdots, 0, 1)$ つまり、$\beta_1 = \cdots = \beta_{n-1} = 0, \beta_n = 1$ とする。このとき、

$$\int < f_\beta, \mu^n >^m \Pi_W(d\mu) = \sum_{\beta'} \frac{m!}{\prod_{j=1}^m (j!)^{\beta'_j}\beta'_j!}\frac{1}{\Gamma(nm)}\times$$

$$\int_0^\infty x^{nm-1}dx\prod_{k=1}^m(\int_0^\infty z^{nk}e^{-xz}\Lambda(dz))^{\beta'_k}\exp\{\int_0^\infty(e^{-xz}-1)\Lambda(dz)\}.$$

ここで、β' は、自然数 m の分割であり、$\sum_{\beta'}$ はこの分割全体についての和である。

(証明は Section 4.)

上の定理1において、特に $\Lambda(dz) = cz^{-\alpha-1}e^{-\lambda z}dz, 0 \le \alpha < 1, \lambda > 0$ であるとき

Corollary 3.

$$I_{n,\beta} = \frac{n!}{\prod_{j=1}^n (j!)^{\beta_j}\beta_j!}\frac{1}{\Gamma(n)}\times$$

$$\prod_{l=1}^n \Gamma(l-\alpha)^{\beta_l} \times c^d \times$$

$$\times\int_0^\infty x^{n-1}(\lambda+x)^{\alpha d-n}\exp\{-c\Gamma(1-\alpha)\int_0^x(\lambda+z)^{\alpha-1}dz\}dx,$$

where $d = \sum_k \beta_k$.

(証明略。Theorem 1 を用いた簡単な計算により示せる。)

Example 1.
上の系において、$\alpha = 0, \lambda = 1, c = \theta$ のとき、系3の右辺の積分は計算できて、

$$I_{n,\beta} = \frac{n!}{\theta(\theta+1)\cdots(\theta+n-1)}\prod_{j=1}^n\frac{\theta^{\beta_j}}{j^{\beta_j}\beta_j!}$$

となる。これが、Ewensのサンプリング公式である [2]。

Example 2.
$\lambda=0, c=\frac{\alpha}{\Gamma(1-\alpha)}$ のとき、

$$I_{n,\beta}=\frac{n!}{\prod_{j=1}^{n}(j!)^{\beta_j}\beta_j!}\frac{\alpha^{d-1}\Gamma(d)}{\Gamma(n)}\prod_{j=1}^{n}(\frac{\Gamma(j-\alpha)}{\Gamma(1-\alpha)})^{\beta_j}.$$

Example 3.
Example 1と同様に,subordinatorはΓ-processの場合を考える。集団遺伝学の無限中立対立遺伝子モデルの定常分布の場合である。この場合にTheorem 2を適用する。
$n=2$の場合のみを考察する。$\beta=(0,1)$であり、関数f_βは、2次元正方形$[0,1]^2$の対角線の定義関数である。このとき、$<f_\beta,\mu^2>$はホモ接合頻度を表す確率変数である。この期待値が$\frac{1}{1+\theta}$であることは集団遺伝学でよく知られている。もちろん$1-<f_\beta,\mu^2>$はヘテロ接合頻度を表す。Theorem 2を適用するとホモ接合頻度の高次モーメントが分かることになる。

$$E[<f_\beta,\mu^2>^m]=\sum_{\beta'}\frac{m!}{\prod_j(j!)^{\beta'_j}\beta'_j!}\frac{\Gamma(\theta)}{\Gamma(\theta+2m)}\theta^{|\beta'|}\prod_k((2k-1)!)^{\beta'_k}.$$

ここで、さらに$m=2$とすると、β'は$\beta'_1=2,\beta'_2=0$のケースと$\beta'_1=0,\beta'_2=1$の2通りしかなく

$$E[<f_\beta,\mu^2>^2]=\frac{\theta+6}{(1+\theta)(2+\theta)(3+\theta)}$$

となり、従ってホモ接合頻度の分散は、

$$V[<f_\beta,\mu^2>]=\frac{2\theta}{(1+\theta)^2(2+\theta)(3+\theta)}$$

となる。これは、F.M.Stewart(1976)の結果と一致する(Kimura [4]による)。
ホモ接合頻度の分散とヘテロ接合頻度の分散は同じであることに注意すること。

$(\Omega_i,B_i,P_i), i=1,2,$を確率空間とし$(\Omega,B,P)$をこの二つの確率空間の直積とする。
$\omega=(\omega_1,\omega_2)\in\Omega,\omega_i\in\Omega_i$とおく。
Lévy measure

$$\Lambda(dz)=c\alpha z^{-\alpha-1}e^{-z}dz,$$

をもつsubordinator$\{Z(t)(\omega_1)\}$と、分布が

$$E_2[e^{-uT(\omega_2)}]=\exp\{\int_0^\infty(e^{-uz}-1)\frac{\theta}{\alpha z}e^{-c\Gamma(1-\alpha)z}dz\},$$

