一般超幾何分布

清水良一

(1968年12月 受付)

Generalized Hypergeometric Distributions

Ryoichi Shimizu

This is "a taxology" of the generalized hypergeometric distributions, of which study has a long history. Classifying the distributions, we intend to clarify the properties of each of the generalized hypergeometric distributions, with the special attention to the relation among themselves as well as to the other distributions. It consists of § 1. introduction, § 2. classification, § 3. derivation of the generalized hypergeometric distributions, § 4. form of the distributions and moments, § 5. truncated distributions, § 6. approximation, and § 7. the distribution of $\pm X \pm p$.

The Institute of Statistical Mathematics

§1. 序

実数 α および，正の整数 x にたいして，

$$(\alpha)_0 = 1$$
$$(\alpha)_x = \alpha(\alpha+1)\cdots\cdot(\alpha+x-1) \tag{1.1}$$

とおく．超幾何級数，

$${}_pF_q(\alpha_1, \cdots\cdot, \alpha_p;\ \beta_1, \cdots\cdot, \beta_q;\ \theta) = \sum_{x=0}^{\infty} \frac{(\alpha_1)_x \cdots\cdot (\alpha_p)_x}{x!\,(\beta_1)_x \cdots\cdot (\beta_q)_x} \theta^x \tag{1.2}$$

が有限の値に収束し，しかも各項が非負であるときに，

$$\Pr\{X = x\} = \frac{(\alpha_1)_x \cdots\cdot (\alpha_p)_x}{x!\,(\beta_1)_x \cdots\cdot (\beta_q)_x} \theta_x \Big/ {}_pF_q(\alpha_1, \cdots\cdot, \alpha_p;\ \beta_1, \cdots\cdot, \beta_q;\ \theta) \tag{1.3}$$

によって確率分布族が定義される．これを一般超幾何級数分布と呼び $F(\alpha_1, \cdots\cdot, \alpha_p; \beta_1, \cdots\cdot, \beta_q|\theta)$ で表わす．$\theta=1$ のときは簡単に $F(\alpha_1, \cdots\cdot, \alpha_p;\ \beta_1, \cdots\cdot, \beta_q)$ と書くことにする．γ が，0でも負の整数でもないとき，$F(\alpha, \gamma; \gamma|\theta)$ や $F(\gamma; \gamma|\theta)$ は γ に無関係である．これらをそれぞれ，$F(\alpha|\theta)$，$F(1; 1|\theta)$ などと書く，一般超幾何級数分布の最も簡単な例は，$n, k>0$，n を整数として，

二項分布	$F(-n\,|-\theta)$	$(\theta > 0)$,
負の二項分布	$F(k\,|\,\theta)$	$(1 > \theta > 0)$,
ポアソン分布	$F(1; 1\,|\,\theta)$	$(\theta > 0)$,
対数級数分布を左にずらしたもの，	$F(1, 1; 2\,|\,\theta)$	$(1 > \theta > 0)$

などである．

以下においてわれわれが興味をもつのは，$F(\alpha, \beta; \gamma)$ であり，これを一般超幾何分布と呼ぶ．これは，(i) γ が，0でも負の整数でもないか，または (ii) $\gamma=-m$ は負の整数であるが，α または β のどちらかが，やはり負の整数，たとえば $\beta=-n$ であり，しかも $n \leqq m$ のときにのみ定義され，このとき，

$$Pr\{X = x\} = \frac{(\alpha)_x\,(\beta)_x}{x!\,(\gamma)_x} \Big/ {}_2F_1(\alpha, \beta; \gamma; 1) \tag{1.4}$$

で与えられる．ただし，つぎの規約を設ける．

規 約 I. $m \geqq n > 0$ が整数のとき，

$$ {}_2F_1(\alpha, -n; -m; \theta) = \sum_{x=0}^{n} \frac{(\alpha)_x (-n)_x}{x!\,(-m)_x} \theta^x \tag{1.5}$$

とし，この場合も含めて，

$$ {}_2F_1(\alpha, \beta; \gamma; \theta) = \sum_{x=0}^{\infty} \frac{(\alpha)_x (\beta)_x}{x!\,(\gamma)_x} \theta^x \tag{1.6}$$

と書く．この規約によって，$F(\alpha, -n; -n|\theta)$ と，$F(\alpha, 1; 1|\theta)$ は区別される．

一般超幾何分布の研究は遠く前世紀以前からはじまっていて，1920 年代ごろまでには，モーメントを求める（〔1〕，〔16〕，〔19〕），ピアソン系の曲線で近似する（〔4〕，〔5〕），分布関数の近似計算（〔4〕），二次元への拡張（〔18〕），などかなり詳細な研究がなされている．またこの分布を導くモデルとして，抽出確率が一定でない場合のサンプリング（せまい意味での超幾何分布）の他，ラプラスの succession law と呼ばれるもの（実は，いわゆる負の超幾何分布）などが知られていたが，この頃までの研究は，主として“形態学”である．ところで (1.4) の右辺は，記号 $\binom{k}{l}$ を適当に定義することによって，

$$ \binom{a}{x}\binom{b}{n-x} \Big/ \binom{a+b}{n} \tag{1.7}$$

と書くことができる（(1.4) と (1.7) がまったく同じという訳ではない）．a, b, n が正の整数のとき，よく知られた超幾何分布であるが，これらが負あるいは非整数の場合でも，これに適するモデルがいろいろと発見された (cf. §3)．C. D. and A. W. Kemp [13] は，a, b, n を実数の範囲内で動かしたとき (1.7) が非負整数上の確率分布になるための条件を求め，これを一般超幾何分布という名で統一し，分類した．この分類は (1.4) についてすでになされていたもの（形態学的分類〔5〕）と本質的に同じであるが，(1.4) と (1.7) のちがいが，あいまいのまま処理された為に不完全であり，十分に整理されていないので見にくい．K. Sarkadi [21] は Kemp の分類の不備を改め，(1.7) をいろいろに変形して，モデルとの対応関係をある程度整理した．さらに $a \geqq n > b > 0$ がすべて整数のとき，(1.7) が $n-b, n-b+1, \cdots, n$ の上の確率分布になっているのに，これが Kemp の分類の中に入っていないのは不つごうであり，これも一般超幾何分布に入れるべきであると主張した．さて，われわれがこゝでやろうとしているのは，(1.4) ないし (1.7) で与えられる，一般超幾何分布の“分類学”である．これは単なる分類ではない．分類された各グループ（タイプと呼ぶ）に属する分布の性質，相互の関係，それらを導くモデル，他の諸分布との関係等々を明確にすることである．以下の各節は，§2. 分類，§3. 分布の導出，§4. 分布の形，モーメント，§5. 打ち切り分布，§6. 他の分布による近似，§7. 上述の K. Sarkadi の主張にたいする考察，などからなる．これらを通して，一般超幾何分布の性質が，かなりはっきりしてくるであろう．

§2. 分 類

一般超幾何分布 $F(\alpha, \beta; \gamma)$ は，超幾何級数 ${}_2F_1(\alpha, \beta; \gamma; \theta)$ が意味をもち，$\theta=1$ で収束，しかも

$$ \frac{(\alpha)_x (\beta)_x}{x!\,(\gamma)_x} \geqq 0, \quad x = 0, 1, 2, \cdots \tag{2.1}$$

であるときに定義される．$\beta=-n$, $\gamma=-m$, $n \leqq m$ のときは，前節の規約 I によって，(2.1) は，$x=0, 1, \cdots, n$ についてだけ考えればよい．α も β も負の整数でないときは，ベキ級数 ${}_2F_1(\alpha, \beta; \gamma; \theta)$ の収束半径は 1 である．$\theta=1$ で収束するためには

$$ \gamma > \alpha + \beta \tag{2.2}$$

が必要で十分である．さらに $F(\alpha,\beta;\gamma)$ が，α,β に関して対称であることを考慮すると，一般超幾何分布 $F(\alpha,\beta;\gamma)$ は，つぎのいずれかの形に書かれる．以下，ξ,η,ζ は正の数，ε,ρ,δ は0と1の間（両端を含まない）の数，n は正の整数を表わす．

