統計数理（1998）
第46巻 第1号 227-240

［原著論文］

統計基礎差分方程式に基づく多変量離散型分布のノンパラメトリックな構築

統計数理研究所　土屋　高宏・松縄　規

（受付 1997年8月20日；改訂 1997年11月19日）

要　旨

多変量離散型統計基礎モデルを，ノンパラメトリックな統計的不確定性関係およびそれからのノンパラメトリックな統計基礎方程式に基づいて，理論的に構築する．そのために，議論の基盤に，多次元離散可測空間とその上での計数測度を考える．その結果，統計基礎方程式は多次元差分方程式で与えられることが示される．更に，基礎差分方程式を和分することにより，代表的な多変量離散型分布を含む多変量統計基礎モデルを，最小不確定性分布の誘導の観点から組織的に構築できることを示す．

キーワード：多変量離散型統計モデル，ノンパラメトリックな統計的不確定性関係，統計基礎差分方程式，全差分，不定全和分，最小不確定性分布．

1. はじめに

著者の一人（Matsunawa (1994)）は母集団分布の存在に触れずに，データに基づき統計的基礎モデルを構築する事を考察した．そのためにはモデル分布型特定のための明確な原理と，統計学の進歩に繋がる基礎方程式を与える事が必要であった．結果として，概略次の様な統計的不確定性関係と統計基礎方程式が重要な意味を持つことに思い至った：ランダム行列Aの統計的基礎モデルを記述する確率分布P^Aの族はあるσ-有限測度μに関し絶対連続な多変量密度fから成るものとする．Δで観測対象行列をモデルで観測した時の誤差行列（＝系の乱れ）を表す．モデルのデータ記述能力としてAに関する性能比強度$\mathcal{P}_A=\nabla_A \ln f$を考える．ここで$\nabla_A$は$A$による gradient 作用素を表す．また，誤差と観測機構（＝モデル）の平均変動（＝不確定性）を各々Σ，Iとし，それらの間の相互作用の平均変動Jとその転置行列J^tの存在を仮定する．この設定の下で，統計的不確定性関係(a) $\Sigma-J^tI^{-1}J\gg 0$（非負値定符号行列），およびその等号条件としての統計基礎方程式(b) $\nabla_A f=\Xi\Delta f$ (μ−a.e.) を得る．ただし，Ξは観測精度行列を表す．不等式(a)は「観測対象とモデルからなる統計的複合システムで両者を同時に決定論的因果関係により厳密に記述する事の不可能性」を示している．統計基礎方程式(b)は統計的意味での客観的基礎モデルの構築を可能にし，この系に大域的な秩序を形成する．Ξ，Δ等を適切に選択すれば，代表的な多変量分布を多数含む最小不確定性分布族が誘導される．そこにはピアソンシステムも含まれる．

ところで，ノンパラメトリックな場合の上記の一般理論はσ-有限測度μが特に計数測度（counting measure）の場合も成立している．ただしその場合，ノンパラメトリックな統計基礎方程式は差分方程式の形を取る．また統計基礎方程式に基づく離散型統計基礎モデルの構築も，連続型分布導出の際の積分の代わりに，和分（summation operation）を用いることになる．こ

れらの接近方法は，筆者等の知る限り，従来の統計理論等において取られて来なかったように見える．本稿では，離散型のノンパラメトリックな統計基礎モデルの構築に研究対象を絞り，上記の接近法を組織的に準備・活用して，その場合の統計基礎差分方程式を与え，それに基づく代表的な多変量離散型統計基礎モデルを構築することを試みる．

本稿の構成は次の通りである．次章では以下の議論に必要な記号および仮定を準備する．第3章では多変量離散型のノンパラメトリックな統計的不確定性関係の表現とそれに基づくノンパラメトリックな統計基礎方程式を与える．また第4章では，離散型分布の構築の例として，多項分布，多変量超幾何分布，order k の行列離散分布などを含む多変量離散型分布を誘導出来ることを示す．

2. 記号，仮定

本章では，ノンパラメトリックな多変量離散型モデルの構築を考察する際に必要な記号と仮定を挙げる．

$A=(\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n)$ を列ベクトル $\boldsymbol{a}_i=(a_{1i}, a_{2i}, \cdots, a_{mi})^t$ $(i=1,2,\cdots,n)$ を持つ $m\times n$ 行列とする．各 i について，$\boldsymbol{a}_i$ から m_i 個の要素を取り出して部分列ベクトル $\boldsymbol{a}_i^*=(a_{1i}^*, a_{2i}^*, \cdots, a_{m_ii}^*)^t$ を構成する．この部分列ベクトルを積み上げる事により，行列 A の列ベクトル化を次のように定義する：

$$|A\rangle=(a_{11}^*, a_{21}^*, \cdots, a_{m_11}^*, a_{12}^*, a_{22}^*, \cdots, a_{m_22}^*, \cdots, a_{1n}^*, a_{2n}^*, \cdots, a_{m_nn}^*)^t.$$

ここで，$\sum_{i=1}^n m_i=:k$ は生成された列ベクトル $|A\rangle$ の次元を表す．また，これを転置した行ベクトルを $\langle A|\,[\equiv|A\rangle^t]$ のように表す．

いま，$|A\rangle$ をその要素が非負整数値からなるランダム行列 A をベクトル化した k-次元ランダムベクトルで，測度空間 $(R^k, \boldsymbol{B}^k, \mu)$ 上で定義されているものとする．ここに，R^k はすべての k-次元非負整数点の集合，$\boldsymbol{B}^k$ は R^k の部分集合の σ-集合体，μ は可測空間 $(R^k, \boldsymbol{B}^k)$ 上の計数測度 (counting measure) を表す．以下で混乱のおそれがないならば，$|A\rangle$ の代わりに A と記す．$\boldsymbol{P}=\{P_\Lambda^A; \Lambda\in\mathcal{S}_\Lambda\}$ を上記の可測空間で定義される多変量ノンパラメトリックな離散型モデルの分布族とする．ここで，Λ は A と関数的に独立なパラメーター空間 $\mathcal{S}_\Lambda$ に属する潜在パラメーターとする．また，$f(A;\Lambda)$ を P_Λ^A の確率関数とする．なお，以下の議論では簡単のために潜在パラメーター Λ を陽に表現しないこともある．