で与えられる確率変数$T(\omega_2)$を考える。このとき、
$Z(tT(\omega_2))(\omega_1)/Z(T(\omega_2))(\omega_1)$は、$(\omega_1,\omega_2)$を固定したとき、区間$[0,1]$上のランダムな分布関数であり、かつ、

$$E_1[e^{-uZ(tT(\omega_2))(\omega_1)}]=\exp\{tT(\omega_2)\int_0^\infty(e^{-uz}-1)c\alpha z^{-\alpha-1}e^{-z}dz\}$$

である。区間$[0,1]$上のこのランダムな離散分布についてのsampling formulaを議論する。
Corollary 3が適用できて次の結果を得る。

Theorem 4.

$$I_{n,\beta} = \frac{n!}{\prod_{j=1}^n (j!)^{\beta_j}\beta_j!}\times$$
$$\frac{(\theta+\alpha)(\theta+2\alpha)\cdots(\theta+(d-1)\alpha)\Gamma(\theta+1)}{\Gamma(\theta+n)}\prod_{j=1}^n(\frac{\Gamma(j-\alpha)}{\Gamma(1-\alpha)})^{\beta_j}.$$

この定理が Pitman-Yor[6] の Proposition 51 である。

(証明略。$T(\omega_2)$ の分布の密度関数が

$$\Gamma(\frac{\theta}{\alpha})^{-1}(c\Gamma(1-\alpha))^{\frac{\theta}{\alpha}}z^{\frac{\theta}{\alpha}-1}e^{-c\Gamma(1-\alpha)z}$$

であることに注意せよ)

Remark 1.2 $Z(tT(\omega_2))(\omega_1)/Z(T(\omega_2))(\omega_1),\quad 0\le t\le 1,$ の *jump* を大きい方から順に並べた減少列のランダム列の分布が *Pitman-Yor* の *two-parameter Poisson-Dirichlet distribution* である。

$\mathbf{K}_n = \sum_\beta I_{n,\beta}|\beta|$ とおく。これは、Young 図形の長さの平均の期待値である。これが n を大きくするときの漸近挙動を明らかにしたい。この問題を考えることが、このノートの第二の目的である。
Example 1 の場合は良く知られていて

$$\mathbf{K}_n = \theta\sum_{i=1}^n \frac{1}{\theta+n-1} \sim \theta\log n,\quad n\to\infty$$

である。一般の場合について次の命題が成り立つ。

Theorem 5.
(1) ある $0<\alpha<1$ と正数 c_1 に対して、

$$\int_0^\infty e^{-zu}z\Lambda(dz) \sim \frac{c_1}{u^{1-\alpha}}, u\to\infty,$$

が成り立つならば、$E[\frac{1}{Z(1)^\alpha}]<\infty$ であり、かつ

$$\mathbf{K}_n \sim c_1\frac{1}{\alpha}E[\frac{1}{Z(1)^\alpha}]\, n^\alpha,\ n\to\infty$$

が成り立つ。
(2) ある正数 c_1 に対して、

$$\int_0^\infty e^{-zu}z\Lambda(dz) \sim \frac{c_1}{u}, u\to\infty,$$

と

$$\int_1^\infty \log z\Lambda(dz) < +\infty,$$

が成り立つならば、$E[|\log Z(1)|] < +\infty$ が成り立ち、かつ

$$\mathbf{K}_n \sim c_1 \log n,\ n \to \infty.$$

となる。
(証明は Section 5)

2 Preliminaries

$[0,1]^n$ の部分集合 A_β^N を次のように定義する。

$$A_\beta^N = \{\mathbf{x}_N = (x_1^N, \cdots, x_n^N) : Nx_i^N \text{ is an integer for each } i, \mathbf{x}_N \in A_\beta\}$$

$[0,1]^n$ の部分集合 $C_{N,\beta}$ を次のように定義し、関数 $f_{N,\beta}$ は $C_{N,\beta}$ の定義関数とする。

$$C_{N,\beta} = \bigcup_{\mathbf{x}_N \in A_\beta^N} \prod_{i=1}^{n} (x_i^N, x_i^N + \frac{1}{N}]$$

$$f_{N,\beta}(x_1, \cdots, x_n) = I_{C_{N,\beta}}(x_1, \cdots, x_n).$$

このとき、次の補題の成り立つことが分かる。

Lemma 2.1.
The sequence $\{f_{N,\beta}\}$ converges to f_β boudedly, as $n \to +\infty$.
Proof. A_β に属する各点では、$\{f_{N,\beta}\}$ の極限が 1 であることをいう。$\mathbf{x} = (x_1, x_2, \cdots, x_n) \in A_\beta$ としよう。$x_1, x_2, \cdots, x_n$ のなかの相異なる 2 つの実数の距離の最小値より $2/N$ が小さくなるように N を大きくすれば $f_{N,\beta}(\mathbf{x}) = 1$ となる。
各 x_i に対して、$x_i \geq k_i/N$ となるような最大の k_i/N を選ぶと、
点 $(k_1/N, k_2/N, \cdots, k_n/N) \in A_\beta^N$ となることを示す。上の定め方から、$x_i = x_j$ であれば $k_i/N = k_j/N$ であることは明らか。$x_i \neq x_j$ であるにもかかわらず、$k_i/N = k_j/N$ となったとしよう。

$$2/N < |x_i - x_j| \leq |x_i - k_i/N| + |k_i/N - k_j/N| + |k_j/N - x_j| < 2/N$$

となって矛盾である。
これで十分である。実際、$\mathbf{x} = (x_1, x_2, \cdots, x_n) \notin A_\beta$ としよう。$\mathbf{x} \in A_{\beta'}, \beta' \neq \beta$ となる β' が存在し、$f_{N,\beta} + f_{N,\beta'} \leq 1, \lim_{N\to\infty} f_{N,\beta'}(\mathbf{x}) = 1$ だから、$\lim_{N\to\infty} f_{N,\beta}(\mathbf{x}) = 0$ となる。