分　類

タイプ A-1,

$$F(-\xi,\ -n;\zeta),\quad \xi>n-1,\quad x=0,1,\cdots,n.$$

タイプ A-2,

$$F(\xi,\ -n;-\zeta),\quad \zeta>n-1,\quad x=0,1,\cdots,n.$$

タイプ B-1,

$$F(-n+\varepsilon,\ -n+\delta;\zeta),\quad x=0,1,\cdots$$

タイプ B-2,

$$F(\varepsilon,\ -n+\delta;-n+\rho),\quad \rho>\varepsilon+\delta,\quad x=0,1,\cdots$$

タイプ B-3,

$$F(\xi,\ \eta,\ \zeta),\quad \zeta>\xi+\eta,\quad x=0,1,\cdots$$

$\gamma>\alpha+\beta$ のとき，

$$_2F_1(\alpha,\ \beta;\gamma;1)=\frac{\Gamma(\gamma)\,\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)} \tag{2.3}$$

である（ガウスの定理；例えば［3］を見よ）．ただし，つぎの規約 II を適用する．

規　約 II.　p,q が非負の整数のとき，

$$\frac{\Gamma(-p)}{\Gamma(-p-q)}=(-1)^q\,(p+1)_q=(-1)^q\frac{(p+q)!}{p!} \tag{2.4}$$

と定義する．あるいは，

$$\frac{\Gamma(-p)}{\Gamma(-p-q)}=\lim_{\varepsilon\to 0}\frac{\Gamma(-p-\varepsilon)}{\Gamma(-p-q-\varepsilon)} \tag{2.5}$$

といっても同じことである．

このことを使うと，タイプ A-1, B-1, B-2, B-3 については，(1.4) は，

$$Pr\{X=x\}=\frac{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}{\Gamma(\gamma)\,\Gamma(\gamma-\alpha-\beta)}\cdot\frac{(\alpha)_x\,(\beta)_x}{x!\,(\gamma)_x} \tag{2.6}$$

となる．タイプ A-2 では $\gamma<\alpha+\beta$ となるが，この場合でも (2.6) の形に書けることは，初等的に示される．さらに，実数 k,l について，記号 $\binom{k}{l}$ を，

$$\binom{k}{l}=\frac{\Gamma(k+1)}{\Gamma(l+1)\cdot\Gamma(k-l+1)}$$

で定義すると，(2.6) は，

$$P_r\{X=x\}=\frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}} \tag{2.7}$$

とも書ける．ここで，

$$a=-\alpha,\quad b=\gamma-\beta-1,\quad n=-\beta$$

である．

逆に，(2.7) は，

$$\alpha=-a,\quad \beta=-n,\quad \gamma=b-n+1$$

とおいて，形式的に (2.6) となるが，これは必ずしも，われわれの定義した一般超幾何分布にはならない．たとえば，$a \geqq n = b+1$ で，a, b, n が正の整数のとき，(2.7) は，$x=1, 2, \cdots n$ で正，加えて1になるが，このとき，$\gamma=0$ となる．これについては，§7 を参照．

§3. 一般超幾何分布を導くモデル

タイプ A-1 $F(-\xi, -n; \zeta)$

(2.7) で書いたとき，$a, b \geqq n > 0$, n は整数．

(I) ポイヤのつぼ： つぼの中に白玉と赤玉がそれぞれ，w 個，r 個入っている．このつぼから玉を1個とりだしてもとに戻す．さらに，いまとりだされたものと同じ色の玉を c 個 ($c \leqq 0$ も許すとする) つぼに入れる．つぼの中の玉は，常に等確独でとりだされるものとする．

(i) $c=-p<0$ のとき，この操作を n 回繰り返して，赤玉が x 回とりだされる確率は，

$$\binom{n}{x} \frac{\prod_{i=0}^{x-1}(r-ip) \prod_{j=0}^{n-x-1}(w-jp)}{\prod_{k=0}^{n-1}(r+w-kp)} \tag{3.1}$$

である．この式の分母子を p^n で割って，

$$\frac{\binom{r/p}{x}\binom{w/p}{n-x}}{\binom{(r+w)/p}{n}}$$

となる [21]．これは $F(-r/p, -n; w/p-n+1)$ である．

(ii) とくに $p=1$ なら，もとに戻さぬ標本抽出であり，せまい意味での超幾何分布，

$$\frac{\binom{r}{x}\binom{w}{n-x}}{\binom{r+w}{n}} \qquad x=0, 1, \cdots, n$$

すなわち，$F(-r, -n; \omega-n+1)$ になる．(古くは，たとえば [17])，この形に帰着できるものは極めて多い．

(II) 連の数 [14]： m 個の 0 と n 個の 1 の配列を考える．$m \geqq n$ として一般性を失わない．ちょうど k 個の 0 の連があるとき，$X=k-1$ とおく．すべての配列が同じ確率をもつとすれば，

$Pr\{X=x\} = Pr\{x+1$ 個の 0 の連がある$\} =$

$$\frac{\binom{m-1}{x}\binom{n+1}{x+1}}{\binom{m+n}{m}} = \begin{cases} \dfrac{\binom{m-1}{x}\binom{n+1}{n-x}}{\binom{m+n}{n}} & \text{if } m>n, \\[2ex] \dfrac{\binom{n}{x}\binom{n}{n-1-x}}{\binom{2n}{n-1}} & \text{if } m=n \end{cases}$$

$m>n$ のとき $F(-m+1, -n; 2)$，$m=n$ のとき $F(-n, -n+1; 2)$ である．

(III) 条件つき分布： X, Y が独立で，それぞれ，二項分布，

$$\binom{k}{x}p^x(1-p)^{k-x}, \quad \binom{l}{y}p^y(1-p)^{l-y}$$

に従うとき，$Z=X+Y=n$ が与えられた時の X の条件につき分布は，

$$pr\{X=x \mid X+Y=n\}=\frac{\binom{k}{x}p^x(1-p)^{k-x}\binom{l}{n-x}p^{n-x}(1-p)^{l-n+x}}{\binom{k+l}{n}p^n(1-p)^{k+l-n}}$$

$$=\frac{\binom{k}{x}\binom{l}{n-x}}{\binom{k+l}{n}}$$

これは $F(-k,\ -n;\ l-n+1)$ である．

タイプ A-2 $F(\xi,\ -n;\ -\zeta)$

(2.7) で書いたとき，$n>0>a,\ b$ その為，負の超幾何分布と呼ばれることがある．

(I) ポイヤのつぼで $c=p>0$ とする．n 回の試行中赤が x 回とりだされる確率は，

$$\binom{n}{x}\frac{\prod_{i=0}^{x-1}(r+ip)\prod_{j=0}^{n-x-1}(w+jp)}{\prod_{k=0}^{n-1}(r+w+kp)} \tag{3.2}$$

である．分母子を p^n で割って整理すれば，

$$\frac{\binom{-r/p}{x}\binom{-w/p}{n-x}}{\binom{-(r+w)/p}{n}}$$

をうる．これは $F(r/p,\ -n;\ -w/p-n+1)$ である [21]．

(II) ポイヤのつぼ (逆サンプリング)： $c=-1$ とする．k 個の赤玉が出るまで試行を繰り返す．その間に出た白玉の個数が x である確率は，

$$\binom{x+k-1}{x}\frac{\prod_{i=0}^{k-1}(r-i)\prod_{j=0}^{x-1}(w-j)}{\prod_{k=0}^{x+k-1}(r+w-k)} \tag{3.3}$$

である．これを整理して，

$$\frac{\binom{-k}{x}\binom{k-r-1}{w-x}}{\binom{-r-1}{w}}$$

すなわち，$F(k,\ -w;\ k-r-w)$ をうる．(当然 $r\geqq k$，したがって，$-(k-r-w)>w-1$ である)．

(III) タイプ A-1 (II) の，0と1の配列の問題で，N 番目の0の前に x 個の1がある確率は，

$$\binom{x+N-1}{x}\binom{m+n-N-x}{n-x}\Big/\binom{m+n}{m}$$
$$=\binom{-N}{x}\binom{-m+N-1}{n-x}\Big/\binom{-m-1}{n}$$

これは，$F(N,\ -n;\ -m+N-n)$ である．

(IV)　条件につき分布： X, Y が独立で，それぞれ負の二項分布，

$$\binom{-k}{x}(-p)^x(1-p)^k \quad \text{および} \quad \binom{-l}{y}(-p)^y\cdot(1-p)^l$$

