

信州大学工学部 正 員 荒木正夫 信州大学工学部 正 員 寒川典昭
 信州大学大学院 学生員 大原基秀 信州大学工学部 森 茂

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。統計モーメントで情報を与える方法は、4次までは分布の形状と関連が付き、しかも解の存在条件が明らかにされているため、今まで用いられてきた。しかし、少ない情報からうまく母集団を推定するために、さらに良い情報の与え方を探索する必要がある。ここでは、統計モーメントを含めて4つの関数を候補として、1変数最大エントロピー分布の母集団に対する適合度を検討する。

2. 情報(制約条件)の与え方

確率密度関数を $p(x)$ とし、密度関数が具備すべき条件と任意関数 $g_r(x)$ の期待値を得られる情報とすると、この条件の下で1変数最大エントロピー分布は次式で与えられる。

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

ここで、 λ_r はラグランジュの未定乗数である。(1)式は $g_r(x)$ の与え方により様々な分布形をとる。そこで、 $g_r(x)$ の具体的な形として、次の4つの代表的な関数を選んだ。

$$g_1^1(x) = x^r \cdots (2), \quad g_1^2(x) = \frac{1}{x^r} \cdots (3), \quad g_1^3(x) = \exp(-x^{\frac{r}{M}}) \cdots (4), \quad g_1^4(x) = (\ln x)^r \cdots (5).$$

ここで、 M はデータの平均値である。また、推定するパラメータの数や高次モーメントのもつ誤差をなるべく少なくするために、 $r = 1, 2$ とした。

3. 適合度の検討

母集団として、任意の形状を表現できるガンマ分布を採用した。パラメータ α, β をもつガンマ分布 $(=G(\alpha, \beta))$ は、密度関数

$$p(x) = \frac{1}{\beta^{\alpha+1} \Gamma(\alpha+1)} x^{\alpha} \exp(-\frac{x}{\beta}), \quad \alpha \geq 0, \beta > 0, x \geq 0 \quad (6)$$

より定義される。今回は、 $G(0,1), G(1,0.7), G(1,1), G(3,1), G(3,2), G(3,3)$ を母集団とし、これより50, 100, 1000個の乱数を発生させ、それぞれの頻度図と母集団に対する適合度を検討した。

Fig. 1 ~ Fig. 4は、 $G(3,1)$ を母集団として50個の乱数を発生させ、それぞれの情報の与え方について、頻度図と最大エントロピー分布、及び母集団を示したものである。従来から採用してきた x^r を情報として与えた場合、他と比べややなだらかな分布形となり、あまり長く尾を引かない。これは、 r を2までとしたことにより、データの平均、分散程度までの情報しか取り込まれなかったためと思われる。逆に $1/x^r$ の場合は、かなり左側に歪んで尖った分布形となり長く尾を引いている。他の2つの場合は、ほぼ両者の中間的な分布形となっており、 $1/x^r, \exp(-x^{\frac{r}{M}}), (\ln x)^r, x^r$ の順に徐々に右側にピークが移り、なだらかな分布形になっている。また、100個、1000個とデータ数を増加させ

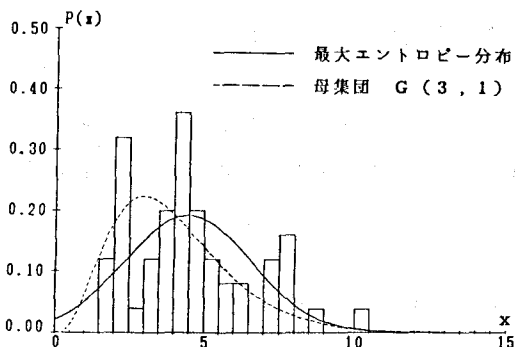


Fig. 1 x を情報として与えた場合の最大エントロピー分布

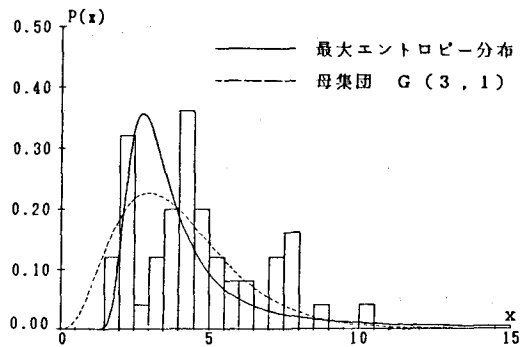


Fig. 2 $1/x$ を情報として与えた場合の最大エントロピー分布

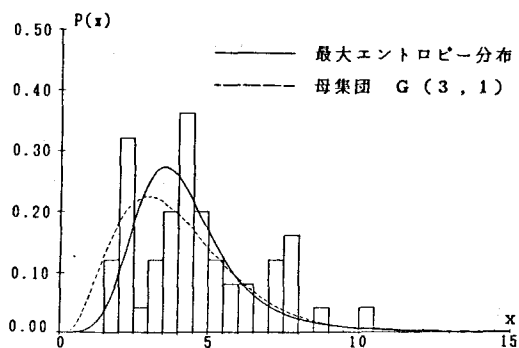


Fig. 3 $\exp(-x/H)$ を情報として与えた場合の最大エントロピー分布

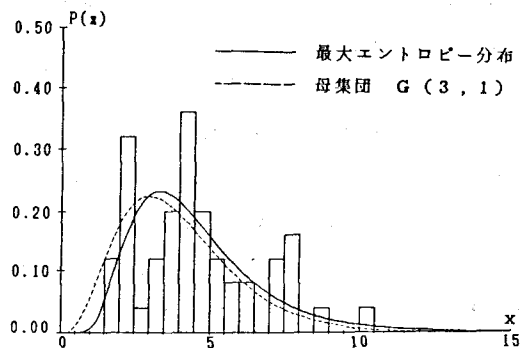


Fig. 4 $(\ln x)^r$ を情報として与えた場合の最大エントロピー分布

た場合でも、前述の傾向が強くなり、 $1/x$ の場合は母集団や頻度図とはかなり異った分布形になってしまったが、他は母集団の分布形に近づいてきている。さらに、適合度の 1 評価方法としてリターンピリオドを求め比較したものを Fig. 5 に示した。これより、リターンピリオドの大きいところでは、母集団に対して、 x の場合のみ危険側、他はすべて、かなり安全側の値をとることがわかった。

これらのことは、 $G(3, 1)$ の場合のみに限らず

検討したすべての母集団に対しても同様であった。以上の結果より、最大エントロピー分布は同じ母集団に対しても、情報の与え方によりかなり異った分布形になることがわかった。また、今回検討した 4 通りの情報の与え方が最大エントロピー分布に及ぼすおおよその傾向は把握できた。

4. おわりに

3. の考察から、 x と他の 1 つを組み合わせた情報が良さそうである。今後は、これらの情報の組み合わせ方と実測水文量への適用を検討していきたい。

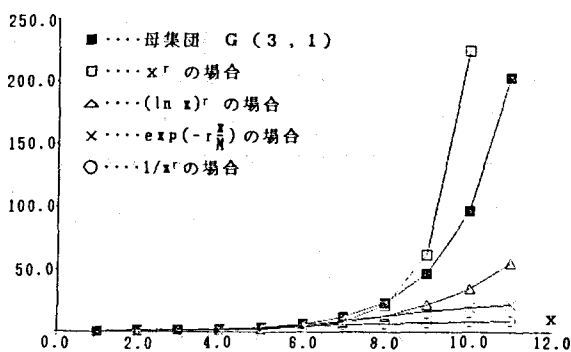


Fig. 5 リターンピリオド