

確率分布の推定母数の不確定性評価法

UNCERTAINTY MEASUREMENT ON POPULATION PARAMETERS OF PROBABILITY DISTRIBUTION

寒川典昭*・荒木正夫**・渡辺輝彦***

By Noriaki SOGAWA, Masao ARAKI and Teruhiko WATANABE

This paper discusses the uncertainty of the estimated population parameters in the hydrologic frequency analysis. Firstly, the changes of the distribution shape and the probable variate of a fixed recurrence interval are investigated using the observed hydrologic data. Next, when the population parameters of the normal distribution are regarded as the random variables, the method of estimation of the decrease degree of the uncertainty, which is measured by the entropy of the distribution in unknown expectation-known variance and unknown expectation-unknown variance, is developed in some detail. Besides, applying this method of estimation to the observed hydrologic data, the validity of this method is investigated. Finally, the possibility of the application of this method to the non-normal population is suggested briefly.

Keywords: uncertainty, population parameters, normal distribution, entropy

1. ま え が き

治水・利水計画を策定するにあたっては、対象となる水文観測資料を用いて頻度解析を行い、計画に必要な確率水流量を算定することが要求される。信頼性の高い水工計画を策定するためには、確率水流量の精度を高める必要があり、従来より水文頻度解析に必要な資料の収集および解析手法の確立に多大の努力が重ねられてきた。その結果、多くの確率分布が水文学に導入され、それらの分布の特徴が解明された。また、パラメーター同定法は、モーメント法、最尤法を基礎として、できるだけ母集団に合う母数の推定法が提示された。しかし、水文資料は小標本である場合が多く、解析に用いられる確率分布の決定に際しても経験的に判断されているにすぎず、求められた確率水流量には多くの不確定さが内在することは否定できない。このような現状から確率水流量の精度をさらに高める努力がなされなければならないことはもちろんであるが、小標本の資料の使用を余儀なくされている以上、どうしてもぬぐいされない不確定性を評価

しておくことはきわめて重要である。

確率水流量の不確定さを考えた場合、大別して2つに分類される。1つは分布形決定に関する不確定さであり、他の1つは分布形が決定された後も存在する推定母数に関する不確定さである。

水文頻度解析には、(i) 正規分布を前提とした分布、(ii) ガンマ分布を前提とした分布、および (iii) 極値分布が採用されている。ところが、対象とする水流量がいかなる母集団に属するかという問題に対しては、単に経験的に、たとえば確率紙上へのプロットから判断されているにすぎず、そこには理論的な根拠、あるいは思想的な妥当性は存在していなかった。

このような問題に対して、高樟・池淵¹⁾は、平均指定の最大エントロピー分布は指数分布、分散指定の場合は正規分布になることに着目し、時間単位別の降水量が平均と標準偏差のどちらが安定しているかを検討して、従う分布形の理論的な説明を試みた。Sonuga^{2),3)}は、標本から得られる情報は2次までのモーメントの形で取り入れ、それ以外はエントロピーが最大となる分布を求めた。われわれ^{4),5)}は、情報理論の分野で Wragg & Dowson⁶⁾ が示した展開法を導入し、任意次数までのモーメントを制約条件とした1変数最大エントロピー分布の特性を研究し、水文学への適用をはかった。さらに、そ

* 正会員 工修 信州大学助手 工学部土木工学科
(〒380 長野市若里500)

** 正会員 工博 信州大学教授 工学部土木工学科 (同上)

*** 工修 (株)オービック (研究当時信州大学大学院工学研究科学生) (〒103 中央区日本橋室町4-2 東邦生命ビル)

れを2変数⁷⁾から多変数⁸⁾へと拡張し、同様な検討を行った。

ところが、分布形が決定されたとしても、それがいかなる資料から母数推定がなされたものであるかにより、確率水文学の信頼性が大きく異なってくる。にもかかわらず、ひとたび頻度解析が行われ、確率水文学が決定されると、その水文学量に従って水工計画が実施され、その水文学量の信頼性がほとんど理論的に吟味されていないのが実情である。

本稿は、このような推定母数に関する不確定性について、正規分布に基づいて議論したものである。

まず最初に、資料数の増加に伴う分布形および確率水文学量の変動を観測資料を用いて検討する。次に、正規分布を支配する母数を確率変数とみなして、平均未知、分散既知の場合、および両者にも未知の場合について、資料数の増加に伴う不確定性の減少度を、推定母数の事後分布のエントロピーの減少度という形で表現する方法を提示・定式化する。最後に、実測水文学資料に適用して本法の妥当性を検証するとともに、非対称母集団に対する適用性を示唆する。

2. 分布形の変動

(1) 対象水文学資料

水工計画に基本的な資料は、降水量および河川流量である。とりわけ、降水量は、水文学資料の中で比較的良好に整備されていることもあり、水工計画にきわめて重要な資料である。そのため、ここでは年降水量および年最大日降水量を対象とする。もちろん、他の水文学量についても同様な検討がなされ得ることはいうまでもない。

使用する観測資料は、長野県内の雨量観測所で得られたものであり、年降水量は19か所、年最大日降水量は28か所である⁹⁾。Table 1に観測所および観測資料数の一覧表を掲げる。資料数の相違は、観測開始年の相違、文献⁹⁾編集後に一部の資料を追加したこと、および欠測によって生じている。

観測資料には地域特性がみられ、平均年降水量は篠ノ井が936 mmで最も少なく、野沢温泉の1926 mmが最も多くなっており、19地点の平均は1221 mmであった。平均年最大日降水量は中野が58 mmで最も少なく、大町の117 mmが最も多くなっており、28地点の平均は70 mmであった。

地域間の相関をみると、年降水量の相関が大きい地点間、たとえば松本と上田、長野と上田等では年最大日降水量の相関も大きく、飯山と上田等の年降水量の相関が小さいところは年最大日降水量の相関も小さくなっている。また、年最大日降水量は局所的な気象変化の影響を受けやすいため、同一地点間では年降水量の相関より

Table 1 Gaging stations of precipitation and number of data (left : annual precipitation, right : annual maximum daily precipitation).