に従うとする．$Z=X+Y=n$ が与えられたときの，X の条件つき分布は，

$$P_r\{X=x \mid X+Y=n\}=\frac{\binom{-k}{x}\binom{-l}{n-x}}{\binom{-k-l}{n}}$$

これは，$F(k, -n; -l-n+1)$ である．k, $l>0$ でよいからタイプ A-2 はすべてこれで尽される (cf. タイプ A-1 (III))．

(V)　二項分布のベータ分布による混合： 二項分布 $\binom{n}{x}p^x(1-p)^{n-x}$ をベータ分布，

$$p^{\xi-1}(1-p)^{\zeta-n}/B(\xi, \zeta-n+1), \quad (\xi>0,\ \zeta-n+1>0)$$

で混合すると，

$$\int_0^1\binom{n}{x}p^x(1-p)^{n-x}p^{\xi-1}(1-p)^{\zeta-n}\,dp/B(\xi, \zeta-n+1)$$

$$=\binom{n}{x}\frac{B(\xi+x, \zeta-x+1)}{B(\xi, \zeta-n+1)}$$

$$=\frac{\Gamma(-\zeta+n)\,\Gamma(-\zeta-\xi)}{\Gamma(-\zeta)\,\Gamma(-\zeta-\xi+n)}\cdot\frac{(\xi)_x\,(-n)_x}{x!\,(-\zeta)_x}$$

これはタイプ A-2 $F(\xi, -n; -\zeta)$ である．タイプ A-2 はすべてこのようにしても得られる．

(VI)　exceedances の分布： 連続な分布 $F(x)$ からの大きさ N の標本について，大きい方から m 番目のものを $Y^{(m)}$ とする．さらに大きさ n の標本をとり，そのうち $Y^{(m)}$ より大きいものの個数を X とする．

$$Pr\{X=x \mid F(Y^{(m)}=p\}=\binom{n}{x}p^x(1-p)^{n-x} \tag{3.4}$$

であるが，$F(Y^{(m)})$ の分布は，(0,1) 上の一様分布からの大きさ N の標本についての，大きい方から m 番目のものの分布に等しく，したがって，その確率密度関数は，F に無関係に，

$$\binom{N}{m}m\,p^{N-m}(1-p)^{m-1} \tag{3.5}$$

である．それ故，

$$P_r\{X=x\}=\int_0^1\binom{n}{x}p^x(1-p)^{n-x}\cdot\binom{N}{m}m\,p^{N-m}(1-p)^{m-1}\,dp$$

$$=\frac{\binom{N}{m}\binom{n}{x}}{\binom{n+N-1}{m+x-1}}\cdot\frac{m}{N+m} \tag{3.6}$$

となる．exceedances の分布として知られているこの分布 ([8], [20]) がタイプ A-2, $F(m, -n; -(n+N-m))$ であることは，(3.5) が，ベータ分布 $p^{N-n}(1-p)^{m-1}/B(N-m+1, m)$ であるとに注意すれば，(V) で述べたものの特別な場合になっていることから分る．

(VII)　予測： 成功の確率が p のベルヌウイ試行を考える．p は事前分布，

$$Pr\left\{p=\frac{i}{k}\right\}=\frac{1}{k+1}, \quad i=0,1,\cdots,k,$$

をもつとする．N 回の試行で m 回の成功が得られた（事象 A と呼ぶことにする）とき，p の事後分布は，

$$Pr\left\{p=\frac{i}{k}\mid A\right\}=\frac{\binom{N}{m}\left(\frac{i}{k}\right)^m\left(1-\frac{i}{k}\right)^{N-m}}{\sum_{j=0}^{k}\binom{N}{m}\left(\frac{j}{k}\right)^m\left(1-\frac{j}{k}\right)^{N-m}} \tag{3.7}$$

したがって A が与えられた時，将来の n 回の試行中の成功の回数 X の条件つき分布は，

$$\begin{aligned}Pr\{X=x\mid A\}&=\sum_{i=0}^{k}Pr\left\{X=x,\quad p=\frac{i}{k}\mid A\right\}\\&=\sum_{i=0}^{k}Pr\left\{X=x\mid p=\frac{i}{k}\right\}\cdot Pr\left\{p=\frac{i}{k}\mid A\right\}\\&=\binom{n}{x}\frac{\sum_0^k\left(\frac{i}{k}\right)^{m+x}\cdot\left(1-\frac{i}{k}\right)^{N+n-m-x}}{\sum_0^k\left(\frac{i}{k}\right)^{m}\left(1-\frac{i}{k}\right)^{N-m}}\end{aligned} \tag{3.8}$$

である．$k\to\infty$ として，この式から，

$$\binom{n}{x}\frac{\int_0^1 p^{m+x}\,(1-p)^{N+n-m-x}\,dp}{\int_0^1 p^{m}\,(1-p)^{N-m}\,dp} \tag{3.9}$$

を得る．これはラプラスの succession law として知られている．分母子をベータ関数で表わして，これが (V) の特別の場合であることが分る．([21]，[26])．

(VIII)　ランダム-ウォーク：　A, B 二人で，つぎのようなゲームを行う．サイコロを振って，偶数が出れば A が，奇数が出れば B が，それぞれ 1 点を相手からもらう．ゲームの回数はランダムに決めることとし，その為にゲームの前に，0 から n までの文字を書いたチップ 1 枚ずつを入れた袋から 1 枚のチップをとり出す．そこに書いてある数字を 2 倍したものをゲームの回数とする．0 と書いてあればゲームをせずにそのまま引分けとする．いまゲームが終了した時に両者の得点が 0 であったとして，袋からとり出されたチップに書かれていた数字 N の条件つき分布を求めて見よう．

X_i，$i=1,2,\cdots\cdots$ は独立で，いずれも二点分布，

$$Pr\{X_i=1\}=Pr\{X_i=-1\}=1/2 \tag{3.10}$$

に従うとする．ゲームを $2k$ 回行った時点での A の得点は，

$$S_k=X_1+\cdots\cdots+X_{2k} \tag{3.11}$$

と表わされ，

$$Pr\{S_k=0\}=\binom{2k}{k}2^{-2k}=\frac{\left(\frac{1}{2}\right)_k}{k!} \tag{3.12}$$

である．

$$Pr\{N=k\}=\frac{1}{(n+1)},\quad k=0,1,\cdots\cdots,n \tag{3.13}$$

だから，

$$Pr\{S_N=0\}=\sum_{k=0}^{n}Pr\{S_N=0\mid N=k\}Pr\{N=k\}$$

$$=\frac{1}{n+1}\sum_{k=0}^{n}\frac{\left(\frac{1}{2}\right)_k}{k!}=\frac{1}{n+1}\,{}_2F_1\left(\frac{1}{2},\ -n;\ -n;\ 1\right)$$

したがって，ベイズの定理から，

$$Pr\{N=x\mid S_N=0\}=\frac{\left(\frac{1}{2}\right)_x(-n)_x}{x!\,(-n)_x}\Big/{}_2F_1\left(\frac{1}{2},\ -n;\ -n;\ 1\right) \tag{3.14}$$

これは $F\left(\frac{1}{2},\ -n;\ -n\right)$ である．

タイプ B-3 $F(\xi,\ \eta;\ \zeta)$

(2.7) で書いたとき，$a<0,\ n<0,\ b>-1$

(I) 負の二項分布のベータ分布による混合： 負の二項分布，

$\binom{-\xi}{x}(-p)^x(1-p)^\xi$ の p を第一種ベータ分布 $\dfrac{p^{\eta-1}(1-p)^{\zeta-\xi-\eta-1}}{B(\eta,\ \zeta-\xi-\eta)}$ で混合する．

$$\int_0^1\binom{-\xi}{x}(-p)^x(1-p)^\xi p^{\eta-1}(1-p)^{\zeta-\xi-\eta-1}\frac{dp}{B(\eta,\ \zeta-\xi-\eta)}$$

$$=(-1)^x\binom{-\xi}{x}\frac{B(\eta+x,\ \zeta-\eta)}{B(\eta,\ \zeta-\xi-\eta)}$$

$$=\frac{\Gamma(\zeta-\xi)\,\Gamma(\zeta-\eta)}{\Gamma(\zeta)\,\Gamma(\zeta-\xi-\eta)}\cdot\frac{(\xi)_x(\eta)_x}{x!\,(\zeta)_x} \tag{3.15}$$

