形状母数の異なる2変数二項分布の理論モデル

名古屋工業大学　○正会員　長尾正志、学生員　小西宏和

1．　はしがき

貯水池の機能評価や操作法の研究をする場合、その入力である河川流量の確率分布を理論モデルで表わす必要がよくある。その際に行列演算などの要求から流量の上限が有限であり、流況に応じて分布形状が変化させうるモデルとして、著者らは2変数の二項分布を採用してきた。しかし、そのモデルは1地点の流量を継続して考える時のように、形状母数の共通するものであった。ここでは、隣接した2流域間の雨量資料を問題にする場合のように、上限が共通で形状母数が異なり、その間に相互相関のある2変数二項分布が解析的に誘導できることを見出したので、その結果について報告する。

2．　確率母関数

Edwards & Gurland の2変数負の二項分布と同様に、上限が共通 r で形状母数の異なる2変数二項分布に対する変量 X_1 、X_2 の確率母関数（p.g.f.）を次式で与え、これを解析の出発点とする。

$$G_{x1,x2}(z_1, z_2) = [A + B_1 z_1 + B_2 z_2 + C z_1 z_2]^r \tag{1}$$

ただし、A、B_1 、B_2 、Cはある定数で、$A + B_1 + B_2 + C = 1$　　(2)

である。これより、X_1 についての周辺分布の確率母関数は、次式のようになる。

$$G_{x1}(z_1) = [G_{x1,x2}(z_1, z_2)]_{z2=1} = [(A + B_2) + (B_1 + C) z_1]^r \tag{3}$$

また、$X_2 = s$ を与えたときの X_1 の条件付き確率母関数は次式で求められる。

$$G_{x1|x2=s}(z_1) = [(\partial^s / \partial z_2{}^s) G_{x1,x2}(z_1, z_2) / (\partial^s / \partial z_2{}^s) G_{x2}(z_2)]_{z2=0}$$
$$= \{(A + B_1 z_1) / (A + B_1)\}^{r-s} \cdot \{(B_2 + C z_1) / (B_2 + C)\}^s \tag{4}$$

3．　周辺分布とその母数

X_1 の周辺分布は、$P_{x1}(i) = [\partial^i G_{x1}(z_1) / \partial z_1{}^i]_{z1=0} / i!$ で求められ、次式となる。

$$P_{x1}(X_1) = {}_rC_{x1} \cdot (A + B_2)^{r-x1} (B_1 + C)^{x1} \quad (X_1 = 0, 1, \cdot\cdot, r) \tag{5}$$

これより、その平均、分散はつぎのように求まる。

$$E(X_1) = [G'_{x1}(z_1)]_{z1=1} = r\{A + B_1 + B_2 + C\}^{r-1}(B_1 + C) = r(B_1 + C) \tag{6}$$

$$V(X_1) = [G''_{x1}(z_1) + G'_{x1}(z_1)]_{z1=1} - \{E(X_1)\}^2$$
$$= r(B_1 + C)(A + B_2) \tag{7}$$

同様に歪係数が3次の積率より誘導される。

$$C_s(X_1) = \{1 - 2(B_1 + C)\} / \{r(B_1 + C)(A + B_2)\}^{1/2} \tag{8}$$

4．　相互相関係数

さらに、X_1 と X_2 との間の相互相関係数を求めた結果はつぎのとおりである。

$$\rho = \{C - (B_1 + C)(B_2 + C)\} / \{(B_1 + C)(A + B_2)(B_2 + C)(A + B_1)\}^{1/2} \tag{9}$$

なお、ρには以下の存在範囲があることが導ける。

$$\max[-\{a_1 a_2 / ((1 - a_1)(1 - a_2))\}^{1/2}, -\{(1 - a_1)(1 - a_2) / (a_1 a_2)\}^{1/2}]$$
$$\leq \rho \leq$$
$$\min[\{a_1 (1 - a_2) / (a_2 (1 - a_1))\}^{1/2}, \{a_2 (1 - a_1) / (a_1 (1 - a_2))\}^{1/2}] \tag{10}$$

さて、以上の諸式を通常の二項分布の表現にするために、形状母数 a_1 を用いて、$B_1 + C = a_1$ 、および $A + B_2 = 1 - a_1$ と表記すると、つぎのようになる。

$$\left.\begin{aligned} &E(X_1) = r a_1 、\quad V(X_1) = r a_1 (1 - a_1) 、\\ &C_s(X_1) = (1 - 2a_1) / \{r a_1 (1 - a_1)\}^{1/2} 、\\ &\rho = (C - a_1 a_2) / \{a_1 a_2 (1 - a_1)(1 - a_2)\}^{1/2} \end{aligned}\right\} \tag{11}$$

