

統計モーメントを情報とした 1VNME 分布と頻度分析

1VNME DISTRIBUTION HAVING STATISTICAL MOMENT INFORMATION AND FREQUENCY ANALYSIS

寒川典昭¹・西 知哉²

Noriaki SOGAWA and Tomoya NISHI

¹正会員 工博 信州大学助教授 工学部社会開発工学科 (〒380-8553 長野市若里四丁目 17-1)

²工修 中央開発株式会社 九州事業部設計部 (〒814-0103 福岡市城南区鳥飼六丁目 3-27)

This paper expanded 1 variate maximum entropy distribution into 1 variate nonstationary maximum entropy distribution (1VNME distribution). The given information to the 1VNME distribution is the statistical moments. By using the 1VNME distribution, nonstationary frequency analysis was applied to monthly, seasonal and annual precipitation. The result obtained was as follows. Although fitting rate of the 1VNME distributions were different a little in monthly, seasonal and annual precipitation, the 1VNME distributions were almost fitted to histograms. Therefore this distribution could confirm that it was efficiency. Besides, because the shape of the distribution changed in time series, it was clear that hydrological value had the nonstationarity. Finally, Although the probabilistic value of hydrological variable had the difference in monthly, seasonal and annual precipitation, it showed the secular change. As a whole, the probabilistic value of hydrological variable decreased in the month of winter and increased in the month of summer.

Key Words : 1VNME distribution, statistical moments, nonstationary frequency analysis, probabilistic value of hydrological variable

1. はじめに

治水・利水計画を策定するとき、対象とする母集団の確率分布を決定し、その確率分布が持つパラメータを推定しなければならない。しかしながら、水文量はしばしば小標本であるため、いかなる確率分布を採用すべきかという問題に多くの不確定さが存在する。その問題を解決するために、寒川・荒木¹⁾は、得られる情報は任意関数の期待値の形で取り入れ、それ以外は不確定さが最大となる分布、すなわち最大エントロピー分布を水文量の頻度分布に導入した。

また、一方水文量の時系列を分析すると、そこには非定常性が存在する場合が多い²⁾。水文量に非定常性が存在する場合、その母集団を記述する確率分布のパラメータも時間に依存する。パラメータの時

間的变化は、水文頻度分析における確率水文量を年毎に変え、計画がいつの時点を対象とするかによって、その規模が変化するという事態が発生する。そのため、水災害・水資源の立場から、水文量及び確率水文量の時系列変化を把握し、得られた結果を水工計画に織り込むことが、水文量の頻度分析の分野で重要な課題となっている。

このような状況下で本研究は、1変数最大エントロピー分布を非定常に拡張した、1変数非定常最大エントロピー分布 (1 Variate Nonstationary Maximum Entropy 分布; 以下 1VNME 分布と呼ぶ) を提案するものである。ここでは、1VNME 分布を用いて、長野県内の5つの気象官署である長野、松本、飯田、諏訪、軽井沢の月、季節、年降水量について、非定常頻度分析を実施した。それは1VNME 分布と31個の移動部分標本の適合度をみて、1VNME 分布の有

効性を検討するものである。また、1VNME 分布から非定常な確率水量を算定して、その経年変化についても把握するものである。尚、非定常と言っても31年間は定常という仮定をおいている。

2. 1VNME 分布の理論式の導出

(1) 分布形の導出

時間 t に依存した確率変数を $x(t)$ 、時間 t に依存したその確率密度関数を $p_t(x(t))$ とすると、エントロピー $-H_t$ は次式で表される。

$$H_t = - \int p_t(x(t)) \ln p_t(x(t)) dx(t) \quad (1)$$

一方、確率密度関数が具備すべき条件と、任意関数 $g_r(x(t))$ の期待値は次のように表現される。

$$\int p_t(x(t)) dx(t) = 1 \quad (2)$$

$$\int g_r(x(t)) p_t(x(t)) dx(t) = E[g_r(x(t))] \quad (3)$$

$r = 1, 2, \dots, N$

ここに、 $E[\cdot]$ は期待値記号である。

(2)、(3)式を制約条件として、(1)式を最大にする分布を求める。この分布は水量から得られる情報は $g_r(\cdot)$ の期待値の形で取り入れ、それ以外はエントロピー、すなわち不確定さを最大にする分布である。水量のような不確定さの大きい統計量に、このような分布を用いることは適切である。そこで上述の問題をラグランジュの未定乗数法で解く。まず、ラグランジュ関数は、