これは $F(\xi,\ \eta;\ \zeta)$ である．負の二項分布を，

$$\binom{-\xi}{x}(-\theta)^x(1+\theta)^{-\xi-x}\qquad\theta>0$$

と書いたときは，第二種のベータ分布，

$$\theta^{\eta-1}(1+\theta)^{\zeta-\xi}/B(\eta,\ \zeta-\xi-\eta)$$

で混合して同じ結果をうる．タイプ B-3 は，すべてこの方法で得られる．

(II) exceedances の分布 (逆サンプリング)： タイプ A-2 (VI) で，第二標本は，$Y^{(m)}$ より大きいものが，k 個現われるまでとることにして，この時までに現われる，$Y^{(m)}$ より小さいものの個数を X とすると，

$$Pr\{X=x\mid F(Y^{(m)})=p\}=\binom{x+k-1}{x}p^x(1-p)^k$$

したがって，前と同様，

$$Pr\{X=x\}=\int_0^1\binom{x+k-1}{x}p^x(1-p)^k\binom{N}{m}m\,p^{N-m}(1-p)^{m-1}dp$$

$$=\frac{\binom{N}{m}\binom{x+k-1}{x}}{\binom{-k+N-1}{m+x-1}}\times\frac{m}{N+m}$$

これが (I) の特別の場合であることは，タイプ A-2 (VI) が (V) の特別の場合であるのと同じである．$F(k,\ N-m+1;\ N+k+1)$.

とくに，$N=m=k=1$ とすると，$F(1,\ 1;3)$

$$Pr\{X=x\}=\frac{1}{(x+1)(x+2)}$$

となるが，たとえば，ルーレットを1回まわし，そのあと，第1回目より大きい角度のところで静止するまでに要した試行回数がこの分布に従うことになる．なお〔15〕を見よ．

(III) ランダム-ウォーク： (i) タイプ A-2 (VIII) のゲームで，ゲームの回数を定める変数 N を，

$$Pr\{N=n\}=\frac{1}{(n+1)(n+2)} \tag{3.16}$$

で与える（例えば，ルーレットを使ってもよい．cf. (II)）．

$$Pr\{S_N=0\}=\sum_{n=0}^{\infty}Pr\{S_N=0\mid N=n\}\cdot Pr\{N=n\}$$

$$=\sum_{n=0}^{\infty}\frac{\left(\frac{1}{2}\right)_n}{n!}\cdot\frac{1}{(n+1)(n+2)}=\frac{1}{2}\sum_{n=0}^{\infty}\frac{(1)_n\left(\frac{1}{2}\right)_n}{n!(3)_n}$$

$$=\frac{1}{2}\,{}_2F_1\left(1,\ \frac{1}{2};3;1\right)$$

従って，ゲーム終了時に，A, B の得点が 0 ($S_N=0$) であることが分ったとき，N の条件つき分布は，

$$Pr\{N=n\mid S_N=0\}=\frac{(1)_n\left(\frac{1}{2}\right)_n}{n!(3)_n}\Big/\,{}_2F_1\left(1,\ \frac{1}{2};3;1\right)$$

これは $F\left(1,\ \frac{1}{2};\ 3\right)$ である．

(ii) A, B のゲームとは独立に，C, D が同じゲームを行う，確率変数の系 Y_i, $i=1,\cdots\cdot$ は互いにまた $X_i, i=1,\cdots\cdot$ とも独立で X_i と同じ二点分布とすると，2n 回のゲームが終了したところで C の得点は，$T_n=Y_1+\cdots\cdot+Y_{2n}$ で与えられる．(i) の場合と同様にして，

$$Pr\{S_N=T_N=0\mid N=n\}=Pr\{S_n=0\}Pr\{T_n=0\}=\left[\frac{\left(\frac{1}{2}\right)_n}{n!}\right]^2$$

となるから，N の分布を前と同じに (3.17) とすれば，

$$Pr\{S_N=T_N=0\}=\sum_{n=0}^{\infty}Pr\{S_n=T_n=0\}Pr\{N=n\}$$

$$=\sum_{n=0}^{\infty}\left(\frac{\left(\frac{1}{2}\right)_n}{n!}\right)^2\frac{1}{(n+1)(n+2)}=\frac{1}{2}\,{}_2F_1\left(\frac{1}{2},\ \frac{1}{2};3;1\right) \tag{3.17}$$

よって，

$$Pr\{N=n\mid S_N=T_N=0\}=\frac{\left(\frac{1}{2}\right)_n\left(\frac{1}{2}\right)_n}{n!(3)_n}\Big/\,{}_2F_1\left(\frac{1}{2},\ \frac{1}{2};3;1\right)$$

これは $F\left(\frac{1}{2},\ \frac{1}{2};\ 3\right)$ である．

(IV) 分枝過程の消滅までの世代数： 個体が単位時間ごとに分裂，消滅を繰り返す過程を考える．1個の個体が分裂してできる個体の数は，すべて互いに独立で，同じ分布 F に従うとする．1個の個体から始まって，n 単位時間後の個体の数を Z_n で表わすと，Z_n は F と Z_{n-1} だけで決まる，: $\{X_{i,j}\}$, $i, j=1, 2,\cdots\cdot$ は互いに独立で，いずれも F に従うなら，

$$Z_0=1,\quad Z_n=X_{n,1}+\cdots\cdots+X_{n,\,Z_{n-1}}$$

と表わされる．$Pr\{X_{i,j}=0\}>0$ のとき，確率変数 N を，

$$N=n\Leftrightarrow Z_0\neq 0,\ Z_1\neq 0,\cdots\cdots,\ Z_n\neq 0,\ Z_{n+1}=0$$

で定義し，これをこの分枝過程の消滅までの世代数と呼ぶ．F が幾何分布，すなわち，

$$Pr\{X_{i,j}=x\}=p^x q \qquad x=0,1,2,\cdots\cdots$$

のとき，N が $F(1,\ 1;3)$，すなわち，

$$Pr\{N=n\}=\frac{1}{(n+1)(n+2)}$$

となる（[25]）．

タイプ B-1　$F(-n+\varepsilon,\ -n+\delta;\zeta)$

タイプ B-2　$F(\varepsilon,\ -n+\delta;\ -n+\rho)$

この二つのタイプは，それぞれタイプ A-1，A-2 の母数 $\beta=-n$ を非整数 $-n+\delta$ で置きかえたものであるが，いまのところ適当なモデルは見当らない．ただタイプ B-1 の極めて特殊なものについて，つぎのものが考えられる．

タイプ B-1　$F\left(-\frac{1}{2},\ -\frac{1}{2};\ 3\right)$

タイプ B-3 (III) (ii) のゲームで，事象，$S_1\neq 0,\cdots\cdots,\ S_{n-1}\neq 0,\ S_n=0,\ T_1\neq 0,\cdots\cdots,\ T_{n-1}\neq 0,\ T_n=0$ を E_n で表わす．これは二組とも，$2n$ 回目に初めて各プレーヤーの得点が 0 になるという事象である．事象 $S_0=0$ を E_0 で表わす．これは確実に起る事象である．すると，

$$P(E_n)=\begin{cases}1 & \text{if } n=0\\ \left|\dfrac{1}{2n}\dbinom{2n-2}{n-1}\cdot 2^{-2(n-1)}\right|^2 & \text{if } n\geqq 1\end{cases}\tag{3.18}$$

である（[14]）．

$$\frac{1}{2n}\binom{2n-2}{n-1}2^{-2(n-1)}=\frac{-\left(-\frac{1}{2}\right)_n}{n!}\qquad n\geqq 1$$

に注意すると，N が (3.17) の分布に従うとき，

$$\begin{aligned}Pr(E_N)&=\sum_{n=0}^{\infty}P(E_n)\,Pr\{N=n\}\\&=\sum_{n=0}^{\infty}\frac{\left(-\frac{1}{2}\right)_n\left(-\frac{1}{2}\right)_n}{n!\,n!}\cdot\frac{1}{(n+1)(n+2)}\\&=\frac{1}{2}\,{}_2F_1\left(-\frac{1}{2},\ -\frac{1}{2};3;1\right)\end{aligned}\tag{3.19}$$