5．　条件付き分布

（4）式の条件付き分布の確率母関数をもとに、条件付き分布を誘導する。なお、条件付き分布とその確率母関数の関係は以下のようである。

$$p_{ij} \equiv Pr[X_1 = j \mid X_2 = i] = [(\partial^j / \partial z^j)\, G_{x1=j|x2=i}(z)]_{z=0} / j! \qquad (12)$$

さて、計算の便宜上、$G_{x1=j|x2=i}(z)$ の内容を、zに関係する項と無関係な項に分けて表現する。

$G \equiv \{(A+B_1)^{i-r}(B_2+C)^{-i}\} \cdot H(z)$, $H(z) \equiv (A+B_1 z)^{r-i}(B_2+Cz)^{i}$

算出の手順としては、（12）式の第j次の微分を求めるために、j＝0，1，2，・・・と順に計算を進め、その結果から最後に一般形を推定する。

Ⅰ　j＝0のとき

$$p_{i0} = [G_{x1=0|x2=i}(z)]_{z=0} / 0! = \{B_2/(B_2+C)\}^{i}\{A/(A+B_1)\}^{r-i} \qquad (13)$$

Ⅱ　j＝1のとき

$H'(0) = (r-i)A^{r-i-1}B_1 B_2^{i} + iA^{r-i}B_2^{i-1}C$　より

$$p_{i1} = [(A+B_1)^{i-r}B_2^{i-1}/\{(B_2+C)^{i}A^{i+1-r}\}][(r-i)B_1B_2 + iAC] \qquad (14)$$

Ⅲ　j＝2のとき

$H''(0)/2! = A^{r-i-2}B_2^{i-2}(B_1B_2)^2 \times [(r-i)(r-i-1)/2! +$
$\{2i(r-i)/2!\}\{AC/(B_1B_2)\} + \{i(i-1)/2!\}\{AC/(B_1B_2)\}^2]$

上記右辺の［　］について、一般的な表現を試みる。　　［　］＝

$${}_iC_0 \cdot {}_{r-i}C_2 + {}_iC_1 \cdot {}_{r-i}C_1\{AC/(B_1B_2)\} + {}_iC_2 \cdot {}_{r-i}C_0\{AC/(B_1B_2)\}^2$$
$$= \sum_{s=0}^{\min(i,j)} {}_iC_s \cdot {}_{r-i}C_{j-s} \cdot \{AC/(B_1B_2)\}^{s}$$

この結果をj＝jまで拡張して、p_{ij}の一般形が以下のように得られる。

$$p_{ij} = \frac{(A+B_1)^{i-r}}{(B_2+C)^{i}} A^{r-i-j}B_2^{i-j}(B_1B_2)^{j} \cdot \sum_{s=0}^{\min(i,j)} {}_iC_s \cdot {}_{r-i}C_{j-s}\{AC/(B_1B_2)\}^{s} \qquad (15)$$

6．　条件付き変量の平均、分散

式（4）の条件付き確率母関数より、$X_2=s$を指定した場合の条件付き変量X_1の平均、および分散が得られる。まず、条件付き平均は

$$\mu_1'(X_1 \mid X_2 = s) = rB_1/(A+B_1) + [C/(B_2+C) - \{B_1/(A+B_1)\}]s \qquad (16)$$

であるから、指定変量$X_2=s$に関して線形である。すなわち、回帰曲線は直線回帰である。ついで、条件付き分散は、

$$\mu_2(X_1 \mid X_2 = s) = rAB_1/(A+B_1)^2 + [B_2C/(B_2+C)^2 - \{AB_1/(A+B_1)^2\}]s \qquad (17)$$

したがって、これも指定変数に関して線形的な関係にある。

7．　積率解としての母数推定

（11）式より、まず形状母数a_1、a_2を、$a_i = 1 - V(X_i)/E(X_i)$　　$(i=1,2)$

として推定した後、上限rを、$r = \{E(X_1)+E(X_2)\}/(a_1+a_2)$　(18)

で求めるが、この式の値に最も近い整数として求め直したものを$\hat{r}$とする。そこで、再度形状母数を次式で計算する。$\hat{a}_i = E(X_i)/\hat{r}$　(i=1,2)　(19)，（11）式より、相関係数に標本相関係数を使うと次式でCが分かる。

$$\hat{C} = \hat{a}_1\hat{a}_2 + \hat{\rho}\{\hat{a}_1\hat{a}_2(1-\hat{a}_1)(1-\hat{a}_2)\}^{1/2} \qquad (20)$$

最後に、B_iおよびAは、次式より推定できる。

$$\hat{B}_i = \hat{a}_i - \hat{C},\ (i=1,2) \qquad (21) \qquad \hat{A} = 1 - \hat{B}_1 - \hat{B}_2 - \hat{C} \qquad (22)$$