$$L = H_t + (\lambda_0(t)) \{1 - \int p_t(x(t)) dx(t)\} + \sum_{r=1}^N \lambda_r(t) \{E[g_r(x(t))] - \int g_r(x(t)) p_t(x(t)) dx(t)\} \quad (4)$$

となる。ここに、 $\lambda_r(t)$ ($r=1, 2, \dots, N$) はラグランジュ乗数である。そこで、(4)式の偏分をとって“0”と置くと、1VNME分布は次式のように求まる。

$$p_t(x(t)) = \exp\{-\lambda_0(t) - \sum_{r=1}^N \lambda_r(t) g_r(x(t))\} \quad (5)$$

(2) パラメータ同定法

a) パラメータ同定法

(5)式のパラメータの同定法を示す。(5)式を(2)式に代入して、 $\lambda_0(t)$ に関して解くと、次式が得られる。

$$\lambda_0(t) = \ln \left\{ \exp \left\{ - \sum_{r=1}^N \lambda_r(t) g_r(x(t)) \right\} dx(t) \right\} \quad (6)$$

(3)式の右辺を $v_r(t)$ と置き、(5)式を(3)式に代入して、

(6)式を使って整理すると次式になる。

$$\begin{aligned} & \int g_r(x(t)) \exp \left\{ - \sum_{r=1}^N \lambda_r(t) g_r(x(t)) \right\} dx(t) \\ &= v_r(t) \int \exp \left\{ - \sum_{r=1}^N \lambda_r(t) g_r(x(t)) \right\} dx(t), \\ & r = 1, 2, \dots, N \end{aligned} \quad (7)$$

(7)式を λ_r に関して解くために、Wragg & Dowson³⁾の方法を採用する。いま、 $\lambda_r(t)$ の近似値を $\alpha_r(t)$ 、残差を $\varepsilon_r(t)$ と置くと、

$$\lambda_r(t) = \alpha_r(t) + \varepsilon_r(t) \quad (8)$$

となる。そこで(8)式を(7)式に代入すると、

$$\begin{aligned} & \int g_r(x(t)) \exp \left\{ - \sum_{r=1}^N (\alpha_r(t) + \varepsilon_r(t)) g_r(x(t)) \right\} dx(t) \\ &= v_r(t) \int \exp \left\{ - \sum_{r=1}^N ((\alpha_r(t) + \varepsilon_r(t)) g_r(x(t))) \right\} dx(t), \\ & r = 1, 2, \dots, N \end{aligned} \quad (9)$$

が得られる。(9)式を変更すると、

$$\begin{aligned} & \int g_r(x(t)) \exp \left\{ - \sum_{r=1}^N \alpha_r(t) g_r(x(t)) \right\} \exp \left\{ - \sum_{r=1}^N \varepsilon_r(t) g_r(x(t)) \right\} dx(t) \\ &= v_r(t) \int \exp \left\{ - \sum_{r=1}^N \alpha_r(t) g_r(x(t)) \right\} \exp \left\{ - \sum_{r=1}^N \varepsilon_r(t) g_r(x(t)) \right\} dx(t), \\ & r = 1, 2, \dots, N \end{aligned} \quad (10)$$