である．よって，

$$Pr\{N=n\mid E_N\}=\frac{\left(-\frac{1}{2}\right)_n\left(-\frac{1}{2}\right)_n}{n!\,(3)_n}\Big/\,{}_2F_1\left(-\frac{1}{2},\ -\frac{1}{2};3;1\right)\tag{3.20}$$

§4　分布の形とモーメント

X が一般超幾何分布 $F(\alpha,\ \beta;\gamma)$ に従うとき，

$$\frac{Pr\{X=x+1\}-Pr\{X=x\}}{Pr\{X=x\}}=-\frac{\Delta_0 x+\gamma-\alpha\beta}{(x+1)(x+\gamma)}\tag{4.1}$$

$$\Delta_0=\gamma-\alpha-\beta+1$$

である．

タイプ A-1, B-1, B-3 では，$\gamma>0$, $\Delta_0>0$ だから分布は一つのモード（単峰）をもつ．$\gamma>\alpha\beta$ なら，モードは $x=0$ で，$Pr\{X=x\}$ が x の単調減少関数になる．タイプ A-2 では，$\gamma=-\zeta<0$ であるが $\Delta_0<0$ のときモード（凸型），$\Delta_0>0$ のとき，アンチモードをもつ，（凹型），$\gamma>\alpha\beta$ なら単調増加で $x=n$ がモード．$\Delta_0=\gamma-\alpha\beta=0$ のとき一様分布になる．また $\gamma\to\beta+1$, すなわち $\zeta\to n-1$ のとき，$x=n$ に退化した分布に近ずく．タイプ B-2 は $\gamma=-n+\rho<0$, $\Delta_0>0$ だから，$x=n$ と $0\leq x<n-1$ に山がある（双峰）．$\gamma>\alpha\beta$, すなわち $\varepsilon(n-\delta)>n-\delta$ のとき，$x=0$ が1つのモードである．またタイプ A-2 の場合に対応して，$\rho\to1$, $\varepsilon\to1$ のとき，$x=n$ に退化した分布に近ずく．

一般超幾何分布 $F(\alpha, \beta; \gamma)$ の確率母関数は，

$$\varphi(\theta)={}_2F_1(\alpha, \beta, \gamma; \theta)/{}_2F_1(\alpha, \beta; \gamma; 1)$$
$$=\frac{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}{\Gamma(\gamma)\cdot\Gamma(\gamma-\alpha-\beta)}\cdot\sum_{x=0}^{\infty}\frac{(\alpha)_x\,(\beta)_x}{x!\,(\gamma)_x}\theta^x \tag{4.2}$$

で与えられる．これは $0<\theta<1$ の範囲で，和の記号下で微分できて，

$$\varphi^{(k)}(\theta)=\frac{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}\cdot\sum_{x=0}^{\infty}x(x-1)\cdots\cdot(x-k+1)\cdot\frac{(\alpha)_x\cdot(\beta)_x}{x!\,(\gamma)_x}\theta^x$$
$$=\frac{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}{\Gamma(\gamma)\,\Gamma(\gamma-\alpha-\beta)}\cdot\frac{(\alpha)_k\,(\beta)_k}{(\gamma)_k}\sum_{x=0}^{\infty}\frac{(\alpha+k)_x\,(\beta+k)_x}{x!\,(\gamma+k)_x}\theta^x \tag{4.3}$$

これは $0<\theta<1$ で成りたつが，$\theta=1$ でも成りたつためには，β（または α）が負の整数であるか，または

$$\gamma>\alpha+\beta+k \tag{4.4}$$

であることが必要で十分．そしてこの時，

$$\nu_k=E\left[X(X-1)\cdots\cdot(X-k+1)\right]=\varphi^{(k)}(1)$$
$$=\frac{(\alpha)_k\,(\beta)_k}{(\gamma)_k}\cdot\frac{\Gamma(\gamma+k)\,\Gamma(\gamma-\alpha-\beta-k)}{\Gamma(\gamma)\,\Gamma(\gamma-\alpha-\beta)}$$
$$=\frac{(\alpha)_k\,(\beta)_k}{(\gamma-\alpha-\beta-k)_k} \tag{4.5}$$

とくに，ν_1 が存在するとき，

$$E(X)=\nu_1=\alpha\beta/\Delta, \tag{4.6}$$
$$\Delta=\gamma-\alpha-\beta-1,$$

また ν_2 が存在するとき，

$$Var\,X=\nu_2+\nu_1-\nu_1^2=\frac{\alpha\beta(\Delta+\alpha)(\Delta+\beta)}{\Delta^2(\Delta-1)} \tag{4.7}$$

タイプ A-1, B-1 では n$\geqq1$ のとき　$E(X)\geqq Var\,X$,

タイプ B-3 では　$E(X)\leqq Var\,X$

である．タイプ B-2 では1次以上のモーメントが存在しない．

§6. 条件つき分布．打ちきり分布

S を集合 $\{0, 1, \cdots\}$ の部分集合する．X が一般超幾何分布 $F(\alpha, \beta; \gamma)$ に従うとして，$X\in S$ が与えられた時の，X の条件つき分布は，

$$Pr\{X=x\mid X\in S\}=\frac{(\alpha)_x\,(\beta)_x}{x!\,(\gamma)_x}\Big/\sum_{x\in S}\frac{(\alpha)_x\,(\beta)_x}{x!\,(\gamma)_x} \tag{5.1}$$

である．これはタイプAでは $\{0,1,\cdots\cdots,n\}\cap S\neq\phi$，タイプBでは $S\neq\phi$ のとき意味をもつ．とくに $S=\{s,s+1,\cdots\cdots\}$ (タイプAでは $s\leqq n$) のとき，(5.1) は，

$$Pr\{X=x\mid X\geqq s\}=\frac{(\alpha)_x(\beta)_x}{x!\,(\gamma)_x}\Big/\sum_{x=s}^{\infty}\frac{(\alpha)_x(\beta)_x}{x!\,(\gamma)_x}$$

$$=\frac{(\alpha+s)_{x-s}(\beta+s)_{x-s}}{x!\,(\gamma+s)_{x-s}}\Big/\sum_{x=0}^{\infty}\frac{(\alpha+s)_x(\beta+s)_x}{(x+s)!\,(\gamma+s)_x}$$

$$x=s,\quad s+1,\cdots\cdots \tag{5.2}$$

となるが，これを $x=s$ の左側を打ちきった分布といい，$F(\alpha,\beta;\gamma\|s)$ であらわす．$F(\alpha,\beta;\gamma\|1)$ を，0-打ちきり分布と呼ぶ．

タイプ A-2 と B-3 では，$\alpha=1$ にとれるが，このときは興味がある．

$$(1+s)_y=(y+s)!/s! \tag{5.2}$$

に注意すれば，(5.2) は $\alpha=1$ のとき，

$$Pr\{X=x\mid X\geqq s\}=\frac{(\beta+s)_{x-s}}{(\gamma+s)_{x-s}}\Big/\sum_{x=0}^{\infty}\frac{(\beta+s)_x}{(\gamma+s)_x} \tag{5.4}$$

となるから，これを s だけ左にずらした．$Y=X-s$ の条件つき分布は，

$$\frac{(1)_y}{y!}\cdot\frac{(\beta+s)_y}{(\gamma+s)_y}\cdot\Big/\sum_{y=0}^{\infty}\frac{(1)_y(\beta+s)_y}{y!\,(\gamma+s)_y}$$

となる．これは $F(1,\ \beta+s;\gamma+s)$ である．

タイプ A-2 $F(1,\ -n;\ -\zeta\|s)$ を s だけ左にずらして，$F(1,\ -n+s;\ -\zeta+s)$． これはタイプ A-2 である．

タイプ B-3 $F(1,\ \eta;\zeta\|s)$ を s だけ左にずらして，$F(1,\ \eta+s;\zeta+s)$． これはタイプ B-3 である．

この性質は，幾何分布 $F(1|\theta)$ のよく知られた，もっと強い性質に対応する．つぎの節で述べるように，タイプ A-2, $F(1,\ -n;\ -\zeta)$，タイプ B-3, $F(1,\ \eta;\zeta)$ は，幾何分布 $F(1|\theta)$ で近似されるが，この幾何分布を，$x=s$ の左側で打ちきって，s だけ左にずらしたものは，もとのと同じ幾何分布になる．