さて，以下の議論での数学的道具立てとなる偏差分 (partial difference) を導入しよう (cf. Milne-Thomson (1933))：

定義 2.1. 次の式の右辺が意味する量を，$|A\rangle=(a_1,\cdots,a_k)^t$ の第 i 要素 $a_i\,(1\le i\le k)$ に関する増分1の $f(a_1,\cdots,a_k;\Lambda)$ の偏差分と呼び，左辺の記号をもって表す：

$$\frac{\delta_p f(a_1,\cdots,a_k;\Lambda)}{\delta_p a_i}:=f(a_1,\cdots,a_i+1,\cdots,a_k;\Lambda)-f(a_1,\cdots,a_k;\Lambda). \tag{2.1}$$

次に，A の行列値ランダム関数 $\Phi(A;\Lambda)$ で観測対象関数，$\Psi(A;\Lambda)$ でその行列値ランダム近似関数，$\Delta\equiv\Delta(A;\Lambda):=\Phi(A;\Lambda)-\Psi(A;\Lambda)$ で測定誤差を表す．さらに，統計モデルの持つデータの記述能力として，次の量で与えられる離散型ランダム行列 A に関する性能比強度を考える：

$$|\mathcal{P}\rangle:=\left|\frac{\delta_p \ln f(A;\Lambda)}{\delta_p A}\right\rangle\left[=:\left(\frac{\delta_p \ln f}{\delta_p a_i}\right), \text{ where } (a_1,\cdots,a_k)^t=|A\rangle\right].$$

また，次の観測対象，観測装置(＝モデル)，それらの間の相互作用に関する平均変動が存在すると仮定する：

$$\Sigma \quad := E_f[|\varDelta\rangle\langle\varDelta|] = \sum_{a_1=0}^{\infty}\sum_{a_2=0}^{\infty}\cdots\sum_{a_k=0}^{\infty}|\varDelta(A\,;\,\Lambda)\rangle\langle\varDelta(A\,;\,\Lambda)|f(A\,;\,\Lambda),$$

$$I \quad := E_f[|\mathcal{P}\rangle\langle\mathcal{P}|] = \sum_{a_1=0}^{\infty}\sum_{a_2=0}^{\infty}\cdots\sum_{a_k=0}^{\infty}\left|\frac{\delta_p \ln f(A\,;\,\Lambda)}{\delta_p A}\right\rangle\left\langle\frac{\delta_p \ln f(A\,;\,\Lambda)}{\delta_p A}\right|f(A\,;\,\Lambda),$$

$$J^t \quad := E_f[|\varDelta\rangle\langle\mathcal{P}|] = \sum_{a_1=0}^{\infty}\sum_{a_2=0}^{\infty}\cdots\sum_{a_k=0}^{\infty}|\varDelta(A\,;\,\Lambda)\rangle\left\langle\frac{\delta_p \ln f(A\,;\,\Lambda)}{\delta_p A}\right|f(A\,;\,\Lambda),$$

$$J \quad := E_f[|\mathcal{P}\rangle\langle\varDelta|] = \sum_{a_1=0}^{\infty}\sum_{a_2=0}^{\infty}\cdots\sum_{a_k=0}^{\infty}\left|\frac{\delta_p \ln f(A\,;\,\Lambda)}{\delta_p A}\right\rangle\langle\varDelta(A\,;\,\Lambda)|f(A\,;\,\Lambda),$$

$$K^{-1} \quad := J^t I^{-1}$$

3. 多変量離散型モデルに関するノンパラメトリックな統計的不確定性関係

以上の記号および仮定の下で，次の定理が成立する：

定理 3.1. $^{\forall}\boldsymbol{x}\in R^k$ に対して次の不等式（＝統計的不確定性関係）が成立する：

$$\boldsymbol{x}^t\Sigma\boldsymbol{x} \geq \boldsymbol{x}^t J^t I^{-1} J\boldsymbol{x}\,.$$

上記不等式で等号が成り立つための必要十分条件は，正則な測定精度行列 $K=I(J^t)^{-1}$ が存在して，

$$\left|\frac{\delta_p \ln f(A\,;\,\Lambda)}{\delta_p A}\right\rangle = K(\Lambda)\cdot|\varDelta(A\,;\,\Lambda)\rangle \quad (\mu\text{−a.e.}) \tag{3.1}$$

で表現される統計基礎方程式が成立する時，そしてその時のみに限られる．□

上記の定理の証明は Matsunawa (1994) に準じて行えるので省略する．

次に，多変量ノンパラメトリックな離散型モデル分布の密度関数の一般型を得るために，連続関数の場合の全微分に対応する全差分 (total difference) を定義する．

定義 3.1. $f(a_1,\ a_2,\cdots,\ a_k)$ の全差分 δ_t を次のように定義する：

$$\delta_t f(a_1,\ a_2,\cdots,\ a_k) := \frac{\delta_p f(a_1,\ a_2,\cdots,\ a_k)}{\delta_p a_1}\delta a_1 + \frac{\delta_p f(a_1,\ a_2,\cdots,\ a_k)}{\delta_p a_2}\delta a_2 + \cdots + \frac{\delta_p f(a_1,\ a_2,\cdots,\ a_k)}{\delta_p a_k}\delta a_k.$$

さて，基礎差分方程式(3.1)から，離散型統計基礎モデルを得ることを考えよう．このために，我々は全差分の逆演算としての不定全和分 (indefinite total summation) を導入する：

定義 3.2.

$$\int^{\oplus}\delta_t f(a_1,\ a_2,\cdots,\ a_k) = \int^{\oplus}\left\{\frac{\delta_p f(a_1,\ a_2,\cdots,\ a_k)}{\delta_p a_1}\delta a_1 + \frac{\delta_p f(a_1,\ a_2,\cdots,\ a_k)}{\delta_p a_2}\delta a_2 + \cdots + \frac{\delta_p f(a_1,\ a_2,\cdots,\ a_k)}{\delta_p a_k}\delta a_k\right\}. \tag{3.2}$$

なお，この段階では，和分に関して定数の不定性が残る．しかし，以下の議論では $f(a_1, a_2,\cdots, a_k)$ が確率関数であることから，その規格化条件によって，モデルが決定できるか否かは吟味可能になる．また，全和分は分配可能ではないことも注意すべきである．分配をして計算したい時は，二変数以上が同時に現れる部分の扱いに気をつけねばならない．

以上の設定の下で，定理 3.1 の基礎差分方程式を和分して，次の多変量離散型最小不確定性分布の一般型を得る：

定理 3.2. 前の定理と同じ記号と条件の下で，離散型ランダム行列 A のノンパラメトリックな統計基礎モデル分布の密度関数は次のように与えられる：

$$f(|A\rangle; \Lambda)=c(\Lambda)\exp\left\{\int^{\oplus}\langle\delta A|K(\Lambda)|\Delta(A;\Lambda)\rangle\right\}. \tag{3.3}$$

