

II-3

情報の与え方と最大エントロピー分布(その2)

信州大学工学部 正 員 寒川典昭

信州大学工学部 正 員 荒木正夫

水資源開発公団 正 員 大原基秀

1. はじめに

最大エントロピー分布は与えられる情報により分布形が決定されるため、少ない資料よりうまく母集団を推定し、変動が少なく妥当な値のリターンピリオドを得るために、適切な情報の与え方を探索する必要がある。従来、実測データへの適用においては統計モーメント x^r のみを情報とした最大エントロピー分布を議論してきたが、新たに昨年、 $1/x^r, \exp(-r\frac{x}{M}), (\ln x)^r$ ($M: x$ の平均、 r : 正の実数) の3種の関数の期待値を個々に情報とした場合を検討した。¹⁾しかし、どの場合も改良の余地が残された。そこで本稿では、これらの情報を組み合わせて与えた場合を1変数最大エントロピー分布について、シミュレーションと実測データへの適用性から検討する。尚、これは2変数以上の多変数への拡張を目差したものである。

2. シミュレーションでの検討

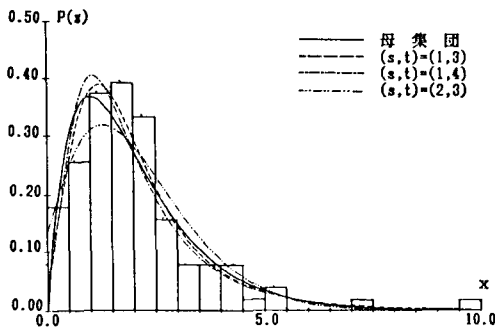
今回は、 x^s と $\exp(-t\frac{x}{M})$ (s, t : 正の実数) の両方の期待値を情報とする場合を検討した。この場合の最大エントロピー分布は次式のようにになる。

$$p(x) = \exp\{-1 - \lambda_0 - \lambda_1 x^s - \lambda_2 \exp(-t\frac{x}{M})\}. \quad (1)$$

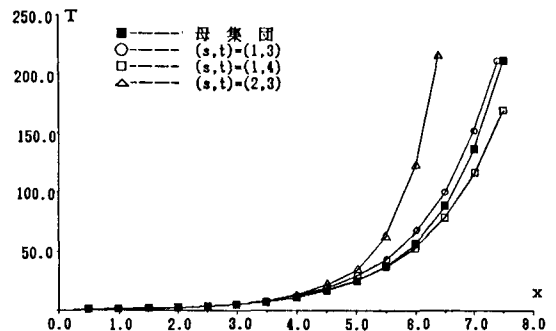
そこで、この分布の特性を知るために、多くの分布形状を表現できるガンマ分布($=G(\alpha, \beta)$)を母集団と仮定し、これより 50, 100, 1000個と発生させた乱数をデータとして(s, t)の値をかえ、それぞれの頻度図と母集団に対する適合度を検討した。このガンマ分布の密度関数は次式より定義される。

$$p(x) = \frac{1}{\beta \alpha^{\alpha+1} \Gamma(\alpha+1)} x^{\alpha} \exp(-\frac{x}{\beta}), \quad \alpha \geq 0, \beta > 0, x \geq 0. \quad (2)$$

Fig. 1及びFig. 2は、それぞれ $G(1,1), G(3,1)$ を母集団として100個の乱数を発生させて適合させた例である。これらより頻度図及び母集団に対し図中で用いた分布は良い適合度が得られることがわかった。一般に、分布形は s の値を大きくするとピークの位置が少しずつ右に下がり、なだらかで丸みを帯びてきてあまり右側に尾を引かなくなるため大きな値のデータに対するリターンピリオドが大きくなる。逆に、 t の値を大きくするとピークの位置が少しずつ左に上がり、尖って左側に歪んでくるが、右側に長く尾を引くようになるためリターンピリオドは小さくなる。また、 s の値が3以上になると、 t の値に関係なく適合度が悪く