Gaging stations	Number of data	Gaging stations	Number of data
Azumino	56 70	Nakano	76 85
Bessyo	51 52	Nojiri	— 46
Iiyama	68 91	Nozawaonsen	65 69
Iwamura	61 77	Ōmachi	80 89
Karuizawa	— 59	Ōtaki	— 46
Kawakami	— 44	Sakaki	— 41
Kinasa	60 64	Shinonoi	67 78
Kitamaki	47 54	Sugadaira	— 57
Kitashiro	— 46	Suzaka	72 77
Matumoto	85 86	Togakushi	— 55
Minamiotari	— 72	Toyoshina	55 71
Sōga	41 59	Ueda	88 89
Nagano	94 94	Wada	51 54
Nakajō	55 58	Yashiro	64 77

小さくなっている。

どちらの資料においても、周期性や従属性は顕著に現れていない。

(2) 分布形の変動

ここでは、年降水量の分布は正規分布、年最大日降水量の分布はグンベル分布を用い、資料数の増加に伴う頻度図、適合された分布形および確率水文学量の変動を検討する。なお、資料は観測年の新しい順に10個刻みで増加させ、分布の母数推定は積率法を用いた。紙面の都合上、以下の図では一部の観測点についてのみ示すことにする。

a) 年降水量 (正規分布)

上述の19の観測所について資料数の増加に伴う頻度図と分布形の変動の様子を描いた。図-1に、その一例として長野の場合を示す。

頻度図をみると、資料数が10個、20個、30個と増加するに伴い、急激な変動を示しながらその形が整っていく様子が顕著に現れており、資料数が40個以上になるとその変動は比較的緩やかなものになる。このような傾向は観測資料数の大小にかかわらず、19の観測所のすべての地点においてみられる。あてはめた分布形も当然のことながら資料数が30個程度の場合を境にして変動の様子に違いがみられ、資料数が少ない場合の変動は激しく、多くなると安定してきていることがわかる。

次に、長野、松本、上田の3つの観測所について、データ数の増加に伴う10年、30年非超過確率水文学量と50年、100年超過確率水文学量の変動の様子を調べた。図-2に、その一例として長野の場合を示す。

対象とする確率水文学量や地点間での差異は大きいだが、長野ではデータ数が40個ないし50個から、振動しながらも一定の値に収束していく傾向がみられる。

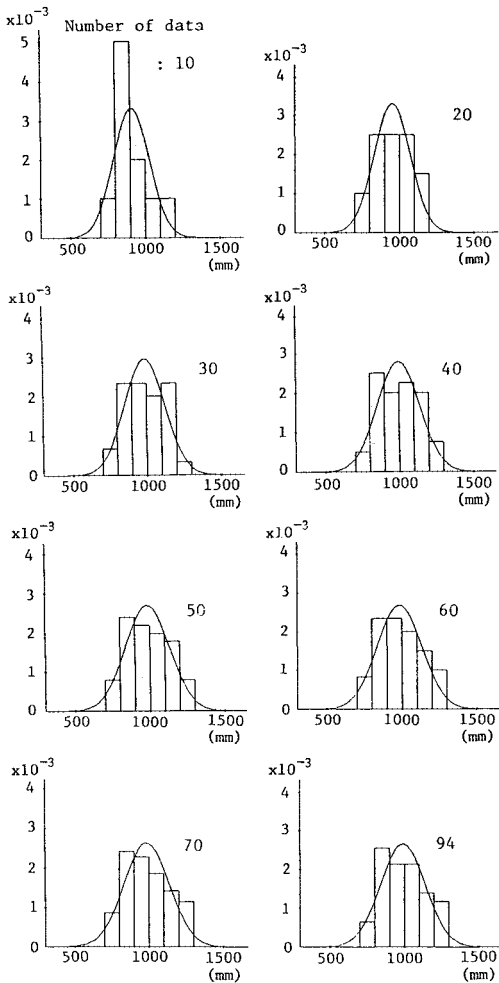


Fig. 1 Change of histogram and fitted normal p. d. f. for annual precipitation with increasing data (Nagano).

b) 年最大日降水量 (グンベル分布)

上述の 28 の観測所について資料数の増加に伴う頻度図と分布形の変動の様子を描いた。図-3 に、その一例として長野の場合を示す。

年降水量の場合と同様に、頻度図および分布形は、資料数が 30 個程度以下の場合には大きく変動し、それを越えると安定してきていることがわかる。

次に、長野、松本、上田の 3 つの観測所について、データ数の増加に伴う 50 年、100 年超過確率水文学の変動の様子を調べた。図-4 に、その一例として長野の場合を示す。

この場合、3 地点ともに、資料数が 30 個を越えると 50 年、100 年超過確率水文学量は安定してきている。

以上、資料数の増加に伴う頻度図、分布形および確率水文学の変動特性を検討したが、大部分は資料数が 30

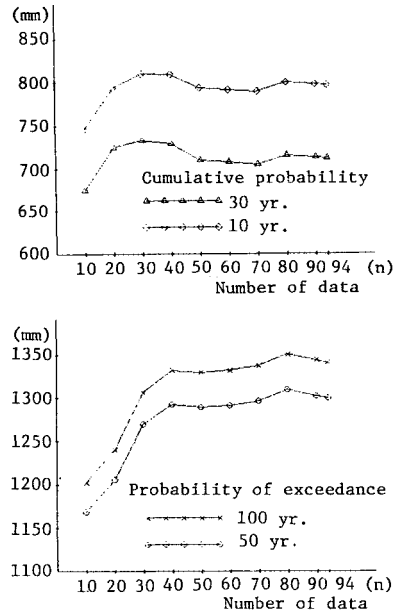


Fig. 2 Change of estimated variate to same fixed recurrence intervals with increasing data of annual precipitation (Nagano).

個ないし 40 個程度以上となると変動は小さくなり安定してきていることがわかる。しかし、資料数が 80 個、90 個程度あっても、分布形の変動、すなわち、推定母数の不確定性が確率水文学に及ぼす影響は無視できないことも付記する。

3. 母数の不確定性とエントロピー

与えられた確率分布が母集団の確率分布と一致しているなら、母数の推定精度が確率水文学の信頼性を支配する。母数を確率変数と考えると、母数の確率分布 (1 次元あるいは多次元) の形状がシャープなほど、その推定精度が高くなることが予想される。そこで、ここでは母数の与えられた事前分布に対し、ベイズ論的立場から得られた事後分布を示し、その分布のエントロピーを計算可能な形に変換する。

なお、以下での議論は、基礎確率変数 (以下では母数を確率変数とみなしているため、それと区別する意味で水文学を表現する確率変数を基礎確率変数とよぶことにする) が確率分布の中で最も基本となる正規分布に従う場合に限った。