この補題によって次の命題を得ることが出来る。
Lemma 2.2.

$$I_n = \lim_{N\to+\infty} \int < f_{N,\beta}, \mu^n > \Pi_W(d\mu)$$

$$= \lim_{N\to+\infty} \frac{n!}{(1!)^{\beta_1}(2!)^{\beta_2}\cdots(n!)^{\beta_n}} \sum{}^{*}\, E^W[\prod_{j=1}^{N} (W_j^{(N)})^{m_j}],$$

where $W_j^{(N)} = W(\frac{j}{N}) - W(\frac{j-1}{N}), j = 1, 2, \cdots, N$. Here, $\sum^*$ means the summation on $\mathbf{m} = (m_1, m_2, \cdots, m_N)$ with nonnegative integers m_i such that $\sum_{j=1}^N m_j = n$, and $\#\{j : m_j = k\} = \beta_k, k = 1, 2, \cdots, n$.

Proof. 最初の等号は、有界収束定理による。二番目の等号を示すには、次の注意を参照せよ。

Remark. 番号のついた N 個の箱に番号のある n 個のボールをいれる。但し、j 個のボールをいれる箱の数は $\beta_j (j = 1, 2, \cdots n)$ 個とする。$\sum^*$ は、N 個の各箱にいれるボールの個数を指定するあらゆる方法についての和をあらわす。$\frac{n!}{(1!)^{\beta_1}(2!)^{\beta_2}\cdots(n!)^{\beta_n}}$ は、n 個のボールを指定された箱にいれる方法の総数である。
(Ethier-Kurtz[1]448 頁を参照せよ。ここに 同じ量があらわれるが、この量を導く方法は、ここでは、Ethier-Kurtz の方法と異なる。)

$$J_N = \frac{n!}{(1!)^{\beta_1}(2!)^{\beta_2}\cdots(n!)^{\beta_n}} \sum{}^* E^W[\prod_{j=1}^N (W_j^{(N)})^{m_j}]$$

とおく。

Lemma 2.3.

$$\begin{aligned} J_N &= \frac{n!}{(1!)^{\beta_1}(2!)^{\beta_2}\cdots(n!)^{\beta_n}} \frac{1}{\Gamma(n)} \int_0^\infty u^{n-1} du \sum{}^* \prod_{k=1}^N (-1)^{m_k} \varphi_{\frac{1}{N}}^{(m_k)}(u) \\ &= \frac{n!}{\prod_{j=1}^n (j!)^{\beta_j}} \frac{N!}{\prod_{j=1}^n \beta_j!(N-d)!} \frac{1}{\Gamma(n)} \int_0^\infty u^{n-1} du \prod_{j=1}^n \{(-1)^j \varphi_{\frac{1}{N}}^{(j)}(u)\}^{\beta_j} \times \{\varphi_{\frac{1}{N}}(u)\}^{N-d}. \end{aligned}$$

Proof.

$$\begin{aligned} E^W[\prod_{j=1}^N (W_j^{(N)})^{m_j}] &= E[\frac{1}{Z(1)^n}(Z(\frac{1}{N}))^{m_1} \cdots (Z(1) - Z(\frac{N-1}{N}))^{m_N}] \\ &= \frac{1}{\Gamma(n)} \int_0^\infty u^{n-1} du E[e^{-Z(1)u}(Z(\frac{1}{N}))^{m_1} \cdots (Z(1) - Z(\frac{N-1}{N}))^{m_N}] \\ &= \frac{1}{\Gamma(n)} \int_0^\infty u^{n-1} du \prod_{k=1}^N E[(Z(\frac{k}{N}) - Z(\frac{k-1}{N}))^{m_k} e^{-(Z(\frac{k}{N}) - Z(\frac{k-1}{N}))u}] \\ &= \frac{1}{\Gamma(n)} \int_0^\infty u^{n-1} du \prod_{k=1}^N (-1)^{m_k} \varphi_{\frac{1}{N}}^{m_k}(u) \\ &= \frac{1}{\Gamma(n)} \int_0^\infty u^{n-1} du \prod_{j=1}^n \{(-1)^j \varphi_{\frac{1}{N}}^{(j)}(u)\}^{\beta_j} \{\varphi_{\frac{1}{N}}(u)\}^{N-d}. \end{aligned}$$

$$\varphi_t(u) = \exp\{t\psi(-u)\}$$

$$\varphi(u) := \varphi_1(u) = \exp\{\psi(-u)\}$$

とおく。ここで、

$$\psi(-u)=\int_0^\infty (e^{-uz}-1)\Lambda(dz)$$

である。このとき、

Lemma 2.4.