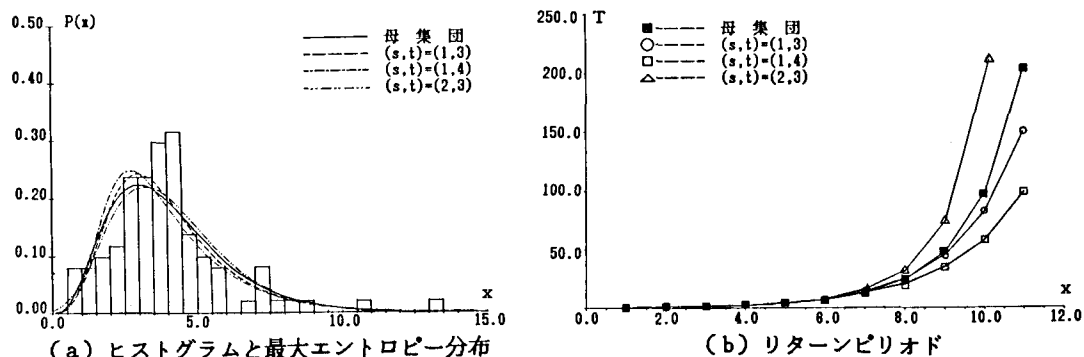


(a) ヒストグラムと最大エントロピー分布



(b) リターンピリオド

Fig. 1 $G(1, 1)$ を母集団とした場合

Fig. 2 $G(3, 1)$ を母集団とした場合

なるため、このような組み合わせの場合 s の値は 2 以下とすることが妥当であると思われる。

3. 実測データへの適用

Fig. 3 は、(1)式の分布を長野の年最大日降水量に対し適用し、従来の x^r ($r=1,2$) の期待値を情報とした場合と比較したものである。(a)から従来の方法に比べ、ヒストグラムに対して両側の裾の部分の適合度が改善されていることがわかる。また (b) はデータを年代の古い順に並べ、前半分までのデータの最大値のリターンピリオドについてデータ数の増加に伴う変動をみたものである。従来の方法より組み合わせで情報を与えた方が、データ数の増加による変動が小さく、小さな値を示し、 s の値が小さい程、また t の値が大きい程、変動幅、値共に小さくなる傾向がみうけられる。これは右側に長く尾を引くようになり、より大きな降雨の生起の可能性を示しているためと思われる。

特に、 $(s,t)=(1,3), (1,4)$ などの場合は変動が少なく、シミュレーションでの乱数の発生状況やデータなどから考察して妥当な値のリターンピリオドを示していると考えられる。

以上より、 x^s と $\exp(-tx^{\frac{s}{t}})$ の両方の期待値を情報とした最大エントロピー分布は、年最大日降水量のような実測データに対しても、ヒストグラムとの適合度や安定したリターンピリオド推定の立場から、かなり有効な分布であることが検証された。

4. おわりに

本稿では、母集団分布及びリターンピリオド推定の立場から、シミュレーションと実測データでの検討により、最大エントロピー分布に対する有効な情報の与え方を見いだすことができた。今後、さらに多くの実測データに対し、他の既存分布との比較などにより、その有効性を実証するとともに、このような情報の与え方を多変数へと拡張したい。最後に、本研究を行なうにあたり協力願った信州大学工学部生の森茂君に謝意を表す。

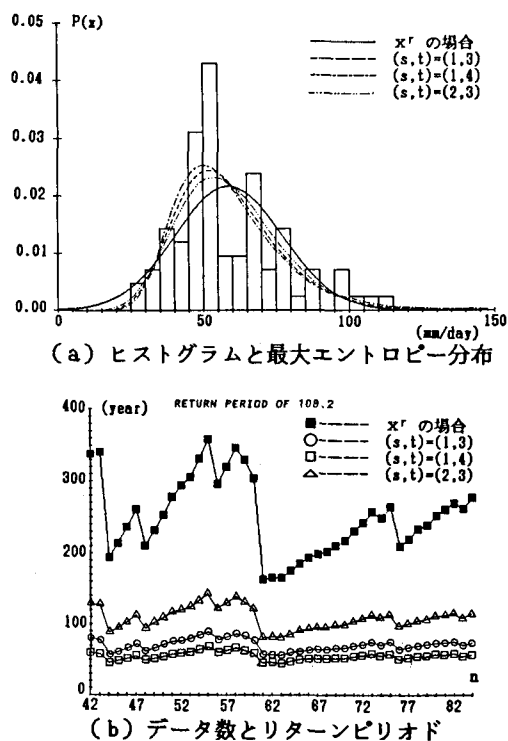


Fig. 3 長野年最大日降水量に対する適合

1) 荒木、寒川、大原、森：情報の与え方と最大エントロピー分布、61年度中部支部講演集、1987年3月。