II－63　渇水時流況表現のための二変数二項分布の母数推定

名古屋工業大学　正会員　長尾正志、岐阜工業高等専門学校　正会員　鈴木正人
学生員　小西宏和、飛島建設　伊藤淳

１．　研究の目的と概要

利水用貯水池の機能評価や合理的な操作法を研究する場合、その入力である渇水時の流況を的確に表現することが基本である。また、計算の便宜上その分布は上限のある離散分布であれば、好都合である。著者らは、渇水時の時系列の強い持続性と周辺分布における正の歪が導入できるモデルとして二変数二項分布の採用を試みてきた。ここでは、二変数二項分布に従う乱数を発生させた数値実験を介して、積率解、最尤解の比較・検討の形で、その母数推定を行なう。

２．　母数推定法

２．１　理論分布　基礎とした分布は、Edwards & Gurland の提案した分布を用いる。この周辺分布、条件付き分布は、上限ｒ、形状母数ａ、１次の自己相関係数ρを使って（１）、（２）式で表わされる。

$$P(i) \equiv P_r[Q_t = i] = {}_iC_r \cdot (1-a)^{r-i} a^i \quad (i=1,2,...,r) \qquad (1)$$

$$g_{ij} \equiv P_r[Q_t = j \mid Q_{t-1} = i] = \sum_{s=0}^{\min(i,j)} {}_sC_i \cdot {}_{j-s}C_{r-i} \cdot \{a(1-\rho)+\rho\}^s a^{j-s} \times (1-a)^{i-s} (1-\rho)^{i+j-2s} \{1-a(1-\rho)\}^{s+r-i-j} \qquad (2)$$

２．２　積率解　総数ｎ個の資料$x_1, x_2, \cdot\cdot, x_n$に対して、３個の未知母数$\tilde{r}, \tilde{a}, \tilde{\rho}$は以下のように導かれる。

$$\tilde{r} = (\bar{x})^2 / (\bar{x} - s^2), \quad \tilde{a} = 1 - s^2 / \bar{x}, \quad (\bar{x}、s：標本の平均、標準偏差) \qquad (3)$$

$$\tilde{\rho} = [\{\sum_{i=1}^{n-1} x_i \cdot x_{i+1} / (n-1)\} - \bar{x}1 \cdot \bar{x}2] / (s1 \cdot s2) \qquad (4)$$

なお、上式中の$\bar{x}1$、$\bar{x}2$およびｓ１、ｓ２は、資料$\{x_i\}$の最後、最初を除いたｎ－１個の標本の平均および標準偏差である。

２．３　最尤解　通常は、３個の未知母数をパラメータとして、対数尤度を最大にするものを選ぶが、ここでは以下の手法をとる。予備計算では、母数のうちで相関母数は比較的精度良く求められるのに反して上限母数はかなり安定性が良くないことが分かっている。そこで、まず周辺分布でｒとａを得たうえで、相関母数ρを求めるという２段階をへた推定を行なう。

（１）　周辺分布の母数推定　（１）式より対数尤度$\log L = \log \prod_{i=1}^{n} P(x_i)$を求め、、まず未知母数ａで微分して０とおくと次式となる。　$r \cdot a = \bar{x}$　（５）　同様にｒに関する式に（５）式を代入してｒのみの関数として（６）式を得る。さらに、これは資料を小さい方から大きい方へ並べ直した順序統計量

$$\sum_{i=1}^{n} \sum_{k=1}^{xi} (r+1-k)^{-1} + n \cdot \log(1 - \bar{x}/r) = 0 \qquad (6)$$

X(j)　(j=1,2,･･･,n) の超過確率をW(j)、また資料の最大値をＲとすると、次式の表現ができる。

$$Er(r) \equiv \sum_{j=0}^{R-1} [W(j) / (r-j)] + \log(1 - \bar{x}/r) = 0 \qquad (7)$$

計算に際しては、資料上限を出発点としてｒを増しながらEr(r)を求め、その絶対値が最小となる整数ｒを解とした。ついで、ａは（５）式より、$a = \bar{x}/r$で求める。

（２）　相関母数の推定　（１）、（２）式より同時分布の対数尤度$LL = LL(\rho)$は次式となる。

$$LL = \sum_{i=1}^{n-1} [\ln xiCr + xi\ln a + (r-xi)\ln a + \ln \sum_{s=0}^{\min(xi,xi+1)} sCx_i \cdot x_{i+1}{-s}Cr{-xi} \cdot \{a(1-\rho)+\rho\}^s$$

$$\times \{1-a(1-\rho)\}^{s+r-x_i-x_{i+1}}\ a^{x_{i+1}-s}\ (1-a)^{x_i-s}\ (1-\rho)^{x_i+x_{i+1}-2s}\ ] \qquad (8)$$