ここで，$\langle\delta A|$ は k-次元差分行ベクトルで，

$$\langle\delta A|:=(\delta a_1, \delta a_2,\cdots, \delta a_k),$$

$K(\Lambda)=I(J^t)^{-1}$ は Λ に依存してもよい $k\times k$ 精度行列，$\int^{\oplus}$ は (3.2) の意味での和分を表す．また，$c(\Lambda)$ は次の規格化条件を満たすスカラー値関数を表す：

$$\sum_{a_1=0}^{\infty}\sum_{a_2=0}^{\infty}\cdots\sum_{a_k=0}^{\infty}f(|A\rangle; \Lambda)=1. \qquad \square$$

この定理によって，Δ, K を適切に設定したとき，$c(\Lambda)$ が決まれば，多変量離散型基礎モデルを構築することが可能となる．

次章で以下の不定和分の公式を用いる．証明は右辺の差分を考えれば容易に得られるので省略する（cf. Miller (1960), Nörlund (1924)）．

公式 3.1.

(i) $$\int^{\oplus}\ln(a+x)\delta x=\begin{cases}\ln(a+x-1)! & (a\text{ が整数のとき})\\ \ln\Gamma(a+x) & (a\text{ が整数でないとき}).\end{cases}$$

(ii) $$\int^{\oplus}\ln p\delta x=\ln p^x \quad (p:\text{非負定数}).$$

(iii) $$\int^{\oplus}\ln(a-x)^{-1}\delta x=\begin{cases}\ln(a-x)! & (a\text{ が整数のとき})\\ \ln\Gamma(a-x+1) & (a\text{ が整数でないとき}).\end{cases}$$

4. 多変量離散型統計基礎モデル構築の例

本章では，前章の定理を利用して最小不確定性分布としての，いくつかの代表的な多変量離散型分布を誘導する．

4.1 多項分布（Multinomial distribution）

$$f(\boldsymbol{x};\boldsymbol{p}, n)=\frac{n!}{x_1!\cdots x_{k+1}!}p_1^{x_1}\cdots p_{k+1}^{x_{k+1}}.$$

ここで，$\boldsymbol{x}=(x_1,\cdots, x_{k+1})^t$ は非負の整数値をとる $k+1$-次元ベクトルで，$\sum_{i=1}^{k+1}x_i=n, 0\le x_i\le n$, また，$\boldsymbol{p}=(p_1,\cdots, p_{k+1})^t$ で，$\sum_{i=1}^{k+1}p_i=1$, $0<p_i<1$.

(誘導). 定理3.2において，Δ, Kを次の様に設定する：

$$K(\Lambda)=I \quad (k\times k),$$

$$|\Delta(\boldsymbol{x}\,;\,\boldsymbol{p},\,n)\rangle=\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_1+1}-\ln\frac{p_{k+1}}{p_1},\cdots,\,\ln\frac{n-\sum_{i=1}^{k}x_i}{x_k+1}-\ln\frac{p_{k+1}}{p_k}\right)^t.$$

(3.3)式より

$$\begin{aligned}f(\boldsymbol{x}\,;\,\boldsymbol{p},\,n)&\propto\exp\left[\int^{\oplus}\left\{\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_1+1}-\ln\frac{p_{k+1}}{p_1}\right)\delta x_1\right.\right.\\&\qquad\left.\left.+\cdots+\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_k+1}-\ln\frac{p_{k+1}}{p_k}\right)\delta x_k\right\}\right]\\&=\exp\left[-\ln\,(n-\textstyle\sum_{i=1}^{k}x_i)!-\ln\,x_1!-\cdots-\ln\,x_k!\right.\\&\qquad\left.-\ln\left(\frac{p_{k+1}}{p_1}\right)^{x_1}-\cdots-\ln\left(\frac{p_{k+1}}{p_k}\right)^{x_k}\right]\\&=\frac{1}{x_1!\cdots x_k!(n-\sum_{i=1}^{k}x_i)!}p_1^{x_1}\cdots p_k^{x_k}p_{k+1}^{-\sum_{i=1}^{k}x_i}.\end{aligned}$$

ここで，

$$(p_1+\cdots+p_{k+1})^n=\sum_{x_1+\cdots+x_{k+1}=n}\frac{n!}{x_1!\cdots x_{k+1}!}p_1^{x_1}\cdots p_{k+1}^{x_{k+1}}.$$

よって

$$c(\boldsymbol{p},\,n)=n!\,p_{k+1}^{n}.$$

4.2 多変量超幾何分布（Multivariate hypergeometric distribution）

$$f(\boldsymbol{x};\boldsymbol{M},N,n)=\frac{\binom{M_1}{x_1}\cdots\binom{M_{k+1}}{x_{k+1}}}{\binom{N}{n}}.$$

ここで，$\boldsymbol{x}=(x_1,\cdots,\,x_{k+1})^t$, $\boldsymbol{M}=(M_1,\cdots,\,M_{k+1})^t$, $\sum_{i=1}^{k+1}x_i=n$, $0\leq x_i\leq M_i$, $\sum_{i=1}^{k+1}M_i=N$, $0<M_i<N$.

(誘導). Δ, Kを次の様に設定する：

$$K(\Lambda)=I \quad (k\times k),$$

$$\begin{aligned}|\Delta(\boldsymbol{x}\,;\,\boldsymbol{M},\,N,\,n)\rangle=&\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_1+1}-\ln\frac{N-\sum_{i=1}^{k}M_i-n+\sum_{i=1}^{k}x_i+1}{M_1-x_1},\right.\\&\left.\cdots,\,\ln\frac{n-\sum_{i=1}^{k}x_i}{x_k+1}-\ln\frac{N-\sum_{i=1}^{k}M_i-n+\sum_{i=1}^{k}x_i+1}{M_k-x_k}\right)^t.\end{aligned}$$