当然のことながら，(5.2) の右辺の形の分布は，一般超幾何分布 $F(\alpha,\ \beta;\gamma)$ を打ちきって得られるものよりも広いことに注意しよう．たとえば，$F(0,\ \beta;\gamma)$ は $x=0$ に退化した分布になるが，(5.2) で $\alpha=0$ とおいたものはそうでない．超幾何数分布 $F(\alpha,\ \beta;\gamma|\theta)$ を $x=s$ の左側で打ちきったものを同様に，$F(\alpha,\ \beta;\gamma|\theta\|s)$ と書くことにすると，この分布は，

$$\frac{(\alpha+s)_{x-s}(\beta+s)_{x-s}}{x!\,(\gamma+s)_{x-s}}\theta^{x-s}\Big/\sum_{x=0}^{\infty}\frac{(\alpha+s)_x(\beta+s)_x}{(x+s)!\,(\gamma+s)_x}\theta^x \tag{5.5}$$

$$x=s,\quad s+1,\cdots\cdots$$

となる．$F(0|\theta)=F(0,\ 1;1|\theta)$ は原点に退化した分布になるが，(5.5) で $\alpha=0,\ s=\beta=\gamma=1$ とすれば，$0<\theta<1$ にたいして，

$$\frac{\theta^x}{x}\Big/-\log(1-\theta) \tag{5.6}$$

$$x=1,2,\cdots\cdots$$

である．これは，いわゆる対数級数分布である．これを $F(0|\theta\|1)$ と書くことができよう．(cf. §6. (III))．

また，タイプ A-1 $F(-\xi,\ -n;1)$ やタイプ B-3 $F(\xi,\ \eta;1)$ で形式的に $\xi=n=0$，ある

いは，$\xi=\eta=0$ としたものは，$x=0$ に退化したものであるが，(5.2) で，$\alpha=\beta=0$，$s=\gamma=1$ としたもの，すなわち，$F(0,\ 0;1\|1)$ と書かれるべきものはツェータ分布，

$$Pr\{X=x\}=\frac{1}{x^2}\Big/\sum_1^{\infty}\frac{1}{x^2}$$

$$=\frac{6}{\pi^2}\frac{1}{x^2} \qquad (5.7)$$

$$x=1,2,\cdots\cdot$$

である．

その他，$\alpha=-m$，$\beta=-n$，$\gamma=0$，$s=1$，$m\geqq n\geqq 3$ とすれば，(5.2) の右辺は (2.7) の形に書き直したとき，

$$\binom{m}{x}\binom{n-1}{n-x}\Big/\binom{m+n-1}{n} \qquad x=1,2,\cdots\cdot,n$$

となるが，これはタイプ A–1 $F(-m+1,\ -n+1;2)$ を 1 だけ右にずらしても得られる．(cf. §7(I))

いずれにしても (5.2) の右辺は一般超幾何分布がさらに拡張されるべきであることを暗示しているが，これについては別の機会に述べよう．

§6. 一般超幾何分布の近似

(I) 二項分布，負の二項分布による近似．

$$\frac{(\alpha)_x}{(\gamma)_x}=\left(\frac{\alpha}{\gamma}\right)^x\prod_{k=0}^{x-1}\frac{1+k/\alpha}{1+k/\gamma} \qquad (6.1)$$

だから，$\alpha,\ \gamma\to\pm\infty$，$\alpha/\gamma\to\theta$ のとき，

$$\frac{(\alpha)_x}{(\gamma)_x}\to\theta^x \qquad (6.2)$$

である．したがって，

(i) $\theta>0$ のとき，

タイプ A–1

$F(-\theta\zeta,\ -n;\zeta)\to F(-n\,|-\theta)$ $\qquad(\zeta\to\infty)$

タイプ A–2

$F(\theta\zeta,\ -n;-\zeta)\to F(-n\,|-\theta)$ $\qquad(\zeta\to\infty)$

(ii) $1>\theta>0$ のとき，

タイプ A–2

$F(\xi,\ -n;-n/\theta)\to F(\xi\,|\,\theta)$ $\qquad(n\to\infty)$

タイプ B–3

$F(\xi,\ \theta\eta;\eta)\to F(\xi\,|\,\theta)$ $\qquad(\eta\to\infty)$

(II) ポアソン分布による近似

$$\frac{(\alpha)_x(\beta)_x}{(\gamma)_x}=\left(\frac{\alpha\beta}{\gamma}\right)^x\prod_{k=0}^{x-1}\frac{(1+k/\alpha)(1+k/\beta)}{1+k/\gamma} \qquad (6.3)$$

だから $\alpha,\ \beta,\ \gamma\to\pm\infty$，$\alpha\beta/\gamma\to\theta$ のとき

$$\frac{(\alpha)_x(\beta)_x}{(\gamma)_x}\to\theta^x \qquad (6.4)$$

である．したがって，$\theta>0$ にたいして，

タイプ A-1

$$F(-\xi,\ -n;\ n\xi\theta) \to F(1,\ 1\,|\,\theta) \qquad (n,\ \xi \to \infty)$$

タイプ A-2

$$F(\xi,\ -n;\ -n\xi\theta) \to F(1,\ 1\,|\,\theta) \qquad (n,\ \xi \to \infty)$$

タイプ B-1

$$F(-n+\varepsilon,\ -n+\delta;\ n^2\theta) \to F(1,\ 1\,|\,\theta) \qquad (n \to \infty)$$

タイプ B-3

$$F(\xi,\ \eta;\ \xi\eta\theta) \to F(1,\ 1\,|\,\theta) \qquad (\xi, \eta \to \infty)$$

(III) 打ちきり分布の近似.

(I), (II) で述べた近似は，そのまま条件つき分布について成りたつ．$F(\alpha,\ \beta;\ \gamma)$ が分布 G で近似されるとき，$X\in S$ という条件つきの分布は，G の $X\in S$ という条件つき分布で 近似される．たとえば，

タイプ B-3

$$F(\xi,\ \theta\eta;\ \eta\,\|\,1) \to F(\xi\,|\,\theta\,\|\,1) \qquad (\eta \to \infty)$$

である．この極限分布は，0-打ちきり負の二項分布である．その他，つぎのような近似も考えられる．(cf. §5)

タイプ A-2

$$F(\xi,\ -n;\ -n/\theta\,\|\,1) \to F(0\,|\,\theta\,\|\,1) \qquad (n \to \infty,\ \xi \to 0)$$

タイプ B-3

$$F(\xi,\ \theta\eta;\ \eta\,\|\,1) \to F(0\,|\,\theta\,\|\,1) \qquad (\eta \to \infty,\ \xi \to 0)$$

この極限分布は対数級数分布であり，0-打ちきり負の二項分布の極限としても得られる．

(IV) ガンマ分布による近似.

X が $F(\alpha,\ \beta;\ \gamma)$ に従うとする．

$$\lambda \equiv (\gamma-\beta)/\gamma > 0$$

であるとして，

$$Y = \frac{\gamma-\beta}{\gamma} X = \lambda X$$

とおく．Y は 0 または正の値をとる格子点分布で，その span は λ である．

α は固定し，$\gamma,\ \beta \to \pm\infty$，$\gamma-\beta \to \infty$，$\lambda \to 0$ とする．

$$\frac{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}{\Gamma(\gamma)\,\Gamma(\gamma-\alpha-\beta)} \sim \lambda^{\alpha}$$

(両辺の比が 1 に近ずくという意味)

また，

$$\frac{(\alpha)_x}{x!} \to \frac{1}{\Gamma(\alpha)}\, x^{\alpha-1},$$

$$\frac{(\beta)_x}{(\gamma)_x} = \left[(1-\lambda)^{1/\lambda}\right]^{\lambda x} \prod_{0}^{x-1} \frac{1+k/\beta}{1+k/\gamma}$$

$$\sim e^{-y}$$

ただし $y = \lambda x$

となるから，結局

$$Pr\{Y=y\} \sim \lambda \cdot \frac{1}{\Gamma(\alpha)}\, y^{\alpha-1} e^{-y}$$

そして，任意の $y>0$ にたいして，

$$\int_0^y \frac{1}{\Gamma(\alpha)} u^{\alpha-1} e^{-u} du = \lim_{\substack{p\to\infty \\ \lambda\to 0}} \frac{1}{\Gamma(\alpha)} \sum_{k=1}^{p} (\lambda k)^{\alpha-1} \cdot e^{-\lambda k} \cdot \lambda$$

$$\sim Pr\{Y \leqq y\}$$

とくに,

タイプ A-2 $F(\xi, -n; -n-\zeta)$ で,

$$Y = \frac{\zeta}{n+\zeta} X \quad \text{が,}$$

タイプ B-3 $F(\xi, \eta; \eta+\zeta)$ で,

$$Y = \frac{\zeta}{\eta+\zeta} X \quad \text{が,}$$

それぞれ, $n\to\infty$, $\eta\to\infty$ のときガンマ分布で近似される.