が得られる。ここで $\varepsilon_r(t)$ は微小であるから、 $\varepsilon_r(t)$ に関してTaylor展開し、 $\varepsilon_r(t)^2$ 以上の項を無視すると、(2.10)式より、

$$\begin{aligned} & \int g_r(x(t)) \exp \left\{ - \sum_{r=1}^N \alpha_r(t) g_r(x(t)) \right\} \left\{ 1 - \sum_{r=1}^N \varepsilon_r(t) g_r(x(t)) \right\} dx(t) \\ &= v_r(t) \int \exp \left\{ - \sum_{r=1}^N \alpha_r(t) g_r(x(t)) \right\} \left\{ 1 - \sum_{r=1}^N \varepsilon_r(t) g_r(x(t)) \right\} dx(t), \\ & r = 1, 2, \dots, N \end{aligned} \quad (11)$$

が得られる。従って、(2.11)式を整理すると、次の $\varepsilon_r(t)$ に関する N 元連立1次方程式が得られる。

$$\begin{aligned} & \sum_{j=1}^N (A_{r,j}(t) - v_r(t) A_{0,j}(t)) \varepsilon_j(t) = A_{0,j}(t) - v_r(t) A_{0,0}(t), \\ & r = 1, 2, \dots, N \end{aligned} \quad (12)$$

ここで、

$$A_{r,j}(t) = \int g_r(x(t)) g_j(x(t)) \exp \left\{ - \sum_{r=1}^N \alpha_r(x(t)) g_r(x(t)) \right\} dx(t) \quad (13)$$

$$g_0(x(t)) = 1 \quad (14)$$

としている。

b) 解の存在条件

本研究で扱う1VNME分布は、常に存在するとは限らない。ここでいう1次、2次モーメント $\{\mu_1, \mu_2\}$ に対する1VNME分布が存在する条件は、

$$(\mu_1)^2 < \mu_2 \leq 2(\mu_1)^2 \quad (15)$$

である⁴⁾。また、分散を σ^2 と書くと、(15)式は次のように書き換えられる。

$$0 < \sigma/\mu_1 \leq 1 \quad (16)$$

Sonuga⁵⁾の提案した2次モーメントまでを用いた1VNME分布は、上式の条件下で議論されていることになる。実際の計算の中にもこの条件を取り入れ、適さないデータは、これを排除するものとする。

c) モーメントの基準化

データから求められる統計モーメントは、高次になるほど大きな値となって取り扱いに困難を生じる。そこで、確率変数の平均が1となるような次式の変数変換を行う。

$$u(t) = x(t)/\mu_1(t) \quad (17)$$

すると、 $u(t)$ に関する確率密度関数 $q(u(t))$ は、Jacobian $J(\cdot)$ を介して

$$q_i(u(t)) = J(x(t)/u(t)) p_i(x(t)) \quad (18)$$

となる。ここに

$$J(x(t)/u(t)) = \mu_1(t) \quad (19)$$

であるから、(18)式は、

$$q_i(u(t)) = p_i(x(t)) \mu_1(t) \quad (20)$$

あるいは

$$p_i(x(t)) = q_i(u(t))/\mu_1(t) \quad (21)$$

と表現できる。

したがって、上述の変数変換を行うと、制約条件式は次のように書き換えられる。

$$\int q_i(u(t)) du(t) = 1 \quad (22)$$

$$\int u(t)^r q_i(u(t)) du = \mu_r(t)/(\mu_1(t))^r \equiv m_r(t) \quad (23)$$

$r = 1, 2, \dots, N$

ここに、 $m_0(t) = m_1(t) = 1$ であるため、(23)式で定義された $m_r(t)$ を基準化されたモーメントと呼ぶこととする。 $m_r(t)$ を用いると、 $N=2$ の場合 $m_2(t)$ のみに依存して1VNME分布が定まる。

次に、 $p_i(\cdot)$ と $q_i(\cdot)$ のエントロピーの関係を求める。(17)式、(21)式を(1)式に代入すると、

$$\begin{aligned} H_i &= - \int q_i(u(t)) \ln q_i(u(t)) du(t) + \ln(\mu_1(t)) \\ &\equiv H'_i + \ln(\mu_1(t)) \end{aligned} \quad (24)$$