(3.3)式より

$$\begin{aligned}&f(\boldsymbol{x}\,;\,\boldsymbol{M},\,N,\,n)\\&\propto\exp\left[\int^{\oplus}\left\{\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_1+1}-\ln\frac{N-\sum_{i=1}^{k}M_i-n+\sum_{i=1}^{k}x_i+1}{M_1-x_1}\right)\delta x_1\right.\right.\\&\qquad\left.\left.+\cdots+\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_k+1}-\ln\frac{N-\sum_{i=1}^{k}M_i-n+\sum_{i=1}^{k}x_i+1}{M_k-x_k}\right)\delta x_k\right\}\right]\\&=\exp\left[-\ln\,(n-\textstyle\sum_{i=1}^{k}x_i)!-\ln\,x_1!-\cdots-\ln\,x_k!\right.\\&\qquad\left.-\ln(N-\textstyle\sum_{i=1}^{k}M_i-n+\sum_{i=1}^{k}x_i)!-\ln(M_1-x_1)!-\cdots-\ln(M_k-x_k)!\right]\end{aligned}$$

$$=\frac{1}{(M_1-x_1)!\,x_1!\cdots(M_k-x_k)!\,x_k!}$$
$$\times\frac{1}{(N-\sum_{i=1}^k M_i-n+\sum_{i=1}^k x_i)!\,(n-\sum_{i=1}^k x_i)!}$$
$$=\binom{M_1}{x_1}\cdots\binom{M_k}{x_k}\frac{1}{M_1!\cdots M_k!\,(M_{k+1}-x_{k+1})!\,x_{k+1}!}.$$

ここで,

$$(1+x)^{M_1}\cdots(1+x)^{M_{k+1}}=(1+x)^N$$

の両辺を展開し，x^n の係数を比較すれば,

$$\sum_{x_1+\cdots+x_{k+1}=n}\binom{M_1}{x_1}\cdots\binom{M_{k+1}}{x_{k+1}}=\binom{N}{n}.$$

よって

$$c(\boldsymbol{M},\,N,\,n)=M_1!\cdots M_{k+1}!\Big/\binom{N}{n}.$$

4.3 負の多項分布（Negative multinomial distribution）

$$f(\boldsymbol{x}\,;\,\boldsymbol{p},\,l)=\binom{l+\sum_{i=1}^k x_i-1}{x_1,\cdots,\,x_k,\,l-1}p_0^l p_1^{x_1}\cdots p_k^{x_k}$$
$$=\frac{\Gamma(l+\sum_{i=1}^k x_i)}{x_1!\cdots x_k!\,\Gamma(l)}p_0^l p_1^{x_1}\cdots p_k^{x_k}.$$

ここで，$\boldsymbol{x}=(x_1,\cdots,\,x_k)^t$, $0\leq x_i$, $\boldsymbol{p}=(p_0,\,p_1,\cdots,\,p_k)^t$, $p_0=1-\sum_{i=1}^k p_i$.

(誘導)．Δ, K を次の様に設定する：

$$K(\Lambda)=I \quad (k\times k),$$
$$|\Delta(\boldsymbol{x}\,;\,\boldsymbol{p},\,l)\rangle=\left(\ln\frac{p_1}{x_1+1}-\ln\frac{1}{l+\sum_{i=1}^k x_i},\cdots,\,\ln\frac{p_k}{x_k+1}-\ln\frac{1}{l+\sum_{i=1}^k x_i}\right)^t.$$

(3.3)式より

$$f(\boldsymbol{x}\,;\,\boldsymbol{p},\,l)\propto\exp\Big[\int^{\oplus}\Big\{\Big(\ln\frac{p_1}{x_1+1}-\ln\frac{1}{l+\sum_{i=1}^k x_i}\Big)\delta x_1$$
$$+\cdots+\Big(\ln\frac{p_k}{x_k+1}-\ln\frac{1}{l+\sum_{i=1}^k x_i}\Big)\delta x_k\Big\}\Big]$$
$$=\exp\Big[\ln p_1^{x_1}+\cdots+\ln p_k^{x_k}-\ln x_1!-\cdots-\ln x_k!+\ln\Gamma(l+\textstyle\sum_{i=1}^k x_i)\Big]$$
$$=\frac{\Gamma(l+\sum_{i=1}^k x_i)}{x_1!\cdots x_k!}p_1^{x_1}\cdots p_k^{x_k}$$
$$=\binom{l+\sum_{i=1}^k x_i-1}{x_1,\cdots,\,x_k,\,l-1}p_1^{x_1}\cdots p_k^{x_k}\Gamma(l).$$

ここで,

$$\sum_{x_1,\cdots,x_k}\binom{l+\sum_{i=1}^k x_i-1}{x_1,\cdots,\,x_k,\,l-1}p_1^{x_1}\cdots p_k^{x_k}\Gamma(l)$$
$$=\sum_{x_1,\cdots,x_k}\binom{l+\sum_{i=1}^k x_i-1}{x_1+\cdots+x_k}\binom{x_1+\cdots+x_k}{x_1,\cdots,\,x_k}p_1^{x_1}\cdots p_k^{x_k}\Gamma(l)$$

$$=\sum_{m=0}^{\infty}\sum_{x_1+\cdots+x_k=m}\binom{l+m-1}{m}\binom{m}{x_1,\cdots,x_k}p_1^{x_1}\cdots p_k^{x_k}\Gamma(l)$$
$$=\sum_{m=0}^{\infty}\binom{-l}{m}(-1)^m(p_1+\cdots+p_k)^m\Gamma(l)$$
$$=\left(1-\sum_{i=1}^{k}p_i\right)^{-l}\Gamma(l).$$

$$1-\sum_{i=1}^{k}p_i=p_0$$

とおけば

$$c(\boldsymbol{p},l)=\frac{p_0^l}{\Gamma(l)}.$$

4.4 多変量負の超幾何分布 (Multivariate negative hypergeometric distribution)

$$f(\boldsymbol{x};\boldsymbol{M},N,n)=\frac{\binom{M_1+x_1-1}{x_1}\cdots\binom{M_{k+1}+x_{k+1}-1}{x_{k+1}}}{\binom{N+n-1}{n}}$$
$$=\frac{n!\Gamma(N)}{\Gamma(N+n)}\prod_{i=1}^{k+1}\frac{\Gamma(M_i+x_i)}{\Gamma(M_i)x_i!}.$$

ここで, $\boldsymbol{x}=(x_1,\cdots,x_{k+1})^t$, $\boldsymbol{M}=(M_1,\cdots,M_{k+1})^t$, $\sum_{i=1}^{k+1}x_i=n$, $0\le x_i\le n$, $\sum_{i=1}^{k+1}M_i=N$.