$$(-1)^k\varphi^{(k)}(u)=\sum_{\beta^k\in\mathcal{P}(k)} C(\beta^{(k)})(\psi'(-u))^{\beta_1^{(k)}}(\psi''(-u))^{\beta_2^{(k)}}\cdots(\psi^{(k)}(-u))^{\beta_k^{(k)}}\varphi(u),$$

where $\mathcal{P}(k)$ stands for the set of partitions β^k of integer k,and $C(\beta^{(k)})$ are positive constants.

$$(-1)^k\varphi^{(k)}_{\frac{1}{N}}(u)=\sum_{\beta^k\in\mathcal{P}(k)} C(\beta^{(k)})(\frac{1}{N})^{d_k}(\psi'(-u))^{\beta_1^{(k)}}(\psi''(-u))^{\beta_2^{(k)}}\cdots(\psi^{(k)}(-u))^{\beta_k^{(k)}}\varphi_{\frac{1}{N}}(u),$$

$$=\frac{1}{N}\psi^{(k)}(-u)\varphi_{\frac{1}{N}}(u)+\frac{1}{N^2}h_k(u),$$

where

$$h_k(u)\le\sum_{\beta^{(k)}\in\mathcal{P}(k)} c(\beta^{(k)})(\psi'(-u))^{\beta_1^{(k)}}\cdots(\psi^{(k)}(-u))^{\beta_k^{(k)}}\varphi_{\frac{1}{N}}(u),$$

and $d_k=\sum_i\beta_i^{(k)}$.
(Note that $\beta_k^{(k)}=1$ implies $C(\beta^{(k)})=1$, and that $\sum_j j\beta_j^{(k)}=k$.)
Proof. Appendix 合成関数の高階導関数による。

Lemma 2.5.

$$\{(-1)^k\varphi^{(k)}_{\frac{1}{N}}(u)\}^{\beta_k}=(\frac{1}{N})^{\beta_k}\{\psi^{(k)}(-u)\varphi_{\frac{1}{N}}(u)\}^{\beta_k}+\frac{1}{N^{\beta_k+1}}H_k(u),$$

where

$$H_k(u)\le\sum c(l)(\psi'(-u))^{l_1}(\psi''(-u))^{l_2}\cdots(\psi^{(k)}(-u))^{l_k}(\varphi_{\frac{1}{N}}(u))^{\beta_k},$$

where $\sum_j jl_j=k\beta_k$, and

$$\prod_{k=1}^n\{(-1)^k\varphi^{(k)}_{\frac{1}{N}}(u)\}^{\beta_k}(\varphi_{\frac{1}{N}}(u))^{N-d}=(\frac{1}{N})^d\prod_{k=1}^n(\psi^{(k)}(-u))^{\beta_k}\varphi(u)+(\frac{1}{N})^{d+1}H(u),$$

where

$$H(u)\le\sum const.(\psi'(-u))^{\beta'_1}(\psi''(-u))^{\beta'_2}\cdots(\psi^{(n)}(-u))^{\beta'_n}\varphi(u),$$

and $(\beta'_1,\beta'_2,\cdots,\beta'_n)$ are partitions of n.

3 Proof of Theorem 1

Lemma 2.3 と Lemma 2.5 により、

$$J_N = \frac{n!}{\prod_{j=1}^{n}(j!)^{\beta_j}\beta_j!}\frac{N!}{(N-d)!}[(\frac{1}{N})^d\frac{1}{\Gamma(n)}\int_0^\infty u^{n-1}du\prod_{k=1}^{n}(\psi^{(k)}(-u))^{\beta_k}\varphi(u)$$

$$+(\frac{1}{N})^{d+1}\frac{1}{\Gamma(n)}\int_0^\infty u^{n-1}duH(u)] \tag{3.1}$$

を得る。
この式の右辺第2項が有限であることをまず注意しよう。
このためには、Lemma2.5 の最後の式によれば

$$\int_0^\infty u^{n-1}du(\psi'(-u))^{\beta_1'}(\psi''(-u))^{\beta_2'}\cdots(\psi^{(n)}(-u))^{\beta_n'}\varphi(u) < +\infty$$

示せば十分である。Lemma 2.4 の最初の式を用いて、

$$\begin{aligned}\int_0^\infty u^{n-1}du\prod_{j=1}^{n}(\psi^{(j)}(-u))^{\beta_j'}\varphi(u) &\leq c(\beta')^{-1}\int_0^\infty u^{n-1}du(-1)^n\varphi^{(n)}(u)\\ &= c(\beta')^{-1}\int_0^\infty u^{n-1}duE[Z(1)^ne^{-uZ(1)}]\\ &= c(\beta')^{-1}E[\frac{Z(1)^n}{Z(1)^n}]\\ &= c(\beta')^{-1}\end{aligned}$$