(V) 正規分布による近似

(i) タイプ A-1, A-2, B-1 および B-3 は (II) で述べたようにポアソン分布で近似されるから,

$$Y = \frac{X - \alpha\beta/\gamma}{\sqrt{\alpha\beta/\gamma}},$$

あるいは,

$$Y = \frac{\gamma X - \alpha\beta}{\sqrt{\alpha\beta\gamma}}$$

が, $\alpha, \beta, \gamma \to \pm\infty$, $\alpha\beta/\gamma \to \infty$ のとき, 正規分布 $N(0,1)$ で近似される.

この節については, 〔4〕, 〔5〕, 〔6〕 および 〔19〕 などを参照.

§7. $\pm X \pm p$ の分布

すでに §1 で述べたように, K. Sarkadi 〔21〕 は, $a \geqq n > b > 0$ がすべて整数のとき,

$$\binom{a}{x}\binom{b}{n-x} \Big/ \binom{a+b}{n}$$

は $n-b, n-b+1, \cdots, n$ 上の分布になるが, これも一般超幾何分布に入れるべきであると主張した. 実際, この分布は, つぼのモデルで説明されるものであり, さらにこれを $n-b$ だけに左にずらした変数 Y は,

$$\frac{\binom{b}{y}\binom{a}{a+b-n-y}}{\binom{a+b}{a+b-n}} \qquad y = 0, 1, \cdots, b$$

すなわちタイプ A-1 $F(-b, -a-b+n; n-b+1)$ に従うものになる. 同じようなものは, タイプ B-3 からも作れるが, K. Sarkadi はこういうものまで含めてこそ, 超幾何から一般超幾何への拡張が完成するのだと主張しているのであるが, 事情はそれ程簡単ではない. 例を上げよう. これらは, いずれも K. Sarkadi のいうようなものではないのである. (なお, §5. の最後にある例をみよ).

(例 1)

$$\binom{-1}{x}\binom{-4}{3-x} \Big/ \binom{-5}{3}$$

は $x=0, 1, 2, 3$ の上の分布 (タイプ A-2 $F(1, -3; -6)$) であるが, 同時に $x=7, 8, 9, 10$ の上の分布でもある.

(例2)

$$\binom{\frac{3}{2}}{x}\binom{\frac{1}{2}}{\frac{3}{2}-x}\Big/\binom{2}{\frac{2}{3}}$$

は $x=1,2,3,\cdots\cdots$ 上の分布である.

(例3)

$$\binom{-1}{x}\binom{-\xi}{-n-\xi-1-x}\Big/\binom{-\xi-1}{-\xi-1-n}$$

は $x=-n-1,\ -n,\cdots\cdots,-1$ 上の分布である.

(例4)

$$\binom{-1}{x}\binom{1}{-1-x}\Big/\binom{0}{-1}$$

は $x=0,1,\cdots\cdots$ 上の分布 (タイプ B-3 $F(1,\ 1;3)$) であるが, 同時に $x=-3,\ -4,\cdots\cdots$ 上の分布でもある.

これらの分布を記述するために, (2.6) の表現を拡張しておこう. 実数 α と, 正の整数 x にたいして,

$$(\alpha)_x=\Gamma(x+\alpha)/\Gamma(\alpha) \tag{7.1}$$

であるから, (2.6) は,

$$Pr\{X=x\}=\frac{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta)}{\Gamma(\gamma-\alpha-\beta)\,\Gamma(\alpha)\,\Gamma(\beta)}\cdot\frac{\Gamma(x+\alpha)\,\Gamma(x+\beta)}{\Gamma(x+1)\,\Gamma(x+\gamma)} \tag{7.2}$$

となる. これは, たとえば $\gamma=0$ でも, また $x<0$ でも意味をもちうるから, (2.6) を離れて, (7.2) によって, これまでより広い分布族を定義することができる. そして, 例1や4からも分るように, 母数 $\alpha,\ \beta,\ \gamma$ の値の他, 分布範囲 A を指定して分布が確定することになる. A はもちろん整数の集合としてよいが, これはかなり複雑なものも作れる. たとえば N を非負整数からなる任意の集合とし, $M=\{y;\ -3-y\notin N,\ y$ は -3 より大きくない整数$\}$ とおくと, (7.2) で, $\alpha=1,\ \beta=1,\ \gamma=3$ としたものは, $A=M\cup N$ 上の分布になる (cf. 例4). しかし, こういうものを記述してみても, あまり意味はないと思うので, X が $F(\alpha,\beta;\gamma)$ に従う変数, p が正の整数のとき, $\pm X\pm p$ という変数で, しかも, (7.2) の形の分布をもつもののうち, 興味あるものについて述べよう. これらを一般超幾何分布と呼ぶべきかどうかは別問題である (cf. §5.).

(I) 正の整数 p にたいして,

$$\frac{\Gamma(1-\alpha)\,\Gamma(1-\beta)}{\Gamma(\alpha-p)\,\Gamma(\beta-p)}=\frac{\Gamma(1-\alpha+p)\,\Gamma(1-\beta+p)}{\Gamma(\alpha)\,\Gamma(\beta)} \tag{7.3}$$

に注意すると, X が $F(\alpha,\ \beta;\ 1+p)$ に従うとき, $Y=X+p$ (X を p だけ右にずらす) の分布は,

$$Pr\{Y=y\}=\frac{\Gamma(1-\alpha)\,\Gamma(1-\beta)}{\Gamma(\alpha-p)\,\Gamma(\beta-p)\cdot\Gamma(1+p-\alpha-\beta)}\times\frac{\Gamma(y+\alpha-p)\,\Gamma(y+\beta-p)}{\Gamma(y+1)\,\Gamma(y+1-p)} \tag{7.4}$$

$$y=p,\ p+1,\cdots\cdots$$

である. これは, (7.2) の形であり, $F(\alpha-p,\ \beta-p;\ 1-p)$ と書かれるべきものである. 同様に, X が, $F(\alpha,\ \beta;\ 1-p)$ に従うとき, $Y=X-p$ (X を p だけ左にずらす) の分布は, $F(\alpha+p,\ \beta+p;\ 1+p)$ と書かれるべきものになる. 具体的には,

タイプ A-1　$F(-\xi,\ -n;\ 1+p)$ を p だけ右にずらして，

$$F(-\xi-p,\ -n-p;\ 1-p)$$

タイプ A-2　$F(\xi,\ -n;\ 1-p)$，$p>n$ を p だけ左にずらして，

$$F(\xi+p,\ -n+p;\ 1+p)\qquad \text{(cf. 例3)}$$

タイプ B-1　$F(-n+\varepsilon,\ -n+\delta;\ 1+p)$ を p だけ右にずらして，

$$F(-n-p+\varepsilon,\ -n-p+\delta;\ 1-p)\qquad \text{(cf. 例2)}$$

タイプ B-3　$F(\xi,\ \eta;\ 1+p)$ を p だけ右にずらして，

$$F(\xi-p,\ \eta-p;\ 1-p)$$

等々が得られる．分布範囲 A が何であるかは明らかであろう．K. Sarkadi が指摘しているのは，これらのうちタイプ A-1, B-3 である．

とくに $p=1$ として得られる，タイプ A-1 $F(-\xi-1,\ -n-1;\ 0)$ は $F(-\xi-1,\ -n-1;\ \gamma)$ から，またタイプ B-2 $F(-n-1+\varepsilon,\ -n-1+\delta;\ 0)$ は $F(-n-1+\varepsilon,\ -n-1+\delta;\ \gamma)$ から，$\gamma\to 0$ として得られる．また，タイプ B-3 についていえば，$\xi>1>\delta>0$, $2>\xi+\delta$ のとき，$F(\xi-1,\ \delta-1;\ 0)$ はタイプ B-2 $F(\xi-1,\ \delta-1;\ \gamma)$ から，$1>\varepsilon$, $\delta>0$ のとき $F(\varepsilon-1,\ \delta-1;\ 0)$ はタイプ B-1 $F(\varepsilon-1,\ \delta-1;\ \gamma)$ から $\gamma\to 0$ としてそれぞれ得られる．