が得られる。

(24)式はエントロピー H_i が最大るとき、変数変換したエントロピー H'_i が最大になることを示している。実際の適用では $q_i(\cdot)$ を求めておき、しかるのちに(21)式により $p_i(\cdot)$ に変換すればよい。以下の本研究では、この基準化されたモーメントを用いて計算を行った。また、最初から N 次モーメントまでを用いて1VNME分布を求めることを N モーメント適合と呼ぶこととする。

3. 用いたデータ

実データとして、長野県内の気象官署である長野・松本・飯田・諏訪・軽井沢の5地点の1月～12月の月降水量を用いた^{6),7)}。ここで、移動部分標本の長さは31年とした。31年としたのは、確率水文量が安定する⁸⁾という理由からである。ここで、 $t=1$ とは、長野の場合、データの最初の年から31年間(1889～1919年)の中間の年(1904年)、 $t=2$ とは2年目から31年間(1890～1920年)の中間の年(1905年)を意味し、以下同様に、 $t=79$ (1967～1997年の中間の年の1982年)である。

表・1のデータを月(1月～12月)、季節(冬、春、夏、秋)、年について解析する。ここで、冬を前年の12月と当該年の1,2月、春を当該年の3,4,5月、夏を当該年の6,7,8月、秋を当該年の9,10,11月としている。そのため、冬の数1つ少なくなっている。なお、研究を始めた年には、1997年までのデータしか存在しなかったため、データの区間は1997年までとしている。

4. 1VNME分布の算定とヒストグラムへの適合度の検討

(1) 1VNME分布の算定

任意関数 $g_r(x(t))$ として $x(t)^r$ を採用した。これは $r=1, 2, 3, 4$ がそれぞれ平均情報、分散、ひずみ度、とがり度に関連した情報を与えるからである。

$$g_r(x(t)) = x(t)^r, \quad r = 1, 2, \dots, N \quad (25)$$

と置くと、(25)式は

$$\int x(t)^r p_i(x(t)) dx = v_r(t), \quad r = 1, 2, \dots, N \quad (26)$$

となる。これは、降水量データから計算された既知量であり、原点周りの r 次モーメントを表している。

表-1 用いたデータ

気象官署	データの区間	データ数	移動部分標本の組
長野	1889年～1997年	109年分	$t = 1 \sim 79$
松本	1898年～1997年	100年分	$t = 1 \sim 70$
飯田	1898年～1997年	100年分	$t = 1 \sim 70$
諏訪	1945年～1997年	53年分	$t = 1 \sim 23$
軽井沢	1926年～1997年	72年分	$t = 1 \sim 42$

表-2 パラメータの存在表

	長野			松本			飯田			諏訪			軽井沢		
	2M	3M	4M	2M	3M	4M	2M	3M	4M	2M	3M	4M	2M	3M	4M
1月		△													
2月		△													
3月		△					△		△				△		
4月		△		△			△		△				△		
5月		△		△			△		△				△		
6月		△		△										△	
7月		△					△		△						
8月		△													
9月		△		△			△								
10月							△		△						
11月															
12月															
冬	△	×		△			△		△				△		
春	△	×		△	×		△	△	△	×			△	×	
夏	△	△		△	△		△	△	△	△	×	△	△	△	△
秋	△	△		△	△		△	△	△	△	△		△	△	△
年	△	△	×	△	△	×	△	×	×	△	△	×	△	×	×