Lemma 2.2 によって、$I_N = \lim_{N\to\infty} J_N$ であるが、(3.1) 式の右辺第2項は 上の議論によって無視できることになり、従って、

$$I_N = \frac{n!}{\prod_{j=1}^{n}(j!)^{\beta_j}\beta_j!}\frac{1}{\Gamma(n)}\int_0^\infty u^{n-1}du\prod_{k=1}^{n}(\psi^{(k)}(-u))^{\beta_k}\varphi(u)$$

となり、Theorem 1 の証明は 完結する。

4 Proof of Theorem 2.

Theorem 1 の証明と同じように
$f_{N,\beta} = \sum_{i=1}^{N} I_{A_i^N\times A_i^N\times\cdots A_i^N}(x_1,x_2,\cdots,x_n)$, where $A_i^N = (\frac{i-1}{N},\frac{i}{N}]$.
とおく。以後、簡単のために、A_i^N を単に A_i と書く。

$$\begin{aligned}\int <f_\beta,\mu^n>^m \Pi_W(d\mu) &= \lim_{N\to\infty}\int <f_{N,\beta},\mu^n>^m \Pi_W(d\mu)\\ &= \lim_{N\to\infty}\int \Big(\sum_{i=1}^{N}\mu(A_i)^n\Big)^m\Pi_W(d\mu)\end{aligned}$$

であり、

$$
\begin{aligned}
\int (\sum_{i=1}^{N} \mu(A_i)^n)^m \Pi_W(d\mu) &= E[(\sum_{i=1}^{N} (\frac{Z(A_i)}{Z(1)})^n)^m] \\
&= E[\frac{1}{\Gamma(nm)} \int_0^\infty x^{nm-1} dx (\sum_{i=1}^{N} Z(A_i)^n)^m e^{-Z(1)x}] \\
&= \sum_{m_1, m_2, \cdots, m_N} \frac{m!}{\prod_{j=1}^{N} m_j!} \frac{1}{\Gamma(nm)} \int_0^\infty x^{nm-1} dx E[\prod_{j=1}^{N} Z(A_j)^{nm_j} e^{-Z(1)x}] \\
&= \sum_{m_1, m_2, \cdots, m_N} \frac{m!}{\prod_{j=1}^{N} m_j!} \frac{1}{\Gamma(nm)} \int_0^\infty x^{nm-1} dx \prod_{j=1}^{N} E[Z(A_j)^{nm_j} e^{-Z(A_j)x}] \\
&= \sum_{m_1, m_2, \cdots, m_N} \frac{m!}{\prod_{j=1}^{N} m_j!} \frac{1}{\Gamma(nm)} \int_0^\infty x^{nm-1} dx (-1)^{nm} \prod_{j=1}^{N} \varphi_{\frac{1}{N}}^{(nm_i)}(x) \\
&\sim \sum_{\mathbf{m}} \frac{m!}{\prod_{j=1}^{N} m_j!} \frac{1}{\Gamma(nm)} \int_0^\infty x^{nm-1} dx \\
&\times (\frac{1}{N})^d \prod_{j=1}^{N} (\int_0^\infty z^{nm_i} e^{-xz} \Lambda(dz)) \varphi(x), \\
&= \sum_{\beta'} \frac{m!}{\prod_{j=1}^{m} (j!)^{\beta'_j} \beta'_j!} \frac{N!}{(N-|\beta'|)! N^d} \frac{1}{\Gamma(nm)} \int_0^\infty x^{nm-1} dx \\
&\times \prod_{k=1}^{m} (\int_0^\infty z^{nk} e^{-xz} \Lambda(dz))^{\beta'_k} \exp(\int_0^\infty (e^{-xz}-1) \Lambda(dz))
\end{aligned}
$$

従って、Theorem 2 の証明は完結する。

5 K_n の漸近挙動

まず次の補題が成り立つことを示そう。

Lemma 5.1.

$$\mathbf{K}_n = \frac{1}{\Gamma(n)} \int_0^\infty x^{n-1} \int_0^x (\int_0^\infty e^{-zu} z \Lambda(dz)) du E[Z(1)^n e^{-xZ(1)}] dx. \tag{5.1}$$

Proof.

Remark 1.1 の説明と同様の考察により

$$\frac{1}{\Gamma(n)} \int_0^\infty x^{n-1} dx (-1)^n \frac{d^n}{dx^n} \exp\{\alpha \int_0^\infty (e^{-xz}-1) \Lambda(dz)\} = 1$$

が成り立つから、Theorem 1 により、

$$\frac{1}{\Gamma(n)} \int_0^\infty x^{n-1} dx \sum_{\beta} \frac{n! \alpha^{|\beta|}}{\prod_{j=1}^{n} (j!)^{\beta_j} \beta_j!} \prod_{k=1}^{n} (\int_0^\infty z^k e^{-xz} \Lambda(dz))^{\beta_k} \times \exp\{\alpha \int_0^\infty (e^{-xz}-1) \Lambda(dz)\} = 1$$