ここで、既知であるａとｒを上式に用い、未知数ρのみの関数と考え、これを最大とするρを解とする。

３． 推定結果とその考察

３．１ 一様乱数発生と検定　発生させた[0 1) 区間の一様乱数に対して１００個を単位として、χ^2検定（８分割で5%水準）と自己相関係数（ρ＝０の両側5%水準）で検定し、およそ7.7%を棄却した。

３．２ 二変数二項分布乱数の発生と検定　検定済一様乱数から二変数二項分布乱数を発生させた。

（１）発生条件　平均を一定1.2 として、（ｒ＝3 ，ａ=0.4) (4 ，0.3) (6 ，0.2)　また相関母数はρ=0.2 ，0.4 ，0.6 ，0.8 とした。１００個を単位とした１０００組について、形状母数、相関母数の違いと推定法の関係などを調べた。

（２）棄却　積率解では、上限母数ｒに非常に大きい正数または小さい負数が出現することがまれにある。また、最尤解では上限母数の決定で、ｒの増加に伴なって対数尤度が増加を続ける場合がある。いずれもこれらを含めると結果が極めて不明確になるので、これらを除外することにした。すなわち最尤解で結果のでないものを含めてどの手法でも、ｒが０以下および仮定したｒの２倍を越すものは全て除いている。

除外された数は相関母数によって異なり、相関が大きくなると増加する。たとえば、ｒ＝4,　ａ＝0.3, ρ＝0.4 とρ＝0.6 ではそれぞれ6%, 12%であった。

３．３ 推定結果の考察　100 個単位で各400 組に対する積率解と最尤解の比較として、表－１に相関母数ρ＝0.4 、表－２にρ＝0.6 の例を示す。

表－１　母数推定における積率解と最尤解の比較
（ｒ=4, ａ=0. 3, ρ=0. 4）

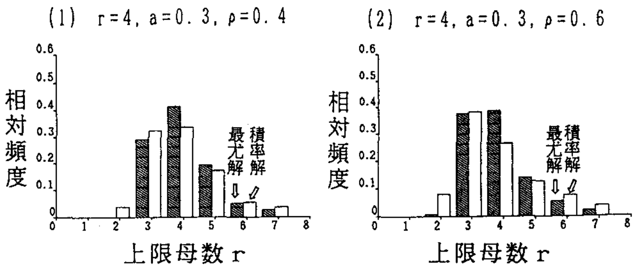
手法	積率解			最尤解		
項目	上限母数 r	形状母数 a	相関母数 ρ	上限母数 r	形状母数 a	相関母数 ρ
平　均	4.1200	0.3205	0.3783	4.1725	0.3069	0.3839
変動係数	0.3184	0.3102	0.2502	0.2661	0.2483	0.2294
歪 係数	1.0806	0.4742	-0.0509	1.2453	-0.0233	-0.0157

表－２　母数推定における積率解と最尤解の比較
（ｒ=4, ａ=0. 3, ρ=0. 6）

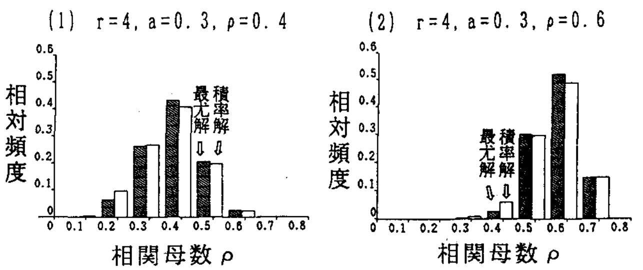
手法	積率解			最尤解		
項目	上限母数 r	形状母数 a	相関母数 ρ	上限母数 r	形状母数 a	相関母数 ρ
平　均	3.9725	0.3378	0.5718	3.9650	0.3226	0.5803
変動係数	0.3521	0.3339	0.1364	0.2680	0.2550	0.1158
歪 係数	1.0299	0.4172	-0.4412	1.2579	0.0072	-0.4402

まず、手法の比較では積率解と最尤解では、推定された上限母数の平均はほとんど変わらないが、変動係数では最尤解の方がかなり小さく解が安定する。また形状母数と相関母数では最尤解の方が真値に近い。さらに、仮定する相関母数が大きいと一般に変動が増し，推定結果が不安定となる。

この様子を、横軸に上限母数、相関母数縦軸に相対頻度をとって、図－１、図－２に示す。ここでも最尤解の方が解の真値の近傍に推定値が集中し、推定結果の信頼性が高いことが分かる。なお、どの方法でも一般に相関母数の推定結果はあまり大きな変動は示さないが、上限母数の変動は大きく、この推定には十分な注意が必要であると判断される。なお、どの場合にも標本単位数や標本数が増せば推定結果は改善されていく。



図－１　上限母数ｒの推定結果



図－２　相関母数ρの推定結果