(1) 正規情報を取り入れた事後確率分布¹⁰⁾

未知な真の状態 ϕ が状態空間 Φ の要素である場合、すなわち $\Phi = \{\phi\}$ で与えられる場合を考える。 ϕ は次の a) では平均 θ を、 b) では平均 θ と精度 h を表現していることになる。

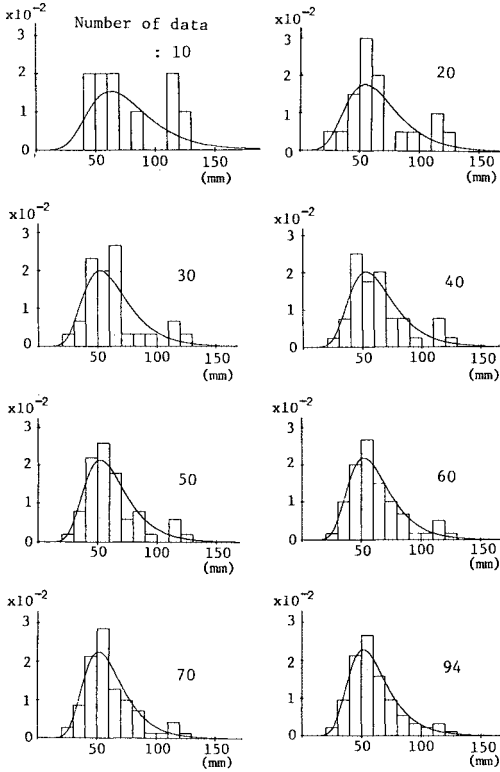


Fig. 3 Change of histogram and fitted Gumbel p.d.f. for annual maximum daily precipitation with increasing data (Nagano).

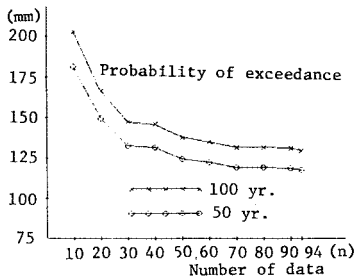


Fig. 4 Change of estimated variate to two fixed recurrence intervals with increasing data of annual maximum daily precipitation (Nagano).

a) 平均未知, 分散既知の場合

いま, n 個の基礎確率変数からなる確率変量 $\tilde{x}(n) = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$ (確率変数であることを明示するとき～を変数の上に付ける) があり, \tilde{x}_i ($i=1, 2, \dots, n$) は $\varphi \in \Phi$ が与えられたとき, 平均 θ , 分散 v をもって互いに独立に正規分布をするものと仮定する. このとき, $\tilde{x}(n)$ の実現値を知ることを大きさ n の正規情報という. 正規情報が利用可能なとき, 平均 θ が未知で分散 v が既知の場合の $\tilde{\theta}$ の事前確率分布の密度関数 $\xi(\theta)$ が正規

分布 $N(m_\theta, v_\theta)$ で与えられるものとする. すなわち,

$$\xi(\theta) = f_N(\theta | m_\theta, v_\theta) \dots \dots \dots (1)$$

ここで, $f_N(\cdot)$ は正規分布の密度関数である.

このとき, $\tilde{x}(n) = x(n)$ が与えられての $\tilde{\theta}$ の事後確率分布の密度関数 $\xi(\theta | x(n))$ は, ベイズ論的立場から次の正規分布で与えられる.

$$\xi(\theta | x(n)) = f_N(\theta | m_{\theta|x(n)}, v_{\theta|x(n)}) \dots \dots \dots (2)$$

ここで,

$$\left. \begin{aligned} m_{\theta|x(n)} &= \frac{\frac{1}{v_\theta} m_\theta + \frac{n}{v} \bar{x}}{\frac{1}{v_\theta} + \frac{n}{v}} = \frac{m_\theta v + n \bar{x} v_\theta}{v + n v_\theta} \\ \frac{1}{v_{\theta|x(n)}} &= \frac{1}{v_\theta} + \frac{n}{v}, \quad v_{\theta|x(n)} = \frac{v v_\theta}{v + n v_\theta} \end{aligned} \right\} \dots \dots \dots (3)$$

b) 平均未知, 分散未知の場合

i) 平均と精度の同時事後確率分布

$\varphi \in \Phi$ が与えられたとき, \tilde{x}_i ($i=1, 2, \dots, n$) は平均 θ , 精度 $h \equiv 1/v$ をもって互いに独立に正規分布をするものと仮定する. このとき, 平均 θ , 精度 h がともに未知の場合の $(\tilde{\theta}, \tilde{h})$ の同時事前確率分布の密度関数 $\xi(\theta, h)$ が, $\alpha', \beta', n', \nu'$ をパラメータとする次の正規—ガンマ分布で与えられるものとする.

$$\begin{aligned} \xi(\theta, h) &= \frac{n'^{1/2}}{(2\pi)^{1/2}} \frac{1}{\Gamma(\nu'/2)} \left(\frac{1}{2} \beta' \nu' \right)^{\nu'/2} \\ &\quad \cdot \exp \left\{ -\frac{1}{2} n' h (\theta - \alpha')^2 \right\} \\ &\quad \cdot \exp \left(-\frac{1}{2} \beta' \nu' h \right) h^{\nu'-1/2} \dots \dots \dots (4) \end{aligned}$$

ここで, $\Gamma(\cdot)$ はガンマ関数である.

このとき,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad S = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \dots \dots \dots (5)$$

で与えられる (θ, h) に関する $\tilde{x}(n)$ に基づく十分統計量を \tilde{x}, \tilde{S} とすると, $\tilde{x} = \bar{x}, \tilde{S} = S$ が得られた後の $(\tilde{\theta}, \tilde{h})$ の同時事後確率分布の密度関数 $\xi(\theta, h | x(n))$ もやはりベイズ論的立場から正規—ガンマ分布で与えられ, そのパラメータは次式から算定される.

$$\left. \begin{aligned} \alpha'' &= \frac{n' \alpha' + n \bar{x}}{n' + n} \\ n'' &= n' + n \\ \beta'' &= \frac{(\nu' \beta' + n' \alpha'^2) + (\nu S + n \bar{x}^2) - n'' \alpha''^2}{\nu' + n} \\ \nu'' &= \nu' + n \end{aligned} \right\} \dots \dots \dots (6)$$

ここで,

$$\nu = n - 1 \dots \dots \dots (7)$$

ii) 平均に関する周辺事後確率分布

$(\tilde{\theta}, \tilde{h})$ の同時事前確率分布が式 (4) に示される正規—ガンマ分布であるとする. このとき, 式 (5) より $\tilde{x} = \bar{x}, \tilde{S} = S$ が得られた後の $\tilde{\theta}$ の周辺事後確率分布の密