を得る。この式の両辺を α で微分して $\alpha=1$ とおくと

$$\frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}dx\sum_\beta \frac{n!|\beta|}{\prod_{j=1}^n (j!)^{\beta_j}\beta_j!}\prod_{k=1}^n (\int_0^\infty z^k e^{-xz}\Lambda(dz))^{\beta_k}\times \exp\{\int_0^\infty (e^{-xz}-1)\Lambda(dz)\} \quad +$$

$$\frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}dx\sum_\beta \frac{n!}{\prod_{j=1}^n (j!)^{\beta_j}\beta_j!}\prod_{k=1}^n (\int_0^\infty z^k e^{-xz}\Lambda(dz))^{\beta_k}\exp\{\int_0^\infty (e^{-xz}-1)\Lambda(dz)\}\int_0^\infty (e^{-xz}-1)\Lambda(dz) \quad = \quad ($$

この左辺の第一項は $\mathbf{K}_n$ に等しいこと、および 左辺第２項は

$$\frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}dx\int_0^\infty (e^{-xz}-1)\Lambda(dz)(-1)^n\frac{d^n}{dx^n}\exp\{\int_0^\infty (e^{-xz}-1)\Lambda(dz)\}$$

$$= -\frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}dx\int_0^x (\int_0^\infty e^{-zu}z\Lambda(dz))duE[Z(1)^n e^{-xZ(1)}]$$

等しいから結論を得る。

簡単のために

$$h(x,u)=\frac{1}{\Gamma(n)}x^{n-1}(\int_0^\infty e^{-zu}z\Lambda(dz))E[Z(1)^n e^{-xZ(1)}]$$

とおく。

$$\mathbf{K}_n = \int_0^M (\int_0^x h(x,u)du)dx+\int_M^\infty\int_0^M h(x,u)dudx+\int_M^\infty\int_M^x h(x,u)dudx$$
$$= K_n^1+K_n^2+K_n^3 \tag{5.2}$$

と３つの部分に分ける。どの部分も非負である。

Lévy measure $\Lambda(dz)$ が ある $0<\alpha<1$ に対して条件

$$\int_0^\infty e^{-uz}z\Lambda(dz)\sim \frac{c_1}{u^{1-\alpha}}, n\to\infty \tag{5.3}$$

をみたす場合を 考えよう。
(5.3) により、任意の正数 ε にたいして、ある正数 M が存在し、$u\geq M$ なら

$$(1-\varepsilon)c_1\frac{1}{u^{1-\alpha}}\leq \int_0^\infty e^{-uz}z\Lambda(dz)\leq (1+\varepsilon)c_1\frac{1}{u^{1-\alpha}} \tag{5.4}$$

が成り立つ。このとき、$\dfrac{1}{Z(1)^\alpha}$ の期待値が有限であることをまず示そう。

Lemma 5.2.
ある $0<\alpha<1$ に対して (5.3) が成り立つとき、$E[\dfrac{1}{Z(1)^\alpha}]<+\infty$ である。

Proof.
任意の正数 ε を固定する。この ε に対して (5.4) がなりたつような M を使って、上のように $\mathbf{K}_n$ を３個の部分に分ける。$\mathbf{K}_n$ は n 以下であるから、n を固定したとき K_n^3 はもちろん有限である。この K_n^3 を下から評価する。

$$K_n^3\frac{\Gamma(n)}{(1-\varepsilon)c_1} \quad \geq \quad \int_M^\infty x^{n-1}\frac{1}{\alpha}\{x^\alpha - M^\alpha\}E[Z(1)^n e^{-xZ(1)}]dx$$

$$
\begin{aligned}
&\geq \frac{1}{\alpha}\{\int_M^\infty x^{n+\alpha-1}E[Z(1)^n e^{-xZ(1)}]dx - M^\alpha \int_M^\infty x^{n-1}E[Z(1)^n e^{-xZ(1)}]dx\} \\
&\geq \frac{1}{\alpha}\{E[\int_{MZ(1)}^\infty u^{n+\alpha-1}e^{-u}du\frac{1}{Z(1)^\alpha}] - M^\alpha\} \\
&\geq \frac{1}{\alpha}\{E[\int_{MZ(1)}^\infty u^{n+\alpha-1}e^{-u}du\frac{1}{Z(1)^\alpha} : Z(1) \leq \frac{1}{M}] - M^\alpha\} \\
&\geq \frac{1}{\alpha}\{\int_1^\infty u^{n+\alpha-1}e^{-u}duE[\frac{1}{Z(1)^\alpha} : Z(1) \leq \frac{1}{M}] - M^\alpha\}
\end{aligned}
$$

この式から、$E[\frac{1}{Z(1)^\alpha}] < +\infty$ であることが分かる。

Lemma 5.3

$$
\frac{1}{\Gamma(n)}\int_M^\infty x^{n+\alpha-1}E[Z(1)^n e^{-xZ(1)}]dx \sim E[\frac{1}{Z(1)^\alpha}]\, n^\alpha, \quad n \to \infty. \tag{5.5}
$$

Proof.

$$
\begin{aligned}
\frac{1}{\Gamma(n)}\int_0^M x^{n+\alpha-1}E[Z(1)^n e^{-xZ(1)}]dx &\leq \frac{1}{\Gamma(n)}M^\alpha \int_0^M x^{n-1}E[Z(1)^n e^{-xZ(1)}]dx \\
&\leq M^\alpha \frac{1}{\Gamma(n)}\int_0^\infty x^{n-1}E[Z(1)^n e^{-xZ(1)}]dx \\
&\leq M^\alpha
\end{aligned}
$$

