II-76　情報の与え方と多変数最大エントロピー分布（その2）

信州大学大学院　学生員　福沢直樹
信州大学工学部　正会員　寒川典昭
〃　〃　荒木正夫

１．はじめに

最大エントロピー分布は，情報の与え方により様々な分布形をとる。従来，寒川ら[1]は多変数最大エントロピー分布（ＭＭＥＤ）の基礎特性を研究したが，実データへの適用に際しては，統計モーメントのみを情報として与え，適合度改善のためにパラメータ数を増加させてきた。統計モーメントの情報は，推定分布の形状との密接な関係があることから有効であるが，その後の研究[2],[3]により，多変数の場合も極値水文量に対しては指数関数の期待値も情報として加えると，少ないパラメータ数でデータに対する全体的適合度の改善及び確率水文量の安定性が得られることが予想される。そこで本稿では，統計モーメントと指数関数の期待値を情報とする３変数最大エントロピー分布（３ＭＥＤ）[4]を千曲川流域３地点流量データに適用し，従来の統計モーメントのみを情報とする３ＭＥＤとデータに対する全体的な適合度を数量的に比較検討する。

２．理論式

一般にＭＭＥＤは，確率変数$x_1, x_2, \cdots, x_n$の同時確率密度関数を$p(x_1,x_2,\cdots,x_n)$とすると，次式のように書かれる。

$$p(x_1,x_2,\cdots,x_n) = \exp\{-\lambda_0-\sum_{r=1}^{N}\lambda_r g_r(x_1,x_2,\cdots,x_n)\} \qquad (1)$$

ここに，$g_r(\cdot)$は任意関数で，$\lambda_r (r=1,2,\cdots,N)$はパラメータである。

式（１）は$g_r(\cdot)$の与え方により様々な形状を取る。２次までの統計モーメントのみを情報として，３変数，ｘ，ｙ，ｚの場合に書き換えると，式（１）は次式となる。

$$p(x,y,z) = \exp\{-\alpha-\sum_{a=1}^{2}\beta_a x^a-\sum_{b=1}^{2}\gamma_b y^b-\sum_{c=1}^{2}\delta_c z^c-\eta xy-\mu yz-\rho xz\} \equiv 3M(2,2,2,1,1,1) \qquad (2)$$

ここに，α，β_a，γ_b，δ_c，η，μ，ρは，パラメータである。αは他のパラメータに依存して決定されるため，パラメータ数には含めない。

次に，統計モーメントのみならず，指数関数の期待値も情報とした３ＭＥＤは次式で表される。

$$p(x,y,z) = \exp\{-\alpha-\beta_1 x^a-\beta_2\exp(-bx/M_x)$$
$$-\gamma_1 y^c-\gamma_2\exp(-dy/M_y)-\delta_1 z^e-\delta_2\exp(-fz/M_z)-\eta xy-\mu yz-\rho xz\} \qquad (3)$$