度関数 $\xi_\theta(\theta|x(n))$ は、 α'' , n''/β'' , ν'' をパラメーターとする一般化された学生分布

$$\xi_\theta(\theta|x(n)) = \frac{\nu''^{\nu''/2}}{B\left(\frac{1}{2}, \frac{1}{2}\nu''\right)} \left\{ \nu'' + \frac{n''}{\beta''} (\theta - \alpha'')^2 \right\}^{-(\nu''+1)/2} \cdot \left(\frac{n''}{\beta''}\right)^{1/2} \dots\dots\dots (8)$$

で与えられる。ここに、 $B(\cdot, \cdot)$ はベータ関数であり、 α'' , β'' , n'' , ν'' は式 (6) より決定される。

iii) 精度に関する周辺事後確率分布

$(\tilde{\theta}, \tilde{h})$ の同時事前確率分布が式 (4) に示される正規-ガンマ分布であるとする。このとき、式 (5) により $\tilde{x} = \bar{x}$, $\tilde{S} = S$ が得られた後の \tilde{h} の周辺事後確率分布の密度関数 $\xi_h(h|x(n))$ は

$$\xi_h(h|x(n)) = \frac{1}{\Gamma(\nu''/2)} \left(\frac{1}{2}\beta''\nu''\right)^{\nu''/2} h^{\nu''/2-1} \cdot \exp\left(-\frac{1}{2}\beta''\nu''h\right) \dots\dots\dots (9)$$

で与えられる。ここに、 α'' , β'' , n'' , ν'' は式 (6) より決定される。

(2) 事後確率分布のエントロピー

a) 平均未知, 分散既知の場合

式 (2) で示される $\tilde{\theta}$ の事後確率分布のエントロピーは

$$H(\theta) = -\int_{-\infty}^{\infty} \xi(\theta|x(n)) \ln \xi(\theta|x(n)) d\theta = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}v_{\theta(x(n))}} \exp\left\{-\frac{(\theta - m_\theta)^2}{2v_{\theta(x(n))}}\right\} \cdot \ln\left[\frac{1}{\sqrt{2\pi}v_{\theta(x(n))}} \exp\left\{-\frac{(\theta - m_\theta)^2}{2v_{\theta(x(n))}}\right\}\right] d\theta = \ln(2\pi e v_{\theta(x(n))}^{1/2+1}) \dots\dots\dots (10)$$

b) 平均未知, 分散未知の場合

i) 平均と精度の同時事後確率分布のエントロピー
式 (6) をパラメーターとする $(\tilde{\theta}, \tilde{h})$ の同時事後確率分布のエントロピーは

$$H(\theta, h) = -\int_0^{\infty} \int_{-\infty}^{\infty} \xi(\theta, h|x(n)) \ln \xi(\theta, h|x(n)) d\theta dh = -\int_0^{\infty} \int_{-\infty}^{\infty} \frac{n''^{1/2}}{(2\pi)^{1/2}} \frac{1}{\Gamma(\nu''/2)} \left(\frac{1}{2}\beta''\nu''\right)^{\nu''/2} \cdot \exp\left\{-\frac{1}{2}n''h(\theta - \alpha'')^2\right\} \cdot \exp\left(-\frac{1}{2}\beta''\nu''h\right) h^{\nu''-1/2} \cdot \ln\left[\frac{n''^{1/2}}{(2\pi)^{1/2}} \frac{1}{\Gamma(\nu''/2)} \left(\frac{1}{2}\beta''\nu''\right)^{\nu''/2} \cdot \exp\left\{-\frac{1}{2}n''h(\theta - \alpha'')^2\right\} \cdot \exp\left(-\frac{1}{2}\beta''\nu''h\right) h^{\nu''-1/2}\right] d\theta dh = -\ln A + B\left[1 - \psi\left(B + \frac{1}{2}\right) + \ln C\right] + 1$$

$$\dots\dots\dots (11) \quad \text{(補遺 1 参照)}$$

ここで,

$$\left. \begin{aligned} A &= \frac{n''^{1/2}}{(2\pi)^{1/2}} \frac{1}{\Gamma(\nu''/2)} \left(\frac{1}{2}\beta''\nu''\right)^{\nu''/2} \\ B &= (\nu'' - 1)/2 \\ C &= \beta''\nu''/2 \end{aligned} \right\} \dots\dots\dots (12)$$

であり、 $\psi(\cdot)$ はディガンマ関数である。

ii) 平均に関する周辺事後確率分布のエントロピー
式 (8) で示される $\tilde{\theta}$ の事後確率分布のエントロピーは

$$H_\theta(\theta) = -\int_{-\infty}^{\infty} \xi_\theta(\theta|x(n)) \ln \xi_\theta(\theta|x(n)) d\theta = -\int_{-\infty}^{\infty} \frac{\nu''^{\nu''/2}}{B(1/2, \nu''/2)} \left\{ \nu'' + \frac{n''}{\beta''} (\theta - \alpha'')^2 \right\}^{-(\nu''+1)/2} \cdot \left(\frac{n''}{\beta''}\right)^{1/2} \ln\left[\frac{\nu''^{\nu''/2}}{B(1/2, \nu''/2)} \left\{ \nu'' + \frac{n''}{\beta''} (\theta - \alpha'')^2 \right\}^{-(\nu''+1)/2} \left(\frac{n''}{\beta''}\right)^{1/2}\right] d\theta = -\ln D + DEJ_\theta \dots\dots\dots (13) \quad \text{(補遺 2 参照)}$$

ここで,

$$\left. \begin{aligned} D &= \frac{\nu''^{\nu''/2}}{B(1/2, \nu''/2)} \left(\frac{\nu''}{\beta''}\right)^{1/2} \\ E &= (\nu'' + 1)/2 \\ J_\theta &= \int_{-\infty}^{\infty} \frac{\ln\left\{\nu'' + \frac{n''}{\beta''} (\theta - \alpha'')^2\right\}}{\left\{\nu'' + \frac{n''}{\beta''} (\theta - \alpha'')^2\right\}^E} d\theta \end{aligned} \right\} \dots\dots\dots (14)$$

iii) 精度に関する周辺事後確率分布のエントロピー
式 (9) で示される \tilde{h} の事後確率分布のエントロピーは

$$H_h(h) = -\int_0^{\infty} \xi_h(h|x(n)) \ln \xi_h(h|x(n)) dh = -\int_0^{\infty} \frac{1}{\Gamma(\nu''/2)} \left(\frac{1}{2}\beta''\nu''\right)^{\nu''/2} h^{\nu''/2-1} \cdot \exp\left(-\frac{1}{2}\beta''\nu''h\right) \ln\left[\frac{1}{\Gamma(\nu''/2)} \cdot \left(\frac{1}{2}\beta''\nu''\right)^{\nu''/2} h^{\nu''/2-1} \exp\left(-\frac{1}{2}\beta''\nu''h\right)\right] dh = -\ln F + G[1 - \psi(G+1) + \ln C] + 1 \dots\dots (15) \quad \text{(補遺 3 参照)}$$