(誘導). $\varDelta$, K を次の様に設定する：

$$K(\Lambda)=I\quad(k\times k),$$
$$|\varDelta(\boldsymbol{x};\boldsymbol{M},N,n)\rangle=\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_1+1}-\ln\frac{N-\sum_{i=1}^{k}M_i+n-\sum_{i=1}^{k}x_i-1}{M_1+x_1},\right.$$
$$\left.\cdots,\ln\frac{n-\sum_{i=1}^{k}x_i}{x_k+1}-\ln\frac{N-\sum_{i=1}^{k}M_i+n-\sum_{i=1}^{k}x_i-1}{M_k+x_k}\right)^t.$$

(3.3)式より

$$f(\boldsymbol{x};\boldsymbol{M},N,n)$$
$$\propto\exp\left[\int^{\oplus}\left\{\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_1+1}-\ln\frac{N-\sum_{i=1}^{k}M_i+n-\sum_{i=1}^{k}x_i-1}{M_1+x_1}\right)\delta x_1\right.\right.$$
$$\left.\left.+\cdots+\left(\ln\frac{n-\sum_{i=1}^{k}x_i}{x_k+1}-\ln\frac{N-\sum_{i=1}^{k}M_i+n-\sum_{i=1}^{k}x_i-1}{M_k+x_k}\right)\delta x_k\right\}\right]$$
$$=\exp\left[-\ln(n-\textstyle\sum_{i=1}^{k}x_i)!-\ln x_1!-\cdots-\ln x_k!\right.$$
$$\left.+\ln\Gamma(N-\textstyle\sum_{i=1}^{k}M_i+n-\sum_{i=1}^{k}x_i)+\ln\Gamma(M_1+x_1)+\cdots+\ln\Gamma(M_k+x_k)\right]$$
$$=\frac{\Gamma(M_1+x_1)\cdots\Gamma(M_k+x_k)\Gamma(M_{k+1}+x_{k+1})}{x_1!\cdots x_k!x_{k+1}!}$$
$$=\binom{M_1+x_1-1}{x_1}\cdots\binom{M_k+x_k-1}{x_k}\binom{M_{k+1}+x_{k+1}-1}{x_{k+1}}\prod_{i=1}^{k+1}\Gamma(M_i)$$
$$=\binom{-M_1}{x_1}\cdots\binom{-M_{k+1}}{x_{k+1}}(-1)^{\sum_{i=1}^{k+1}x_i}\prod_{i=1}^{k+1}\Gamma(M_i).$$

ここで,

$$(1-x)^{-M_1}\cdots(1-x)^{-M_{k+1}}=(1-x)^{-N}$$

の両辺を展開し，x^n の係数を比較すれば，

$$\sum_{x_1+\cdots+x_{k+1}=n}\binom{-M_1}{x_1}\cdots\binom{-M_{k+1}}{x_{k+1}}(-1)^{x_1+\cdots+x_{k+1}}=(-1)^n\binom{-N}{n}$$
$$=\binom{N+n-1}{n}.$$

よって

$$c(\boldsymbol{M},N,n)=1\Big/\left\{\binom{N+n-1}{n}\prod_{i=1}^{k+1}\Gamma(M_i)\right\}.$$

4.5　多重ポアソン分布（Multiple Poisson distribution）

$$f(\boldsymbol{x};\boldsymbol{\theta})=e^{-\Sigma_{i=1}^k\theta_i}\prod_{i=1}^{k}\frac{\theta_i^{x_i}}{x_i!}.$$

ここで，$\boldsymbol{x}=(x_1,\cdots,x_k)^t$, $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_k)^t$, $0\le x_i$, $0\le\theta_i$.

（誘導）．Δ, K を次の様に設定する：

$$K(\Lambda)=I\quad(k\times k),$$
$$|\Delta(x;\theta)\rangle=[\ln\theta_1-\ln(x_1+1),\cdots,\ln\theta_k-\ln(x_k+1)]^t.$$

(3.3)式より

$$f(\boldsymbol{x};\boldsymbol{\theta})\propto\exp\left[\int^{\oplus}[\{\ln\theta_1-\ln(x_1+1)\}\delta x_1+\cdots+\{\ln\theta_k-\ln(x_k+1)\}\delta x_k]\right]$$
$$=\exp[\ln\theta_1^{x_1}+\cdots+\ln\theta_k^{x_k}-\ln x_1!-\cdots-\ln x_k!]$$
$$=\prod_{i=1}^{k}\frac{\theta_i^{x_i}}{x_i!}.$$

ここで，

$$\sum_{x_1=0}^{\infty}\cdots\sum_{x_k=0}^{\infty}\left(\prod_{i=1}^{k}\frac{\theta_i^{x_i}}{x_i!}\right)=\left(\sum_{x_1=0}^{\infty}\frac{\theta_1^{x_1}}{x_1!}\right)\cdots\left(\sum_{x_k=0}^{\infty}\frac{\theta_k^{x_k}}{x_k!}\right)$$
$$=e^{\theta_1}\cdots e^{\theta_k}.$$

よって

$$c(\boldsymbol{\theta})=e^{-\Sigma_{i=1}^k\theta_i}.$$

4.6　Order k の行列離散分布（Matrix discrete distributions of order k）

本節で，文字 k はここまで使用してきたベクトルの次元ではないことに注意しておく．k を例えば s と表示してもよいが（cf. Johnson et al. (1992, 1997)），order k の分布という習慣化した用法があるので，ここではそれにならうことにする．

Order k（$k>1$）の行列離散分布を誘導するために，次の，(2.1)に基づいて得られる，偏差分の逆演算としての和分公式を準備する：

公式 4.1.

$$\int^{\oplus}\ln\frac{\sum_{\substack{\Sigma_{l=1}^k lx_{ijl}=x_{ij}+\varepsilon_{i\mu j\nu}\\ i=1,\cdots,u\\ j=1,\cdots,v}}g(x_{111},\cdots,x_{uvk})}{\sum_{\substack{\Sigma_{l=1}^k lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}g(x_{111},\cdots,x_{uvk})}\delta x_{\mu\nu}=\ln\sum_{\substack{\Sigma_{l=1}^k lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}g(x_{111},\cdots,x_{uvk}).$$

ただし，$g(x_{111},\cdots,x_{uvk})>0$，$\varepsilon_{i\mu j\nu}$ は次で定義される二値関数である：

$$\varepsilon_{i\mu j\nu}=\begin{cases}1 & i=\mu \text{ and } j=\nu\\ 0 & \text{otherwise.}\end{cases}$$

ここに，μ，ν はそれぞれ $1\leq\mu\leq u$，$1\leq\nu\leq v$ を満たす整数．この記号は多次元のとき($v=1$ のとき) 通常のクロネッカーのデルタに帰着する．

証明は，やはり右辺の偏差分を考えることにより実行できる．

以下で，大きさ $u\times v$ のランダム行列 $X=(X_{ij})$ の実現値 x_{ij} は前述の設定に従い，非負整数値を取るものとする．この時，$x_{111},x_{112},\cdots,x_{uvk}$ を x_{ij} 又は，$x_{ij}+1$ の k-分割の成分として考え得る非負の整数値を表すものとする．