(II)　タイプ A-2, B-3 では，$\alpha=1$ にとれるが，この場合は興味がある．(cf. §5)
p を正の整数とする．

タイプ A-2　$F(1,\ -n;\ -n-2p)$ は，

$$Pr\{X=x\}=\frac{\Gamma(-2p)\,\Gamma(-n-2p-1)}{\Gamma(-2p-1)\,\Gamma(-n)}\cdot\frac{\Gamma(x-n)}{\Gamma(x-n-2p)}$$

$$=\frac{2p+1}{(n+1)_{2p}}\cdot(x-n-1)\cdots\cdots(x-n-2p)\tag{7.5}$$

$$x=0,1,\cdots\cdots,n$$

u の多項式 $f(u)=(u-n-1)\cdots\cdots(u-n-2p)$ が，$u=n+p+1/2$ に関して対称であることに注意して，

$$Y=2n+2p+1-X\tag{7.6}$$

とおくと，

$$Pr\{Y=y\}=Pr\{X=2n+2p+1-y\}$$

$$=\frac{2p+1}{(n+1)_{2p}}\,(y-n-1)\cdots\cdots(y-n-2p)$$

$$=\frac{\Gamma(-2p)\,\Gamma(-n-2p-1)}{\Gamma(-2p-1)\,\Gamma(-n)}\cdot\frac{\Gamma(y-n)}{\Gamma(y-n-2p)}\tag{7.7}$$

$$y=n+2p+1,\cdots\cdots,2n+2p+1$$

(7.5) と (7.7) を比較せよ．この関係は，(7.5) を (2.7) の形に表現した場合にも成りたつと思うかも知れないが，実はそうでない，(7.5) は，

$$Pr\{X=x\}=\frac{\binom{-1}{x}\binom{-2p-1}{n-x}}{\binom{-2p-2}{n}}\tag{7.8}$$

$$x=0,1,\cdots\cdots,n$$

そして，Y の分布は，

$$Pr\{Y=y\}=Pr\{X=2n+2p+1-y\}$$
$$=-\frac{\binom{-1}{y}\binom{-2p-1}{n-y}}{\binom{-2p-2}{n}} \tag{7.9}$$
$$y=n+2p+1,\cdots\cdot,\ 2n+2p+1.$$

である．(k が非負整数のとき，$\binom{-1}{k}=(-1)^k$ であること，および，x が偶 (or 奇) なら y が奇 (or 偶) になることに注意)．$F(1,\ -n;\ -n-2p)$ の代りに，$F(1,\ -n;\ -n-2p-1)$ を考え，(7.6) の代りに，

$$Y=2n+2n+2-X \tag{7.10}$$

とすれば，話は逆になる (cf. 例1)．つぎにタイプ B-3 $F(1,\ n;\ n+p)$ では，

$$Pr\{X=x\}=\frac{\Gamma(n+p-1)\ \Gamma(p)}{\Gamma(p-1)\ \Gamma(n)}\cdot\frac{\Gamma(x+n)}{\Gamma(x+n+p)}$$
$$=(p-1)\cdot(n)_{p-1}\cdot\frac{1}{(x+n)\cdots\cdot(x+n+p-1)} \tag{7.11}$$
$$x=0,1,2,\cdots\cdot$$

である．

$$Y=-2n-p+1-X \tag{7.12}$$

の分布は

$$Pr\{Y=y\}=(-1)^p\cdot(p-1)\cdot(n)_{p-1}\frac{1}{(y+n)\cdots\cdot(y+n+p-1)}$$
$$=(-1)^p\frac{\Gamma(n+p-1)\ \Gamma(p)}{\Gamma(p-1)\ \Gamma(n)}\cdot\frac{\Gamma(y+n)}{\Gamma(y+n+p)} \tag{7.13}$$
$$y=-2n-p+1,\ -2n-p,\cdots\cdot$$

となる．(7.11) と比較せよ．この場合は (2.7) の形で表現しても同じ結果をうる．すなわち，どちらで表現しておいても，p が偶数のときは，Y の分布は X の分布と同じ形の確率関数をもつが，p が奇数のときは符号だけが変わる．

おわりにいろいろと助言を与えて下さった，統計数理研究所の渋谷政昭，高橋宏一両氏に心からお礼いたします．

統計数理研究所

参 考 文 献

[1] A. A. K. Ayyangar, "A note on the incomplete moments of the hypergeometrical series", Biometrika, 26(1934), 264-265.

[2] N. J. Baily, "On estimating the size of mobile populations from recapture data", Biometrika, 38(1951), 293-306.

[3] W. N. Bailey, Generalized hypergeometric series, Cambridge tracts in mathematics and mathematical physics, no. 32, Cambridge University Press, (1935).

[4] B. H. Camp, "Probability integrals for a hypergeometric series", Biometrika, 17(1925), 61-67.

[5] O. L. Davies, "On asymptotic formulae for the hypergeometric series", Biometrika, 26 (1934), 59-107 and 295-322.

[6] Z. Govindarajulu, "Normal approximations to the cassical discrete distributions", G. P. Patil editor, Classical and contagious discrete distributions.

[7] M. Greenwood, Junr., "On errors of random sampling in certain cases not suitable for the application of a 'normal 'curve' of frequency", Biometrika, 9(1913), 69-90.

[8] E. J. Gumbel and H. von Schelling, "The distribution of the number of exceedances", Ann. Math. Statisti., 21(1950), 247-262.

[9] W. Hopkins, "An instance of negative hypergeometric sampling in practice", Bull. de l'Inst. International Statisti., 34(1953), 298-306,

[10] J. O. Irwin, "The generalized Waring distribution applied to accident theory", Jour. Roy. Statisti. Society, 131(1968), 205-225.

[11] G. Ishii and R. Hayakawa, "On the compound binomial distribution", Ann. Inst. Statisti. Math., 12(1960), 69-80.

[12] J. A. Keats and F. M. Lord, "A theoretical distribution for mental test scores", Psychometrika, 27(1962), 59-72.

[13] C. D. Kemp and A. W. Kemp, "Generalized hypergeometric distribution", Jour. Roy. Statisti. Society, 18(1956), 202-211.

[14] W. Feller, "An introduction to probability theory and its applications, I", John Wiley, 1957.

[15] W. Feller, "An introduction to probability theory and its applications, II", John Wiley, 1966.

[16] K. Pearson, "On the moments of the hypergeometrical series," Biometrika, 16(1924), 157-162.

[17] K. Pearson, "On the curves which are most suitable for describing the frequency of random samples of a population", 5(1906), 172-176.

[18] K. Pearson, "On a certain double hypergeometrical series and its representation by continuous frequency surfaces", Biometrika, 16(1924), 272-188.

[19] P. V. Romanovsky, "On the moments of the hypergeometrical series", Biometrika, 17 (1925), 57-60.

[20] K. Sarkadi, "On the distribution of the number of exceedances", Ann. Math. Statisti., 28(1957), 1021-1023.

[21] K. Sarkadi, "Generalized hypergeometric distributions", Publ. Math. Inst. of Hungarian Academy of Science, II (1957), 59-69.

[22] 渋谷政昭編，負の二項分布，日本規格協会 COSCO 資料，(1963).

[23] J. K. Skellam, "A Probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials", Jour. Roy. Statisti. Society, (1948), 257-261.

[24] M. Sibuya, I. Yoshimura, R. Shimizu; "Negative Multinomial Distribution", Ann. Inst. Statisti. Math., 16 (1964), 409-426.

[25] 志村利雄・髙橋宏一，"Galton-Watson 過程の消滅までの世代数分布のモーメントと分布例"，統計数理研究所彙報，14(1966), 161-166.

[26] J. V. Uspensky, Introduction to mathematical probability, McGraw-Hill, 1937.