ここで,

$$\left. \begin{aligned} F &= \frac{1}{\Gamma(\nu''/2)} \left(\frac{1}{2}\beta''\nu''\right)^{\nu''/2} \\ G &= \nu''/2 - 1 \end{aligned} \right\} \dots\dots\dots (16)$$

4. 実測資料への適用と考察

(1) 平均未知, 分散既知の場合

式 (3) の 2 番目の式において、 v_θ は θ の分散に關する事前情報であり、また ν は既知で一定としている

から、これらの値を与えることにより、式(10)は資料数 n の関数として計算可能になる。しかしながら、 v_θ 、 v は主観的に与えねばならず、その値によって $H(\theta)$ の値は異なってくる。そこで、われわれはそれらを次のようにそれぞれ5種類考え、その組合せ(計25個)ごとにエントロピーを算定した。

$$v_\theta = \left. \begin{aligned} & \frac{1}{5} \frac{v'}{n_{\max}/2}, \frac{1}{3} \frac{v'}{n_{\max}/2}, \frac{v'}{n_{\max}/2} \\ & \frac{3v'}{n_{\max}/2}, \frac{5v'}{n_{\max}/2} \end{aligned} \right\} \dots\dots(17)$$

$$v = v'/5, v'/3, v', 3v', 5v'$$

ここで、 n_{\max} は最大観測資料数、 v' は n_{\max} での資料に関する分散である。

式(3)と式(10)から、資料数が増加するとエントロピーが減少することは容易にわかるが、 v_θ と v の組合せも考慮してグラフ化した一例が図-5である。どの地点でも、 v_θ および v の値が大きいほどエントロピーの値が大きくなり、資料数が小さいほどエントロピーの低下分が大きくなっていることがわかる。また、 v_θ が小さく v が大きいほどエントロピーは直線的に低下していることがわかる。なお、エントロピーは分布形がシャープなほど小さな値をとり、エントロピーの値が小さいほど母数の推定精度が高いと判断される。

(2) 平均未知、分散未知の場合

3. では、この場合3通りのエントロピーを計算可能な形に変換して示したが、ここでは平均と精度の同時事後確率分布のエントロピーの計算結果を示す。

式(11)を計算するのに必要なパラメーターは次のように与えられる。

n' 、 α' 、 β' 、 ν' は事前情報であるので、すべての資料を用いて、

$$\left. \begin{aligned} n' &= n_{\max} \\ \alpha' &= \frac{1}{n'} \sum_{i=1}^n x_i \\ \beta' &= \frac{1}{n'-1} \sum_{i=1}^n (x_i - \alpha')^2 \\ \nu' &= n' - 1 \end{aligned} \right\} \dots\dots(18)$$

と定める。また、 \bar{x} 、 S は n 個の資料から計算された平均と分散である。

図-6は、上記のように与えられたパラメーターにより計算された結果をすべての地点について示したものである。エントロピーはほぼ直線的に減少しているが、その値や減少度合には観測所間で差がある。エントロピー値の大小関係により観測所間の母数推定の信頼性の相対的比較を行う場合、エントロピー値は事前情報にも依存するため比較上最適な事前情報の与え方を検討しなければならず、図-6のエントロピー値からこの比較を論ず

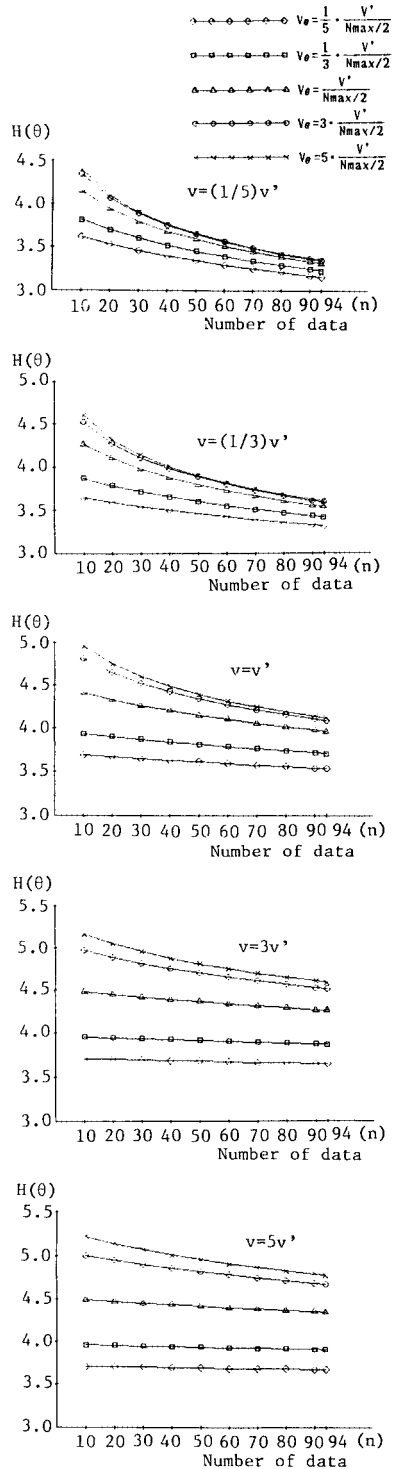


Fig. 5 Change of entropy in connection with the distribution of population parameter (Nagano) (θ =unknown, v =known).