ただし、 $x(t)$ は降水量なので存在区間は $[0, \infty]$ とする。

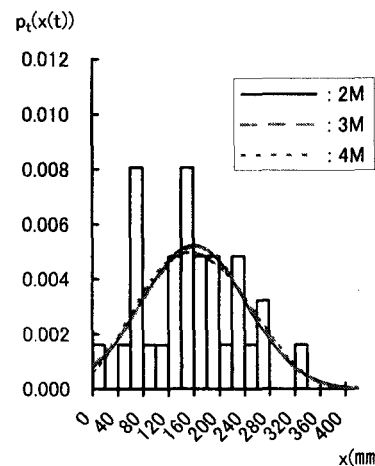
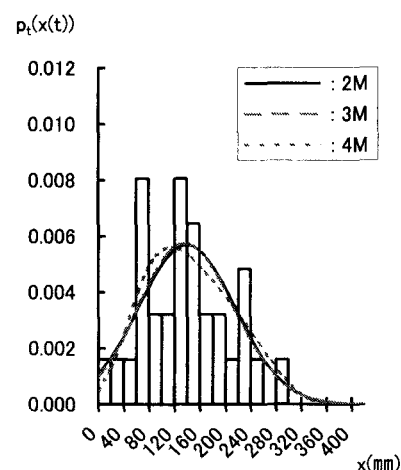
$\nu_r(t)$ を(12)式に入れて、(12)式の連立方程式を $\varepsilon_r(t)$ について解き、(8)式より $\alpha_r(t)$ を更新していき、 $\varepsilon_r(t)$ が十分小さく($\leq 10^{-8}$)なったときに、 $\alpha_r(t)$ を解とみなして、(5)式により1VNME分布が求まる。

この求めた $p(x(t))$ をヒストグラムに重ね合わせて、適合性を判断する。なお、 $\alpha_r(t)$ の初期値は、経験的にその収束性の良さから“0”とした。

(2)パラメータの値

本項では、4次までの統計モーメントが、それを持つ確率分布の形状と密接に関係していることを考えて、 $N=4$ として4モーメント適合までを求めた。表-1で示した実データを用いて、平均、分散に関係した情報を取り入れた2モーメント適合(2M)と、ひずみ度に関係した情報を加えた3モーメント適合(3M)、さらに、尖り度に関係した情報を加えた4モーメント適合(4M)を得た。以下、その結果について考察する。

まず、表-2にパラメータの存在表を示す。表中の×はすべての t についてパラメータが同定できなかった場合、△は一部が同定できなかった場合を示す。これを見ると、季節や年などでは同定できなかった所が多数存在する。分布形が正規分布に近くなると

図-1 ヒストグラムと1VNME分布 ($t=1$)図-2 ヒストグラムと1VNME分布 ($t=20$)

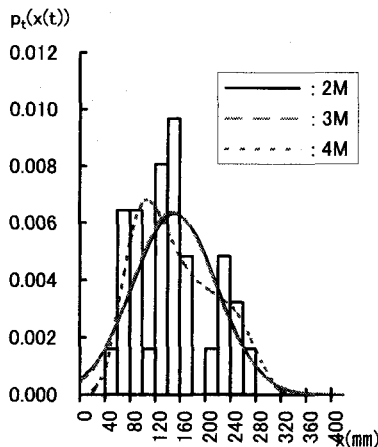


図-3 ヒストグラムと1VNME分布 (t=59)

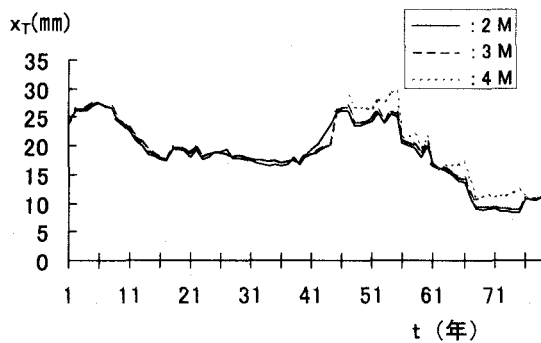


図-4 非超過確率降水量の経年変化 (1月, T=30年)

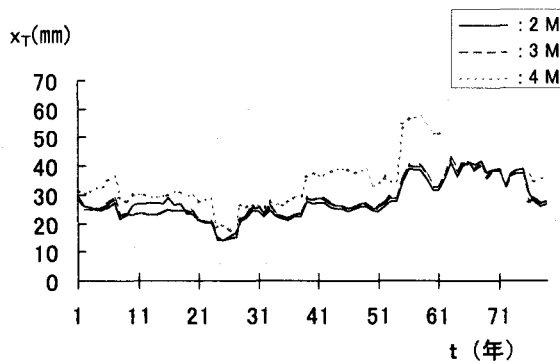


図-5 非超過確率降水量の経年変化 (7月, T=30年)