いま，

$$K(\Lambda)=I\quad(uv\times uv),$$

$$(4.1)\qquad |\Delta(X;\Lambda)\rangle=\left[\ln\frac{\sum_{\substack{\sum_{l=1}^{k} lx_{ijl}=x_{ij}+\varepsilon_{i1j1}\\ i=1,\cdots,u\\ j=1,\cdots,v}} g(x_{111},\cdots,x_{uvk};\Lambda)}{\sum_{\substack{\sum_{l=1}^{k} lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}} g(x_{111},\cdots,x_{uvk};\Lambda)},\right.$$

$$\left.\cdots,\ln\frac{\sum_{\substack{\sum_{l=1}^{k} lx_{ijl}=x_{ij}+\varepsilon_{iuj v}\\ i=1,\cdots,u\\ j=1,\cdots,v}} g(x_{111},\cdots,x_{uvk};\Lambda)}{\sum_{\substack{\sum_{l=1}^{k} lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}} g(x_{111},\cdots,x_{uvk};\Lambda)}\right]^t\quad(uv\times 1)$$

と設定することにより，いくつかの order k の行列離散分布が誘導できる．

4.6.1 Order k の行列負の二項分布 (Matrix negative binomial distribution of order k)

$$f(X;p,Q,r)=p^r\sum_{\substack{\sum_l lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\binom{x_{111}+\cdots+x_{uvk}+r-1}{x_{111},\cdots,x_{uvk},\ r-1}q_{111}^{x_{111}}\cdots q_{uvk}^{x_{uvk}}$$

$$=p^r\sum_{\substack{\sum_l lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\Gamma(\sum_i\sum_j\sum_l x_{ijl}+r)}{\prod_i\prod_j\prod_l(x_{ijl}!)\Gamma(r)}q_{111}^{x_{111}}\cdots q_{uvk}^{x_{uvk}}.$$

ここで，$X=(x_{ij})_{u\times v}$，$x_{ij}=0,1,\cdots$ $(1\leq i\leq u,\ 1\leq j\leq v)$，$Q=\{q_{ijl},\ 1\leq i\leq u,\ 1\leq j\leq v,\ 1\leq l\leq k\}$，$0<q_{ijl}<1$，$r>0$，$q_{111}+\cdots+q_{uvk}+p=1$.

(誘導)．(4.1)式で

$$g(x_{111},\cdots,x_{uvk};\Lambda)=\frac{\Gamma(\sum_i\sum_j\sum_l x_{ijl}+r)}{\prod_i\prod_j\prod_l(x_{ijl}!)\Gamma(r)}q_{111}^{x_{111}}\cdots q_{uvk}^{x_{uvk}}$$

とおくと

$$f(X;p,Q,r)\propto\sum_{\substack{\sum_l lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\Gamma(\sum_i\sum_j\sum_l x_{ijl}+r)}{\prod_i\prod_j\prod_l(x_{ijl}!)\Gamma(r)}q_{111}^{x_{111}}\cdots q_{uvk}^{x_{uvk}}$$

$$=\sum_{\substack{\sum_l lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\binom{x_{111}+\cdots+x_{uvk}+r-1}{x_{111},\cdots,x_{uvk},\ r-1}q_{111}^{x_{111}}\cdots q_{uvk}^{x_{uvk}}.$$

ここで，規格化定数を決めるために次の S_1 を計算する：

$$S_1=\sum_{x_{11}=0}^{\infty}\cdots\sum_{x_{uv}=0}^{\infty}\sum_{\substack{\sum_l lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\binom{x_{111}+\cdots+x_{uvk}+r-1}{x_{111},\cdots,x_{uvk},\ r-1}q_{111}^{x_{111}}\cdots q_{uvk}^{x_{uvk}}.$$

$$x_{ijl}=n_{ijl},\ x_{ij}=n_{ij}+\sum_{l=1}^{k}(l-1)n_{ijl}\quad(1\leq i\leq u,\ 1\leq j\leq v)$$

とおくと (cf. Philippou et al. (1989)),

$$\begin{aligned}S_1&=\sum_{n_{11}=0}^{\infty}\cdots\sum_{n_{uv}=0}^{\infty}\sum_{\substack{\Sigma_l n_{ijl}=n_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\binom{n_{111}+\cdots+n_{uvk}+r-1}{n_{111},\cdots,\ n_{uvk},\ r-1}q_{111}^{n_{111}}\cdots q_{uvk}^{n_{uvk}}\\&=\sum_{n_{11}=0}^{\infty}\cdots\sum_{n_{uv}=0}^{\infty}\sum_{\substack{\Sigma_l n_{ijl}=n_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\binom{n_{111}+\cdots+n_{uvk}+r-1}{n_{111}+\cdots+n_{uvk}}\\&\quad\times\binom{n_{111}+\cdots+n_{uvk}}{n_{111},\cdots,\ n_{uvk}}q_{111}^{n_{111}}\cdots q_{uvk}^{n_{uvk}}\\&=\sum_{n_{11}=0}^{\infty}\cdots\sum_{n_{uv}=0}^{\infty}\binom{n_{11}+\cdots+n_{uv}+r-1}{n_{11}+\cdots+n_{uv}}\\&\quad\times\sum_{\substack{\Sigma_l n_{ijl}=n_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\binom{n_{111}+\cdots+n_{uvk}}{n_{111},\cdots,\ n_{uvk}}q_{111}^{n_{111}}\cdots q_{uvk}^{n_{uvk}}.\end{aligned}$$

ここで, $n_{11}+\cdots+n_{uv}=n$ とおくと,

$$\begin{aligned}S_1&=\sum_{n=0}^{\infty}\binom{n+r-1}{n}\sum_{n_{111}+\cdots+n_{uvk}=n}\binom{n_{111}+\cdots+n_{uvk}}{n_{111},\cdots,\ n_{uvk}}q_{111}^{n_{111}}\cdots q_{uvk}^{n_{uvk}}\\&=\sum_{n=0}^{\infty}\binom{-r}{n}(-1)^n(q_{111}+\cdots+q_{uvk})^n\\&=\{1-(q_{111}+\cdots+q_{uvk})\}^{-r}\\&=p^{-r}.\end{aligned}$$

よって

$$c(p,\ Q,\ r)=p^r.$$

4.6.2 Order k の行列ポアソン分布 (Matrix Poisson distribution of order k)

$$f(X\,;\,\Lambda)=\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\exp\Bigl(-\sum_i\sum_j\sum_l\lambda_{ijl}\Bigr)\frac{\prod_i\prod_j\prod_l\lambda_{ijl}^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}.$$

ここで, $X=(x_{ij})_{u\times v},\ x_{ij}=0,\ 1,\cdots\ (1\leq i\leq u,\ 1\leq j\leq v),\ \Lambda=\{\lambda_{ijl},\ 1\leq i\leq u,\ 1\leq j\leq v,\ 1\leq l\leq k\}$.