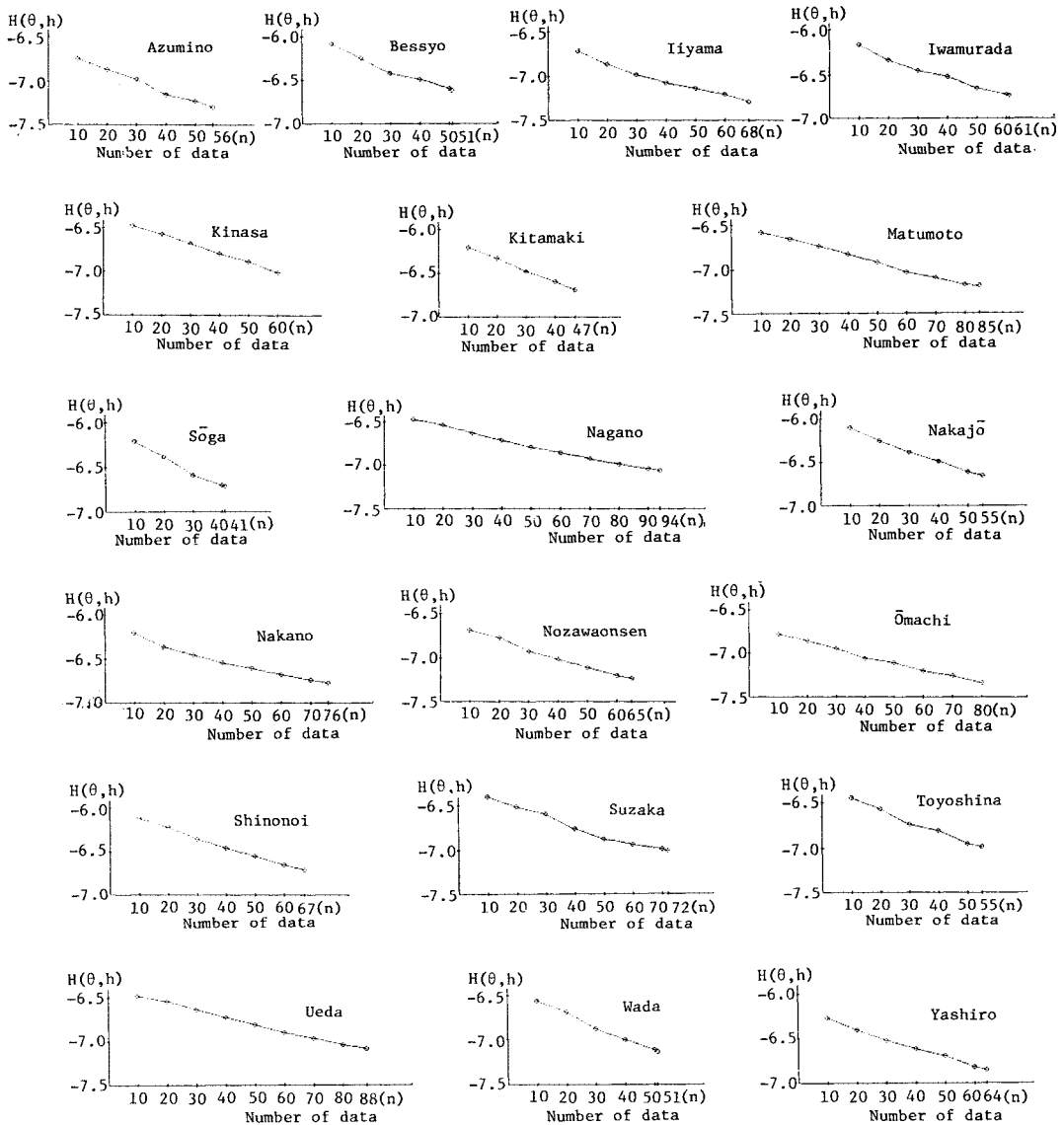


Fig. 6 Change of entropy in connection with the distribution of population parameters (θ =unknown, ν =unknown).

ることは難しい。しかし、連続変数の分布の2つのエントロピーの差、すなわちエントロピーがどの程度変化したかを示す量は情報量であるから、折れ線の傾きが途中で急になる場合には、そのとき増加させた資料はそれ以前の状態で与える情報量が多かったことを示しており、傾きが緩やかなところでは増加させた資料の情報量が少なく、あまり有効でなかったことを示している。

5. 非正規母集団への適用性について

3. では、正規情報が利用可能な場合、すなわち、基礎確率変数が正規分布に従う場合に、母数の事前確率分布を共役事前確率法則（基礎確率変数の密度関数から計

算された尤度関数を密度関数のもつ母数の関数とみなし、その尤度関数から母数と母数に関する十分統計量の項を取り出して構成された母数に関する確率密度関数を共役事前確率法則という）で与えることにより、バイズ論の立場から事後確率分布を計算し、推定母数の不確定性を事後確率分布のエントロピーで評価した。このような立場から母数の不確定性を議論する場合、基礎確率変数に応じた母数の事前分布が与えられる必要がある。

水文頻度解析によく用いられる確率分布のうち、母数の共役分布が求められているものに次のものがあり、基礎確率変数の密度関数を $f(x)$ で表わし、共役分布を未知母数のサフィックスを付けて $f(\lambda)$ と示しておく¹²⁾。

i) 基礎確率変数が指数分布の場合

$$f_x(x) = \lambda \exp(-\lambda x) \dots\dots\dots (19)$$

$$f_\lambda(\lambda) = \nu(\nu\lambda)^{k-1} \exp(-\nu\lambda) / \Gamma(k) \dots\dots\dots (20)$$

ここで、 ν, k は既知とする。

ii) 基礎確率変数が対数正規分布の場合

$$f_x(x) = \frac{1}{(2\pi)^{1/2} \zeta x} \exp\left\{-\frac{1}{2} \left(\frac{\ln x - \lambda}{\zeta}\right)^2\right\} \dots\dots\dots (21)$$

ここで、 ζ は既知とする。

$$f_\lambda(\lambda) = \frac{1}{(2\pi)^{1/2} \sigma} \exp\left\{-\frac{1}{2} \left(\frac{\lambda - \mu}{\sigma}\right)^2\right\} \dots\dots\dots (22)$$

ここで、 σ, μ は既知とする。

このように、基礎確率変数が指数分布や対数正規分布に従う場合、式 (20), (22) を事前確率分布として事後確率分布を計算し、推定母数の不確定性をエントロピーで測ることができる。

基礎確率変数が、グンベル分布、ガンマ分布、あるいはわれわれが研究してきた最大エントロピー分布等をとる場合、著者達の知るかぎり今のところ母数の事前確率分布として有用な共役分布が明らかにされていない。このような場合、事前確率分布として一様分布を仮定することにより同様な理論展開が行われ得るが、その妥当性については議論の余地がある。