4モーメント適合が存在しない場合が多い。すなわち2, 3モーメント適合でヒストグラムが十分に説明できるときには, 4モーメント適合が得られなかったということであると言える。

(3) ヒストグラムと推定分布

ここでは, 表-2のパラメータを用いて得られた推定分布をヒストグラムに重ね合わせた結果について示す。ここで, 注意しなければならないのは, 得られた1VNME分布はそのヒストグラムに適合していれば良いというのではなく, 母集団の確率密度関数に適合していなければならないということであるが,

推定分布は水文量の母集団分布を推定するのであるから, 母集団から得られたデータに良く適合する推定分布を優れた分布と判断するのは妥当な立場であると考え, ここでは紙面の都合上洪水の生起する可能性が高い夏期の降水量として7月について長野の場合を考察する。

この場合には, 最初平らだった分布形はだんだん時間と共に左に寄ってきている。つまり, 降水量は減少してきていると言える。そして, 最後の方で尖ってきている。モーメント数による適合では, 2モーメント適合と3モーメント適合はほぼ同じだが, 4モーメント適合は少しずれて, よりヒストグラムを良く説明している。図-1にt=1, 図-2にt=20, 図-3にt=59の場合のヒストグラムと1VNME分布を掲載している。他の地点の1VNME分布も地点によって異なる傾向を示したが, tに依存した非定常性を示していた。

5. 非定常確率水文量の特性

求めた確率水文量の経年変化を見るために, 縦軸をT(5,10,20,30)年非超過確率降水量 $x_T(t)(mm)$, 横軸をt(年)とした時系列グラフを作成した。この場合, 確率降水量が増加すれば洪水の安全度が上がり, 確率降水量が減少すれば, より危険な方向に向かうということである。以下に2例を示す。

長野1月の非超過リターンピリオドT=30年の場合について図-4に掲載している。この場合, 平均するとほぼ20mmである。Tが前半のうちは4モーメント適合が求まらなかったため, 4モーメント適合による確率降水量は分からないが, 2, 3モーメント適合による確率降水量を見てみると, 100年に15mmほど減少している。長野7月の非超過リターンピリオドT=30年の場合について図-5に掲載している。全体的に2, 3モーメント適合よりも4モーメント適合の方が大きくなっている。2, 3, 4モーメント適合はすべて増加傾向であり, 100年で20~30mm程度の増加であった。

表-2に, 長野のT=30年のときの2, 3, 4モーメント適合から得られたそれぞれの非定常確率水文量の回帰直線の傾きを示す。*はモーメント適合が得られなかった場合, 空欄の所は, モーメント適合が得られたが十分な長さが得られなかった場合を意味する。T=5,10,20年のときも含めて, 冬の月で減少傾向, 夏の月などは増加傾向にあった。長野では傾きの変化が小さく, 他の地域よりも安定している。

表-3に, 松本のT=30年のときの2, 3, 4モーメント適合から得られたそれぞれの非定常確率水文量の回帰直線の傾きを示す。*の記号と空欄の意味は表-2に同じである。松本の場合も長野と同様に, 冬の

表-2 非定常確率水文量の回帰直線の傾き(長野,
T=30年, 単位: mm/年)

	2M	3M	4M
1月	-0.1495	-0.1498	
2月	-0.0470	-0.0145	
3月	-0.0370	-0.0364	0.0121
4月	0.1002	0.0335	
5月	0.0752	0.0562	-0.0697
6月	0.3041	0.3098	0.2858
7月	0.1877	0.2048	0.2622
8月	0.1329	0.1380	0.1089
9月	0.0523	0.0265	0.0157
10月	-0.0666	-0.0706	-0.0133
11月	0.0000	0.0092	0.0803
12月	-0.1097	-0.1152	-0.1895
冬	-0.3925	-0.5161	*
春	0.1105	0.0249	*
夏	0.8077	0.2901	
秋	0.0801	0.0602	
年	-0.8949		*