(誘導). (4.1)式で

$$g(x_{111},\cdots,\ x_{uvk}\,;\,\Lambda)=\frac{\prod_i\prod_j\prod_l\lambda_{ijl}^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}$$

とおくと

$$f(X\,;\,\Lambda)\propto\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\prod_i\prod_j\prod_l\lambda_{ijl}^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}.$$

ここで, 次の S_2 を計算する:

$$S_2=\sum_{x_{11}=0}^{\infty}\cdots\sum_{x_{uv}=0}^{\infty}\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\prod_i\prod_j\prod_l\lambda_{ijl}^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}.$$

$$x_{ijl}=n_{ijl},\ x_{ij}=n_{ij}+\sum_{l=1}^{k}(l-1)n_{ijl}\quad(1\leq i\leq u,\ 1\leq j\leq v)$$

とおくと，

$$\begin{aligned}S_2&=\sum_{n_{11}=0}^{\infty}\cdots\sum_{n_{uv}=0}^{\infty}\sum_{\substack{\sum_l n_{ijl}=n_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\prod_i\prod_j\prod_l\lambda_{ijl}^{n_{ijl}}}{\prod_i\prod_j\prod_l n_{ijl}!}\\&=\prod_{i=1}^{u}\prod_{j=1}^{v}\left\{\sum_{n_{ij}=0}^{\infty}\sum_{n_{ij1}+\cdots+n_{ijk}=n_{ij}}\prod_{l=1}^{k}\left(\frac{\lambda_{ijl}^{n_{ijl}}}{n_{ijl}!}\right)\right\}\\&=\prod_{i=1}^{u}\prod_{j=1}^{v}\prod_{l=1}^{k}\left(\sum_{n_{ijl}=0}^{\infty}\frac{\lambda_{ijl}^{n_{ijl}}}{n_{ijl}!}\right)\\&=\exp\left(\sum_i\sum_j\sum_l\lambda_{ijl}\right).\end{aligned}$$

よって

$$c(\Lambda)=\exp\left(-\sum_i\sum_j\sum_l\lambda_{ijl}\right).$$

4.6.3 Order k の行列対数級数分布（Matrix logarithmic series distribution of order k）

$$f(X\,;\,p,\,Q)=\sum_{\substack{\sum_l lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{(\sum_i\sum_j\sum_l x_{ijl}-1)!}{(-\ln p)\prod_i\prod_j\prod_l x_{ijl}!}\prod_i\prod_j\prod_l q_{ijl}^{x_{ijl}}.$$

ここで，$X=(x_{ij})_{u\times v}$, $x_{ij}=0,\ 1,\cdots$ $(1\leq i\leq u,\ 1\leq j\leq v)$, $\sum_{i,j}x_{ij}\geq 1$, $Q=\{q_{ijl},\ 1\leq i\leq u,\ 1\leq j\leq v,\ 1\leq l\leq k\}$, $p=1-q_{111}-\cdots-q_{uvk}$.

(誘導). (4.1)式で

$$g(x_{111},\cdots,\,x_{uvk}\,;\,\Lambda)=\frac{(\sum_i\sum_j\sum_l x_{ijl}-1)!}{\prod_i\prod_j\prod_l x_{ijl}!}\prod_i\prod_j\prod_l q_{ijl}^{x_{ijl}}$$

とおくと

$$f(X\,;\,p,\,Q)\propto\sum_{\substack{\sum_l lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{(\sum_i\sum_j\sum_l x_{ijl}-1)!}{\prod_i\prod_j\prod_l x_{ijl}!}\prod_i\prod_j\prod_l q_{ijl}^{x_{ijl}}.$$

ここで，

$$\begin{aligned}&\sum_{x_{11}=0}^{\infty}\cdots\sum_{x_{uv}=0}^{\infty}\sum_{\substack{\sum_l lx_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{(\sum_i\sum_j\sum_l x_{ijl}-1)!}{\prod_i\prod_j\prod_l x_{ijl}!}\prod_i\prod_j\prod_l q_{ijl}^{x_{ijl}}\\&=\sum_{n_{11}=0}^{\infty}\cdots\sum_{n_{uv}=0}^{\infty}\sum_{\substack{\sum_l n_{ijl}=n_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{(\sum_i\sum_j\sum_l n_{ijl}-1)!}{\prod_i\prod_j\prod_l n_{ijl}!}\prod_i\prod_j\prod_l q_{ijl}^{n_{ijl}}\\&=\sum_{n=1}^{\infty}\sum_{n_{111}+\cdots+n_{uvk}=n}\frac{(\sum_i\sum_j\sum_l n_{ijl}-1)!}{\prod_i\prod_j\prod_l n_{ijl}!}\prod_i\prod_j\prod_l q_{ijl}^{n_{ijl}}\\&=\sum_{n=1}^{\infty}\frac{(\sum_i\sum_j\sum_l q_{ijl})^n}{n}\\&=-\ln\left(1-\sum_i\sum_j\sum_l q_{ijl}\right)\\&=-\ln p.\end{aligned}$$

よって

$$c(p,\, Q)=\frac{1}{-\ln p}.$$

注 4.1. 本節で誘導した order k の行列離散分布はこれまでの文献では見掛けない．これらの分布は，$v=1$ ととれば，Philippou et al. (1989) による order k の多変量分布に，さらに，$u=1$ ととれば，一変量の order k のマルチパラメーター分布(cf. Philippou (1988))に帰着する．

注 4.2. Order k の行列負の二項分布と order k の行列ポアソン分布の関係，および order k の行列負の二項分布と order k の行列対数級数分布の関係は，Philippou et al. (1989) の結果とパラレルに導くことができる．このことを以下に示す．