6. 結 論

本稿では、確率水文学の不確定性のうち推定母数の不確定性に着目し、資料数の増加に伴う不確定性の減少をエントロピーを用いて表現することを試みた。結果を要約すると次のとおりである。

(1) 正規分布、グンベル分布の変動は、資料数が30個ないし40個程度までかなり大きく不安定であるが、それらを越えると相対的に安定してくることがわかった。

(2) 正規分布、グンベル分布から計算される確率水文学の変動は、資料数が40個ないし50個程度から、変動しながらも一定値に収束していくことがわかった。

(3) 正規分布の推定母数の不確定性は、平均未知、分散既知の場合の平均に関する事後確率分布、平均、分散ともに未知の場合の平均と精度の同時事後確率分布、平均に関する周辺事後確率分布、精度に関する周辺事後確率分布のエントロピーを測ることにより定式化した。

(4) 平均未知、分散既知の場合の実測資料への適用の結果、どの地点でも v_0 および v の値が大きいかほどエントロピーの値は大きくなり、資料数が小さいほどエントロピーの低下分が大きくなっていることがわかった。また、 v_0 が小さく v が大きいほどエントロピーは直線的に低下していることもわかった。

(5) 平均、分散ともに未知の場合の平均と精度の同

時事後確率分布のエントロピー式を実測資料へ適用した結果、エントロピーは資料数に比例してほぼ直線的に減少しているが、その値や減少度合には観測所間で差異を生じることがわかった。

(6) 非正規母集団への適用性については、今のところ基礎確率変数が指数分布をする場合と対数正規分布をする場合について母数の事前確率分布として有用な共役分布が知られており、本稿での議論が適用可能であることを示した。

なお、本研究を行うにあたり、計算等で信州大学大学院工学研究科学生 上原 剛君の協力を得たことを記し、謝意を表す。また、計算にあたって、信州大学情報処理センター HITAC M-240H を使用した。

補遺1. 式 (11) の誘導

$$\begin{aligned}
 H(\theta, h) = & - \int_0^\infty \int_{-\infty}^\infty \frac{n^{\nu/2}}{(2\pi)^{1/2}} \frac{1}{\Gamma(\nu/2)} \left(\frac{1}{2} \beta^\nu \nu^\nu\right)^{\nu/2} \\
 & \cdot \exp\left\{-\frac{1}{2} n^\nu h(\theta - \alpha^\nu)^2\right\} \\
 & \cdot \exp\left(-\frac{1}{2} \beta^\nu \nu^\nu h\right) h^{\nu-1/2} \\
 & \cdot \ln\left[\frac{n^{\nu/2}}{(2\pi)^{1/2}} \frac{1}{\Gamma(\nu/2)} \left(\frac{1}{2} \beta^\nu \nu^\nu\right)^{\nu/2}\right. \\
 & \cdot \exp\left\{-\frac{1}{2} n^\nu h(\theta - \alpha^\nu)^2\right\} \\
 & \left. \cdot \exp\left(-\frac{1}{2} \beta^\nu \nu^\nu h\right) h^{\nu-1/2}\right] d\theta dh \dots\dots (23)
 \end{aligned}$$

ここで、

$$\left. \begin{aligned}
 A = & \frac{n^{\nu/2}}{(2\pi)^{1/2}} \frac{1}{\Gamma(\nu/2)} \left(\frac{1}{2} \beta^\nu \nu^\nu\right)^{\nu/2} \\
 B = & (\nu-1)/2, \quad C = \beta^\nu \nu^\nu/2, \quad S = n^\nu/2
 \end{aligned} \right\} \dots\dots (24)$$

とおくと、

$$\begin{aligned}
 H(\theta, h) = & - \int_0^\infty \int_{-\infty}^\infty A \exp(-Ch) \exp\{-Sh(\theta - \alpha^\nu)^2\} h^B \\
 & \cdot \ln A - Ch - Sh(\theta - \alpha^\nu)^2 + B \ln h \} d\theta dh \\
 = & -(\ln A) I_{0h} + AC J_{0h} + AS K_{0h} - AB L_{0h} \\
 & \dots\dots\dots (25)
 \end{aligned}$$

ここで、

$$\begin{aligned}
 I_{0h} = & \int_0^\infty \int_{-\infty}^\infty A \exp(-Ch) \exp\{-Sh(\theta - \alpha^\nu)^2\} h^B d\theta dh \\
 = & 1 \dots\dots\dots (26)
 \end{aligned}$$

$$\begin{aligned}
 J_{0h} = & \int_0^\infty \int_{-\infty}^\infty \exp(-Ch) \exp\{-Sh(\theta - \alpha^\nu)^2\} h^{B+1} d\theta dh \\
 = & \frac{\pi^{1/2}}{S^{1/2}} \int_0^\infty \exp(-Ch) h^{B+1/2} dh = \frac{\pi^{1/2}}{S^{1/2}} \frac{\Gamma(B+3/2)}{C^{B+3/2}} \\
 & \dots\dots\dots (27)
 \end{aligned}$$

$$\begin{aligned}
 K_{0h} = & \int_0^\infty \int_{-\infty}^\infty \exp(-Ch) \exp\{-Sh(\theta - \alpha^\nu)^2\} h^{B+1} \\
 & \cdot (\theta - \alpha^\nu)^2 d\theta dh = \frac{1}{2} \frac{\pi^{1/2}}{S^{3/2}} \int_0^\infty \exp(-Ch) h^{B+1/2} dh
 \end{aligned}$$

$$= \frac{1}{2} \frac{\pi^{1/2}}{S^{3/2}} \frac{\Gamma(B+1/2)}{C^{B+1/2}} \dots\dots\dots (28)$$

$$L_{oh} = \int_0^\infty \int_\infty^\infty \exp(-Ch) \exp\{-Sh(\theta-\alpha^n)^2\} h^B \cdot \ln h d\theta dh = \frac{\pi^{1/2}}{S^{1/2}} \int_0^\infty \exp(-Ch) h^{B-1/2} \ln h dh = \frac{\pi^{1/2}}{S^{1/2}} \frac{1}{C^{B+1/2}} \Gamma\left(B+\frac{1}{2}\right) \left\{ \psi\left(B+\frac{1}{2}\right) - \ln C \right\} \dots\dots\dots (29)$$