であることと

$$
\begin{aligned}
\frac{1}{\Gamma(n)}\int_0^\infty x^{n+\alpha-1}E[Z(1)^n e^{-xZ(1)}]dx &= \frac{1}{\Gamma(n)}E[\int_0^\infty u^{n+\alpha-1}e^{-u}du\frac{1}{Z(1)^\alpha}] \\
&= \frac{\Gamma(n+\alpha)}{\Gamma(n)}E[\frac{1}{Z(1)^\alpha}] \\
&\sim E[\frac{1}{Z(1)^\alpha}]\, n^\alpha, \quad n \to \infty
\end{aligned}
$$

から、求める結論を得る。この議論の最後の段階で、Stirling の公式を用いている。

(5.2) により、次の補題を示せば、Theorem 5,(1) の証明が完結する。

Lemma 5.4

$$
K_n^1 + K_n^2 \leq \int_0^M \Big(\int_0^\infty e^{-zu}z\Lambda(dz)\Big)du < +\infty, \tag{5.6}
$$

$$
K_n^3 \sim \frac{c_1}{\alpha}E[\frac{1}{Z(1)^\alpha}]\, n^\alpha,\ n \to \infty. \tag{5.7}
$$

Proof.

(5.6) の成り立つことはあきらか。

(5.4) を用いて、

$$
(1-\varepsilon)c_1\frac{1}{\Gamma(n)}\int_M^\infty x^{n-1}\Big(\int_M^x \frac{1}{u^{1-\alpha}}du\Big)E[Z(1)^n e^{-xZ(1)}]dx
$$

$$\leq \quad K_n^3 \quad \leq \quad (1+\varepsilon)c_1 \frac{1}{\Gamma(n)} \int_M^\infty x^{n-1} \Big(\int_M^x \frac{1}{u^{1-\alpha}} du \Big) E[Z(1)^n e^{-xZ(1)}] dx.$$

ここで、

$$\frac{1}{\Gamma(n)} \int_M^\infty x^{n-1} \Big(\int_M^x \frac{1}{u^{1-\alpha}} du \Big) E[Z(1)^n e^{-xZ(1)}] dx$$
$$= \frac{1}{\alpha} \frac{1}{\Gamma(n)} \Big\{ \Big(\int_M^\infty x^{n+\alpha-1} E[Z(1)^n e^{-xZ(1)}] dx \Big) - M^\alpha \quad \times \quad \int_M^\infty x^{n-1} E[Z(1)^n e^{-xZ(1)}] dx \Big) \Big\}$$

$$0 \leq \frac{1}{\Gamma(n)} M^\alpha \int_M^\infty x^{n-1} E[Z(1)^n e^{-xZ(1)}] dx \leq M^\alpha$$

だから、(5.4) によって、(5.7) の成り立つことが分かる。証明終わり。

Lévy measure $\Lambda(dz)$ が 2 つの条件

$$\int_0^\infty e^{-uz} z \Lambda(dz) \sim \frac{c_1}{u}, n \to \infty, \tag{5.8}$$

$$\int_1^\infty \log z \Lambda(dz) < +\infty, \tag{5.9}$$

をみたす場合を考える。
(5.8) により、任意の正数 ε にたいして、ある正数 M が存在し、$u \geq M$ なら

$$(1-\varepsilon)c_1 \frac{1}{u} \leq \int_0^\infty e^{-uz} z \Lambda(dz) \leq (1+\varepsilon)c_1 \frac{1}{u} \tag{5.10}$$

が成り立つ。
次の補題が成り立つことは明らかである。
Lemma 5.5

$$\frac{(1-\varepsilon)c_1}{\Gamma(n)} \Big(\int_M^\infty x^{n-1} \log x E[Z(1)^n e^{-xZ(1)}] dx \Big) - (1-\varepsilon)c_1 \log M \leq$$
$$\leq \quad K_n^3 \quad \leq \quad \frac{(1+\varepsilon)c_1}{\Gamma(n)} \int_M^\infty x^{n-1} \log x E[Z(1)^n e^{-xZ(1)}] dx. \tag{5.11}$$

Lemma 5.6
(5.8) と (5.9) を仮定するとき、

$$E[|\log Z(1)|] < +\infty, \tag{5.12}$$

が成り立つ。

Proof.

$$E[|\log Z(1)| : Z(1) \geq 1] \leq E[\log(e \vee Z(1))]$$

であるから、Sato[8](p. 159, Theorem 25.3) により、条件 (5.9) より

$$E[|\log Z(1)| : Z(1) \geq 1] < \infty$$

となる。他方、Lemma 5.5,(5.11) により、

$$+\infty > K_n^3 + (1-\varepsilon)c_1 \log M \geq \frac{(1-\varepsilon)c_1}{\Gamma(n)} \int_M^\infty x^{n-1} \log x E[Z(1)^n e^{-xZ(1)}]dx,$$