表-3 非定常確率水文量の回帰直線の傾き(松本,
T=30年, 単位: mm/年)

	2M	3M	4M
1月	-0.0188	-0.0060	-0.0474
2月	0.0850	-0.0033	-0.1065
3月	-0.0580	-0.0690	-0.1363
4月	0.0390	-0.1196	-0.4268
5月	-0.1008	-0.0963	-0.0431
6月	0.4761	0.4959	0.6812
7月	0.2889	0.2340	-0.0229
8月	0.0148	-0.0515	-0.1840
9月	0.0010	-0.3190	-0.1141
10月	0.0236	0.0294	0.0318
11月	0.0260	0.0274	0.0140
12月	-0.0507	-0.0394	-0.0398
冬	-0.3065	-0.2415	-0.1182
春	0.3708		*
夏	1.8120	1.0269	
秋	0.4020	0.3417	-0.4180
年	0.0239		*

月で減少傾向, 夏の月で増加傾向にあると言える。

紙面の都合で他の地域の非定常確率水文量の回帰直線の傾きの値は省略するが, 飯田の場合も同様に冬の月で減少傾向, 夏の月で増加傾向という形をとっている。諏訪ではほとんどの場合減少傾向と見られ, 傾きが急になっている。軽井沢ではやはり冬の月では減少傾向, 夏の月では増加傾向となっている。

6. あとがき

得られた結果について要約すると以下のようになる。

- 1) 得られた 1VNME 分布は, ほぼヒストグラムに適合しており, この分布が有効であることを確認でき, また, その形状は時系列的に変化してきていることから, 水文データの非定

常性を確認できた。

- 2) 2, 3 モーメント適合は一致して扁平になることがおおいが, 4 モーメント適合になると適合度は大きく改善されて, 特殊な形状を持つヒストグラムにも, よく適合する。
- 3) 1VNME 分布から非定常な確率水文量を求め, 各地点の確率水文量の経年変化を見た。この変化を延長することで将来の確率水文量の予測などに用いることが期待される。
- 4) 地点や月・季節・年によって多少の差はあるが, 確率水文量は経年変化を示していた。すなわち, 渇水の安全度(確率水文量)が時間(年)と共に変化することを示していた。

次に, 今後解決すべき課題について述べる。

- 1) 1VNME 分布には, データから得られる情報の状態によって, 解が求まらない場合が存在した。このような時の対処について検討する。
- 2) 最適な任意関数 $g_r(\cdot)$ を見つけ, 同様に非定常降水量頻度分析を行なう。

参考文献

- 1) 寒川典昭, 荒木正夫: 水文事象の頻度分析への MEP 導入について, 土木学会論文報告集, 第 335 号, pp.89-95, 1983.
- 2) 寒川典昭, 中村哲, 山田広樹: 長野県における月降水量時系列の経年変化と確率水文量, 第 4 回水資源に関するシンポジウム前刷集, pp.775-780, 1992.
- 3) Wragg, A. and Dowson, D.C.: Fitting continuous probability density function over $[0, \infty)$ using information theory ideas, IEEE Transactions on Information Theory, Vol.IT-16, pp.226-230, 1970.
- 4) Einbu, J.M.: On the existence of a class of maximum entropy probability density function, IEEE Transactions on Information Theory, Vol.IT-23, No.6, pp.772-775, Nov., 1977.
- 5) Sonuga, J.O.: Principle of maximum entropy in hydrologic frequency analysis, Journal of Hydrology, Vol.17, pp.177-191, 1972.
- 6) 長野地方気象台: 信州の気候百年誌, pp.174-182, 1988.
- 7) 財団法人日本気象協会長野センター: 長野県気象月報, 1988~1997.
- 8) 寒川典昭, 荒木正夫, 渡辺輝彦: 確率分布の推定母数の不確定性評価法, 土木学会論文集, 第 375 号/II-6, pp.133-141, 1986.

(2002. 9. 30 受付)