式 (26), (27), (28), (29) を式 (25) に代入して整理すると,

$$H(\theta, h) = -\ln A + \frac{\pi^{1/2} A \Gamma\left(B+\frac{1}{2}\right)}{S^{1/2} C^{B+1/2}} \left[B \left\{ 1 - \psi\left(B+\frac{1}{2}\right) + \ln C \right\} + 1 \right] = -\ln A + B \left\{ 1 - \psi\left(B+\frac{1}{2}\right) + \ln C \right\} + 1 \dots\dots\dots (30)$$

補遺 2. 式 (13) の誘導

$$H_\theta(\theta) = -\int_\infty^\infty \frac{\nu^n \nu^{n/2}}{B(1/2, \nu^n/2)} \left\{ \nu^n + \frac{n^n}{\beta^n} (\theta - \alpha^n)^2 \right\}^{-\nu^n+1/2} \cdot \left(\frac{n^n}{\beta^n} \right)^{1/2} \ln \left[\frac{\nu^n \nu^{n/2}}{B(1/2, \nu^n/2)} \left\{ \nu^n + \frac{n^n}{\beta^n} (\theta - \alpha^n)^2 \right\}^{-\nu^n+1/2} \left(\frac{n^n}{\beta^n} \right)^{1/2} \right] d\theta \dots\dots\dots (31)$$

ここで,

$$D = \frac{\nu^n \nu^{n/2}}{B(1/2, \nu^n/2)} \left(\frac{\nu^n}{\beta^n} \right)^{1/2} \left. \begin{aligned} E &= (\nu^n + 1)/2 \end{aligned} \right\} \dots\dots\dots (32)$$

とおくと,

$$H_\theta(\theta) = -\int_\infty^\infty D \left[\nu^n + \frac{n^n}{\beta^n} (\theta - \alpha^n)^2 \right]^{-E} \left[\ln D - E \cdot \ln \left\{ \nu^n + \frac{n^n}{\beta^n} (\theta - \alpha^n)^2 \right\} \right] d\theta = -(\ln D) I_\theta + DE J_\theta \dots\dots\dots (33)$$

ここで,

$$I_\theta = \int_\infty^\infty D \left[\nu^n + \frac{n^n}{\beta^n} (\theta - \alpha^n)^2 \right]^{-E} = 1 \dots\dots\dots (34)$$

$$J_\theta = \int_\infty^\infty \frac{\ln \left\{ \nu^n + \frac{n^n}{\beta^n} (\theta - \alpha^n)^2 \right\}}{\left\{ \nu^n + \frac{n^n}{\beta^n} (\theta - \alpha^n)^2 \right\}^E} d\theta \dots\dots\dots (35)$$

式 (34), (35) を式 (33) に代入すると,

$$H_\theta(\theta) = -\ln D + DE J_\theta \dots\dots\dots (36)$$

補遺 3. 式 (15) の誘導

$$H_h(h) = -\int_0^\infty \frac{1}{\Gamma(\nu^n/2)} \left(\frac{1}{2} \beta^n \nu^n \right)^{\nu^n/2} h^{\nu^n/2-1} \cdot \exp\left(-\frac{1}{2} \beta^n \nu^n h\right) \ln \left\{ \frac{1}{\Gamma(\nu^n/2)} \left(\frac{1}{2} \beta^n \nu^n \right)^{\nu^n/2} \cdot h^{\nu^n/2-1} \exp\left(-\frac{1}{2} \beta^n \nu^n h\right) \right\} dh \dots\dots\dots (37)$$

ここで,

$$\left. \begin{aligned} F &= \frac{1}{\Gamma(\nu^n/2)} \left(\frac{1}{2} \beta^n \nu^n \right)^{\nu^n/2} \\ G &= \nu^n/2 - 1 \end{aligned} \right\} \dots\dots\dots (38)$$

とおくと,

$$H_h(h) = -\int_0^\infty F h^G \exp(-Ch) (\ln F + G \ln h - Ch) dh = -(\ln F) I_h - FG J_h + CF K_h \dots\dots\dots (39)$$

$$I_h = \int_0^\infty F h^G \exp(-Ch) dh = 1 \dots\dots\dots (40)$$

$$J_h = \int_0^\infty h^G \exp(-Ch) \ln h dh = \frac{\Gamma(G+1)}{C^{G+1}} \{\psi(G+1) - \ln C\} \dots\dots\dots (41)$$

$$K_h = \int_0^\infty h^{G+1} \exp(-Ch) dh = \frac{\Gamma(G+2)}{C^{G+2}} \dots\dots\dots (42)$$

式 (40), (41), (42) を式 (39) に代入して整理すると,

$$H_h(h) = -\ln F + G \{1 - \psi(G+1) + \ln C\} + 1 \dots\dots\dots (43)$$

参 考 文 献

- 1) 高棹琢馬・池淵周一: エントロピー的にみた降雨・流出変換特性とそのモデル化, 京都大学防災研究所年報, No. 23, B-2, pp. 193~209, 1980年.
- 2) Sonuga, J. O. : Principle of maximum entropy in hydrologic frequency analysis, Journal of Hydrology, Vol. 17, pp. 177~191, 1972.
- 3) Sonuga, J. O. : Entropy principle applied to the rainfall-runoff process, Journal of Hydrology, Vol. 30, pp. 81~94, 1976.
- 4) 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, No. 335, pp. 89~95, 1983年.
- 5) 寒川典昭・荒木正夫・小林紀之: 種々の水流量への 1 変数 MEP 分布の適用性, 信州大学工学部紀要, No. 58, pp. 27~40, 1985年.
- 6) Wragg, A. and Dowson, D. C. : Fitting Continuous probability density functions over $[0, \infty)$ using information theory ideas, IEEE Transactions on Information Theory, Vol. IT-16, pp. 226~230, 1970.
- 7) 寒川典昭・荒木正夫・寺島 彰: 2 変数 MEP 分布とその特性に関する研究, 水理講演会論文集, No. 28, pp. 397~402, 1984年.
- 8) 寒川典昭・荒木正夫・佐藤健次: 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 (投稿中).
- 9) 建設省北陸地方建設局千曲川工事事務所編集: 千曲川・犀川河川要覧 (雨量資料編), 1976年.
- 10) 宮沢光一: 情報・決定理論序説, 岩波書店, pp. 126~150, 1971年.
- 11) たとえば, 大泉充郎・本多波雄・野口正一: 情報理論, オーム社, p. 80, 1962年.
- 12) たとえば, 伊藤 学・亀田宏行訳: 土木・建築のための確率・統計の基礎, 丸善株式会社, p. 346, 1977.

(1986. 3. 6・受付)