ただし，ａ，ｂ，ｃ，ｄ，ｅ，ｆは最大４の正の整数とし，M_x，M_y，M_zはそれぞれｘ，ｙ，ｚの期待値である。

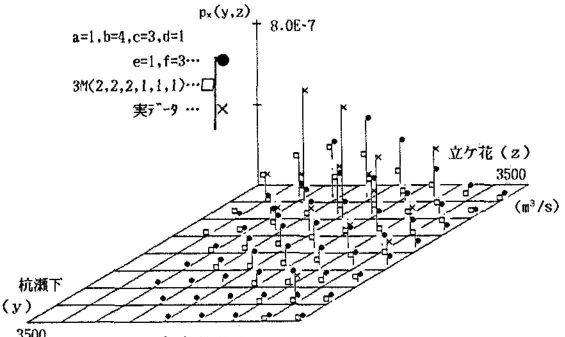
３．実データへの適用と考察

３ＭＥＤの適用問題として，図１に示した支川の千曲川と犀川が合流し，本川千曲川となる合流河川の治水計画を考えた。データは，立ケ花地点３２年間の年最大流量（ｚ）と，それに対応する小市（ｘ），杭瀬下（ｙ）の最大流量である。ここでは，２変数の場合の研究[3]から良い結果が得られている，ａ＝１，ｂ＝４，ｃ＝３，ｄ＝１を固定し，ｅ，ｆを変動させることにより（３）式から１０パターンの最大エントロピー分布を同定した。また得られた分布形の全体的な適合度に対する評価規準として，Ｌ－Ｌ（対数尤度）とＡＩＣを用いた。表１は，今回の上位３パターンと従来の3M(2,2,2,1,1,1)のＬ－Ｌ，ＡＩＣの値である。結果を見ると，用いたデータセットに対しては，指数関数も情報として加えた分布の方が良い適合度を示している。今回の適用では，ｅ＝１，ｆ＝３が最も良い分布となった。図２は，この分布と3M(2,2,2,1,1,1)について実測データとの適合度をメッシュ確率で比較したものである。３者が比較可能な４９メッシュ中，３３

メッシュでｅ＝１，ｆ＝３の分布の方が3M(2,2,2,1,1,1)より実測データに近くなっており，この事からも適合度の改善が読み取れる。

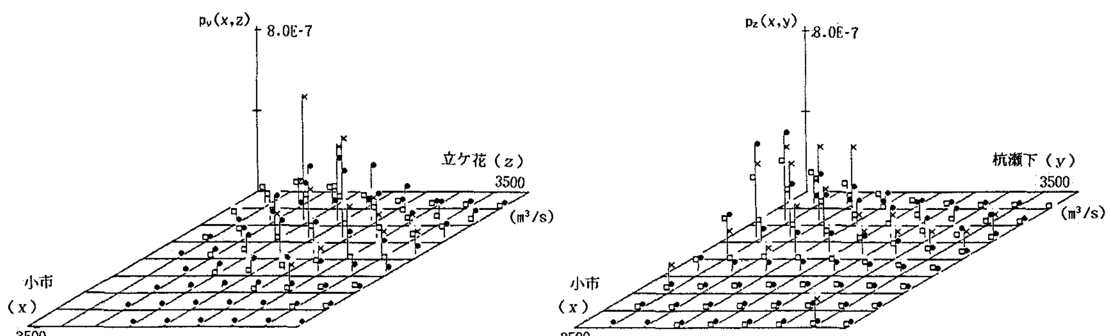
表1　Ｌ－Ｌ，ＡＩＣの値

順位	与えた情報		Ｌ－Ｌ	ＡＩＣ
1	ａ＝１、ｂ＝４，ｃ＝３，ｄ＝１	ｅ＝１，ｆ＝３	-0.7745179D+03	+0.1567036D+04
2		ｅ＝１，ｆ＝４	-0.7748014D+03	+0.1567603D+04
3		ｅ＝１，ｆ＝２	-0.7748352D+03	+0.1567670D+04
4	3M（２，２，２，１，１，１）		-0.1120627D+04	+0.2259255D+04



（a）杭瀬下－立ケ花（y－z）

図１　流量観測所の位置



（b）小市－立ケ花（x－z）　（c）小市－杭瀬下（x－y）

図２　メッシュ確率図

４．おわりに

得られた結果をみると，極値水文量に対しては多変数の場合も統計モーメントに指数関数の期待値を組み合わせた情報が良さそうである。今後は，ａ，ｂ，ｃ，ｄの固定を外すこと，確率水文量の安定性を評価すること，及び他のデータセットへの適用を検討することを考えている。なお，計算，作図等で，北野建設（株）丸山貴彦君（当時，信州大学工学部学生）の協力を得た．記して謝意を表する．

１）寒川・荒木・佐藤：多変数最大エントロピー分布とその基礎特性に関する研究，土木学会論文集，第375号／Ⅱ-6，1986年11月．

２）寒川・荒木・船橋：情報と母数の数と最大エントロピー分布（その２），第43回年講，Ⅱ-4，1988年10月．

３）清水・寒川・荒木：指数関数を情報に加えた２変数最大エントロピー分布，第43回年講，Ⅱ-3，1988年10月．

４）福沢・寒川・荒木・丸山：情報の与え方と多変数最大エントロピー分布，土木学会中部支部昭和63年度研究発表会，Ⅱ－3，1989年3月．