Order k の行列負の二項分布において，$r\to\infty$のとき，$q_{ijl}\to 0$, $rq_{ijl}\to\lambda_{ijl}$ ($0<\lambda_{ijl}<\infty$, $1\le i\le u$, $1\le j\le v$, $1\le l\le k$) と仮定すると，

$$\begin{aligned}
f(X\,;\,p,\,Q,\,r)&=p^r\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\Gamma(\sum_i\sum_j\sum_l x_{ijl}+r)}{\prod_i\prod_j\prod_l(x_{ijl}!)\Gamma(r)}q_{111}^{x_{111}}\cdots q_{uvk}^{x_{uvk}}\\
&=\left(1-\frac{r\sum_i\sum_j\sum_l q_{ijl}}{r}\right)^r\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\Gamma(\sum_i\sum_j\sum_l x_{ijl}+r)}{\Gamma(r)r^{\Sigma_i\Sigma_j\Sigma_l x_{ijl}}}\frac{\prod_i\prod_j\prod_l(rq_{ijl})^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}\\
&=\left(1-\frac{\sum_i\sum_j\sum_l rq_{ijl}}{r}\right)^r\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{(\sum_i\sum_j\sum_l x_{ijl}+r-1)(\sum_i\sum_j\sum_l x_{ijl}+r-2)\cdots r}{r^{\Sigma_i\Sigma_j\Sigma_l x_{ijl}}}\\
&\quad\times\frac{\prod_i\prod_j\prod_l(rq_{ijl})^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!},
\end{aligned}$$

ここで，関係式 $\Gamma(t+1)=t\Gamma(t)$ を逐次使った．したがって，上記の極限過程の下で，

$$f(X\,;\,p,\,Q,\,r)\xrightarrow[r\to\infty]{}\exp\left(-\sum_i\sum_j\sum_l\lambda_{ijl}\right)\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\prod_i\prod_j\prod_l\lambda_{ijl}^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}$$

となり，order k の行列ポアソン分布を得る．また，$\sum_{i,j}x_{ij}\ge 1$ ($1\le i\le u$, $1\le j\le v$) の条件の下で，$r\to 0$ と仮定すると，

$$\begin{aligned}
f(X\,;\,p,\,Q,\,r|\textstyle\sum_{i,j}x_{ij}\ge 1)&=\frac{f(X\,;\,p,\,Q,\,r,\,\sum_{i,j}x_{ij}\ge 1)}{1-f(0_{u\times v}\,;\,p,\,Q,\,r)}\\
&=\frac{p^r}{1-p^r}\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\Gamma(\sum_i\sum_j\sum_l x_{ijl}+r)}{\prod_i\prod_j\prod_l(x_{ijl}!)\Gamma(r)}\prod_i\prod_j\prod_l q_{ijl}^{x_{ijl}}\\
&=\frac{rp^r}{1-p^r}\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\frac{\Gamma(\sum_i\sum_j\sum_l x_{ijl}+r)}{r\Gamma(r)}\frac{\prod_i\prod_j\prod_l q_{ijl}^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}\\
&=\frac{rp^r}{1-p^r}\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\left(\sum_i\sum_j\sum_l x_{ijl}+r-1\right)\left(\sum_i\sum_j\sum_l x_{ijl}+r-2\right)\cdots(r+1)\\
&\quad\times\frac{\prod_i\prod_j\prod_l q_{ijl}^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}\\
&\xrightarrow[r\to 0]{}-\frac{1}{\ln p}\sum_{\substack{\Sigma_l l x_{ijl}=x_{ij}\\ i=1,\cdots,u\\ j=1,\cdots,v}}\left(\sum_i\sum_j\sum_l x_{ijl}-1\right)!\frac{\prod_i\prod_j\prod_l q_{ijl}^{x_{ijl}}}{\prod_i\prod_j\prod_l x_{ijl}!}
\end{aligned}$$

となり，order k の行列対数級数分布を得る．

5. おわりに

本稿では，従来殆ど組織的な研究がなされなかった離散型統計基礎モデルの構築に新しい接近方法を提案し，いくつかの具体的な事例を与えた．この研究の根底には，真の母集団分布の存在を意識すること無しに，離散型の統計基礎モデルを構築する理論を展開したいという意図があった．このことを実現するために，統計的不確定性関係の等号条件として，簡潔な統計基礎差分方程式を与え，それを基に和分を実行してモデルの理論的構築を行った．これは連続型のノンパラメトリックな統計基礎モデルの構築において基礎方程式が偏微分方程式で与えられ，基礎モデルがそれを積分して得られることと類推関係にある．即ち，μを基礎のσ-有限測度とすると，離散，連続を問わず，基礎方程式は$\nabla_A f = \Xi \Delta f$ (μ−a.e.) と統一して表現出来る．このように簡潔な方程式で，一変量のみならず多変量分布も含む，良く知られた確率モデルが導出可能なことはある意味で驚きである．特に本稿で議論した離散型分布の導出は，個別の分布に対する確率関数に関する漸化式や，確率母関数を用いた離散分布の通常の導出方法とは異なるものであり，分布の統一的な導入や分類およびこの分野の関連する研究に寄与できると思う．

謝　辞

査読者諸氏の本稿に対する注意深い検討と有益なコメントに感謝します．

参考文献

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

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*, Wiley, New York.

Matsunawa, T. (1994). Origin of distributions——nonparametric statistical uncertainty relation and a statistical fundamental equation——, *Proc. Inst. Statist. Math.*, **42**(2), 197–214 (in Japanese).

Miller, K. S. (1960). *An Introduction to the Calculus of Finite Differences and Difference Equations*, Henry Holt, New York.

Milne-Thomson, L. M. (1933). *The Calculus of Finite Differences*, Chelsea, New York.

Nörlund, N. E. (1924). *Vorlesungen über Differenzenrechnung*, Springer, Berlin.

Philippou, A. N. (1988). On multiparameter distributions of order k, *Ann. Inst. Statist. Math.*, **40**, 467–475.

Philippou, A. N., Antzoulakos, D. L. and Tripsiannis, G. A. (1989). Multivariate distributions of order k, *Statist. Probab. Lett.*, **7**, 207–216.

Discrete Multivariate Model Building Based on a Statistical Fundamental Difference Equation

Takahiro Tsuchiya and Tadashi Matsunawa
(The Institute of Statistical Mathematics)

Discrete multivariate statistical model building theory is nonparametrically discussed based on a statistical fundamental difference equation. The difference equation is derived as a special case of the general nonparametric statistical uncertainty relation obtained by the second author of the present paper previously. By utilizing the equation the general form of the probability function to a nonparametric minimum statistical uncertainty distribution is given. The model building process is accomplished with the help of the newly introduced total difference operation and the indefinite total summation operation. Further, several prototype multivariate discrete models, including multinomial distribution, multivariate hypergeometric distribution and matrix discrete distributions of order k, are constructed by making use of the general theorem given in this paper.

Key words: Discrete multivariate statistical models, nonparametric statistical model building, statistical fundamental difference equation, nonparametric statistical uncertainty relation, minimum statistical uncertainty distribution, total difference operation, indefinite total summation operation.