

情報の与え方と多変数最大エントロピー分布

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

〃 〃 丸山貴彦

1. はじめに

最大エントロピー分布の特徴の一つは、変数の数に制限をもたないことである。従って、寒川ら¹⁾は多変数最大エントロピー分布(MMED)を提案し、その基礎特性を検討してきたが、実データへの適用に関しては、情報として統計モーメントのみに限った。ところが、1変数²⁾、2変数³⁾の場合から類推すると、多変数の場合も指数関数の情報を期待値に加えれば、全体的適合度の改善及び確率水文量の安定性が得られそうである。本稿は、このような立場からのMMEDに対する初めての検討であるため、3変数最大エントロピー分布(3MED)の全体的な適合度についてのみ考察したので、これについて報告する。

2. 理論式

確率変数 $x_1, x_2, \dots, x_n$ の同時確率密度関数を $p(x_1, x_2, \dots, x_n)$ とすると、MMEDは次式で与えられる。

$$p(x_1, x_2, \dots, x_n) = \exp\left\{-\lambda_0 - \sum_{r=1}^N \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} \quad (1)$$

ここに、 $g_r(\cdot)$ は任意関数であり、 $\lambda_r (r=1, 2, \dots, N)$ はパラメタである。 λ_0 は λ_r に依存して決定されるため、パラメタ数には含めない。

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与え、3変数、 x, y, z の場合に変換すると、式(1)は

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

となり、これを $3M(2, 2, 2, 1, 1, 1)$ と略記する。ここに、 $\alpha, \beta_a, \gamma_b, \delta_c, \eta, \mu, \rho$ は、パラメタである。

次に、 $g_r(\cdot)$ に指数関数も情報に加えると3MEDは次式となる。

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

ここに、 M_x, M_y, M_z はそれぞれ x, y, z の平均値である。

3. 実データへの適用

2本の支川が合流し本川となる合流河川の治水計画は、基本的には3変数統計問題であるため、この問題への3MEDの適用を考える。3地点流量データとしては、図1に示した千曲川流域内立ヶ花地点32年間の年最大流量(z)と、それに対応する小市、杭瀬下の最大流量(x, y)を用いた。いま、式(3)の a, b, c, d, e, f を最大4として、従来の研究³⁾から良い結果が得られている $a=1, b=4, c=3, d=1$ を固定してパラメタ同定を試みた結果、10種類の3MEDが得られた。また、推定分布のデータに対する全体的な適合度判定基準として、 $L-L, AIC$ を用いた。

表1は、上位5パターンと、 $3M(2, 2, 2, 1, 1, 1)$ の $L-L, AIC$ の値である。結果をみると、指数関数の期待値を情報に加えるとかなり全体的な適合度の改善が読み取れ、ここでの計算例では、 $e=1, f=3$ の分布が最良と判断された。図2は、この分布の密度関数立体図である。

以上のことから、このように極値水文量に対しては、統計モーメントと指数関数の期待値を組み合わせることが、有効な情報となりそうである。

順位	情報	L-L	AIC
1	$x^1 \exp(-4x/M_x) y^3 \exp(-y/M_y) z^1 \exp(-3z/M_z) xy yz xz$	-0.77451790+03	0.15670360+04
2	$x^1 \exp(-4x/M_x) y^3 \exp(-y/M_y) z^1 \exp(-4z/M_z) xy yz xz$	-0.77480140+03	0.15676030+04
3	$x^1 \exp(-4x/M_x) y^3 \exp(-y/M_y) z^1 \exp(-2z/M_z) xy yz xz$	-0.77483520+03	0.15676700+04
4	$x^1 \exp(-4x/M_x) y^3 \exp(-y/M_y) z^3 \exp(-4z/M_z) xy yz xz$	-0.79460990+03	0.16072220+04
5	$x^1 \exp(-4x/M_x) y^3 \exp(-y/M_y) z^3 \exp(-3z/M_z) xy yz xz$	-0.80048390+03	0.16189680+04
6	$x^1 x^2 y^1 y^2 z^1 z^2 xy yz xz$	-0.11206270+04	0.22592550+04

表1 L-L, AICの値

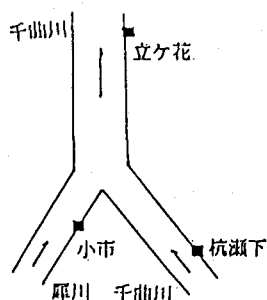
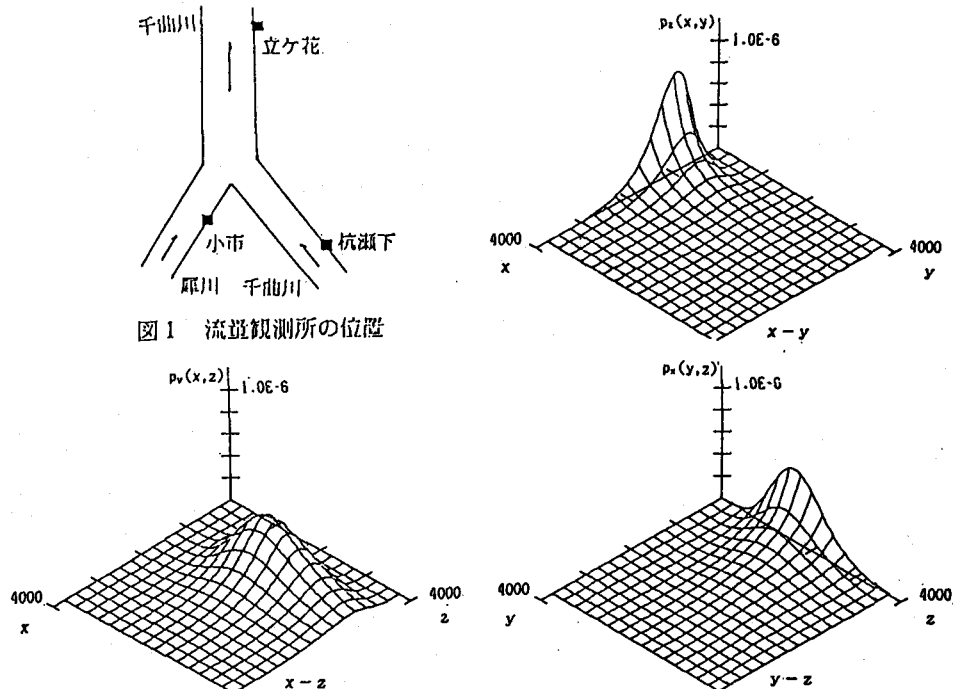


図1 流量観測所の位置

図2 $a=1, b=4, c=3, d=1, e=1, f=3$ の密度関数立体図

4. おわりに

本稿の結果に、1変数, 2変数の場合に得られた結果を考え合わせると、3変数以上の場合にも指数関数の期待値を情報に加える効果は、分布の全体的な適合度の改善という面から大きいと言えよう。恐らくは、確率水文量の安定に対してもよい効果が期待されるが、これについては今後研究していきたい。

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・大原：情報の与え方と最大エントロピー分布（その2），第42回年講，II-3，1987年9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。