である。ここで、$xZ(1) = t$ とおき、$dx = \frac{1}{Z(1)}dt$ だから、($M > 1$ としてよいから)

$$\begin{aligned}
\int_M^\infty x^{n-1} \log x E[Z(1)^n e^{-xZ(1)}]dx &= E[\int_{MZ(1)}^\infty t^{n-1} \log \frac{t}{Z(1)} e^{-t} dt] \\
&\geq E[\int_{MZ(1)}^\infty t^{n-1} \log \frac{t}{Z(1)} e^{-t} dt : Z(1) < 1] \\
&\geq E[\int_{MZ(1)}^\infty t^{n-1} \log t e^{-t} dt : Z(1) < 1] \\
&+ \int_M^\infty t^{n-1} e^{-t} dt E[|\log Z(1)| : Z(1) < 1]
\end{aligned}$$

となり、この不等式から、$E[|\log Z(1)| : Z(1) < 1] < \infty$ を得る。こうして、(5.12) を得る。

Lemma 5.7

$$\frac{c_1}{\Gamma(n)} (\int_M^\infty x^{n-1} \log x E[Z(1)^n e^{-xZ(1)}]dx \sim c_1 \log n, n \to \infty. \tag{5.13}$$

Proof.
$x = \frac{(n-1)u}{Z(1)}$ と変数変換すると、

$$\begin{aligned}
\frac{c_1}{\Gamma(n)} (\int_M^\infty x^{n-1} \log x E[Z(1)^n e^{-xZ(1)}]dx &= \frac{c_1}{\Gamma(n)} (n-1)^n E[\int_{Z(1)M}^\infty u^{n-1} \times \\
&\quad \{\log(n-1) + \log u - \log Z(1)\} e^{-(n-1)u} du] \\
&= \frac{c_1}{\Gamma(n)} (n-1)^n \{\log(n-1) E[\int_{Z(1)M}^\infty u^{n-1} e^{-(n-1)u} du] \\
&+ E[\int_{Z(1)M}^\infty u^{n-1} \log u e^{-(n-1)u} du] \\
&+ E[\int_{Z(1)M}^\infty u^{n-1} (-\log Z(1)) e^{-(n-1)u} du]\} \\
&\equiv \frac{c_1}{\Gamma(n)} (n-1)^n \{I_1 + I_2 + I_3\}
\end{aligned}$$

となる。Saddlepoint method の通常の計算によって、

$$I_1 \sim \log(n-1) e^{-(n-1)} \frac{\sqrt{2\pi}}{\sqrt{n-1}},$$
$$I_2 \sim e^{-(n-1)} \frac{\sqrt{2\pi}}{\sqrt{n-1}} o(1),$$
$$I_3 \sim e^{-(n-1)} \frac{\sqrt{2\pi}}{\sqrt{n-1}} E[-\log Z(1)], n \to \infty,$$

を得る。Lemma 5.6 から $E[-\log Z(1)]$ は有限の値であり、上の 3 つの項の中で,I_1 が主要項であることが分かる。

最後に、Stirling の公式を使って、(5.13) を得る。

こうして、Theorem 5 (2) の証明を完結する準備が整った。
(5.11),(5.12),(5.13) によって

$$K_n^3 \sim c_1 \log n, n \to \infty,$$

であることがわかる。(5.6) は、今の場合も成り立っていることに注意すると、

$$\mathbf{K}_n \sim c_1 \log n, n \to \infty$$

を得る。 こうして、Theorem 5 の証明は完結した。

6 Appendix : 合成関数の高階導関数

1 次元ユークリッド空間 R^1 の区間で、十分なめらかな関数 $g(x)$ と指数関数の合成関数 $e^{ag(x)}$ の合成関数の高階導関数を求める。ここで、a は定数である。次の等式

$$e^{-ag}\frac{d^n}{dx^n}e^{ag} = Y_n(ag_1, ag_2, \cdots, ag_n), \quad g_i = D_x^i g, \quad i = 1, 2, \cdots, n$$

$$Y_n(u_1, u_2, \cdots, u_n) = \sum_{\beta} \frac{n!}{\prod_{j=1}^n \beta_j!} \prod_{i=1}^n (\frac{u_i}{i!})^{\beta_i}. \tag{7.1}$$

が成り立つことが知られている [7],[9]。ここで $\sum_\beta$ は、自然数 n のあらゆる分割 β についての和を表す。

References.

[1] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence* , John Wily and Sons(1986).

[2] W. J. Ewens, *Mathematical Population Genetics*, Springer Verlag(1980).

[3] S. Feng and F. M. Hoppe, Large deviation principles for some random combinatorial structures in population genetics and Brownian motion, Ann. Applied Probab. **8**, 975-994(1998).

[4] M. Kimura, The Neutral Theory of Molecular Evolution(1983).

[5] M. Perman, Order statistics for jumps of normalized subordinators, Stochastic Processes and their Applications vol. **46**, 267-281(1993).

[6] J. Pitman and M. Yor, The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator, Ann. Probab. **25**, 855-900(1997).

[7] S. Roman, The formula of Faà di Bruno, American Mathematical Monthly **87**, 805-809(1980).

[8] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press(1999).

[9] 清水昭信 他、Faà di Bruno の公式とその応用、preprint.