

II-47 非正規母集団の推定母数の信頼性(その2)

信州大学大学院 学生員 上原 剛
 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

1. はじめに

水工計画で用いられている確率水文学には、小標本の資料から計算されているために、多くの不確定さが内在しているものと思われる。この確率水文学の不確定さは、与えられた分布形が母集団の分布形と一致しているならば、推定母数の信頼性によって決定される。これまで我々は、正規分布(平均:未知、分散:既知)、(平均、分散:未知)¹⁾・指数分布・対数正規分布(λ :未知、 ξ :既知)²⁾の未知母数を確率変数とみなして推定母数の信頼性をエントロピーで表現し、実測資料へ適用してきた。本稿では、対数正規分布(λ :既知、 ξ :未知)・グンベル分布(a :既知、 b :未知)について検討する。

2. 資料数増加に伴う確率水文学の変動性

図-1は、月降水量については対数正規分布より、年最大日降水量についてはグンベル分布より求めた超過100年確率水文学と資料数との関係を図にしたものである。この図から、資料数が少ない時程、確率水文学の変動が大きいことがわかる。

3. 理論式

(1) 対数正規分布(λ :既知、 ξ :未知)の場合

n 個の確率変数 $\tilde{x}(n) = (\tilde{x}_1, \dots, \tilde{x}_n)$ が、 ξ と無関係な一定で既知の λ をもって、次式の対数正規分布に従うとする。

$$f(X) = \frac{1}{\sqrt{2\pi}\xi X} \exp\left\{-\frac{1}{2} \frac{(\ln X - \lambda)^2}{\xi^2}\right\} \quad (1)$$

確率変数 ξ の事前分布に

$$f(\xi) = \frac{1}{\xi_2 - \xi_1} \quad , \quad (\xi_1 \leq \xi \leq \xi_2) \quad (2)$$

なる一様分布を仮定すると、 $\tilde{x}(n) = x(n)$ が与えられての ξ の事後分布は、次式ようになる。

$$f(\xi | X(n)) = \frac{1}{K \xi^B} \exp\left(-\frac{A}{\xi}\right) \quad (3)$$

従って、 ξ の事後分布のエントロピーは、次式で与えられる。

$$H(\xi) = \ln K + \frac{BI}{K} - \frac{1}{K} \left[e^{At} \sum_{r=0}^{B-1} (-1)^r \frac{(B-1)! t^{(B-r-1)}}{(B-r-1)! A^r} \right]_{t=\frac{1}{\xi_1}}^{t=\frac{1}{\xi_2}} \quad (4)$$

ただし、(3)、(4)式のA、B、K、Iは、次式で表される。

$$A = -\frac{1}{2} \left\{ \sum_{i=1}^n (\ln X_i)^2 - 2\lambda \sum_{i=1}^n \ln X_i + n\lambda^2 \right\} \quad , \quad B = \frac{n}{2} \quad ,$$

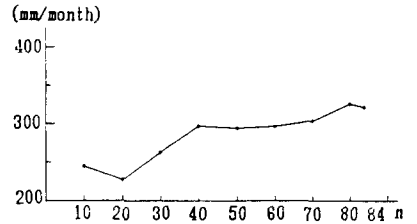
$$K = \left[\frac{e^{At}}{A} \sum_{r=0}^{B-2} (-1)^r \frac{(B-2)! t^{(B-r-2)}}{(B-r-2)! A^r} \right]_{t=\frac{1}{\xi_1}}^{t=\frac{1}{\xi_2}} \quad , \quad I = \int_{\xi_1}^{\xi_2} \frac{1}{\xi^B} \exp\left(-\frac{A}{\xi}\right) \ln \xi \, d\xi \quad (5)$$

(2) グンベル分布(a :既知、 b :未知)の場合

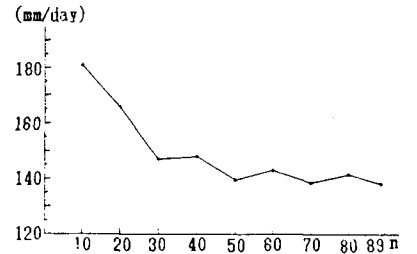
確率変数 $\tilde{x}(n)$ が、 b と無関係な一定で既知の a をもって、次式のグンベル分布に従うとする。

$$f(X) = a \exp\{-a(X-b) - e^{-a(X-b)}\} \quad (6)$$

確率変数 b の事前分布に



a) 上田(6月の月降水量、対数正規分布)



b) 上田(年最大日降水量、グンベル分布)

図-1. 資料数と確率水文学

$$f(b) = \frac{1}{b_2 - b_1}, \quad (b_1 \leq b \leq b_2) \quad (7)$$

なる一様分布を仮定すると、 $\tilde{x}(n) = x(n)$ が与えられての \tilde{b} の事後分布は、次式のようになる。

$$f(b|X(n)) = \frac{1}{K} \exp(abn - Ce^{ab}) \quad (8)$$

従って、 \tilde{b} の事後分布のエントロピーは、次式で与えられる。

$$H(b) = \ln K - \frac{nJ}{K} - \frac{1}{aK} \left[e^{-Ct} \sum_{r=0}^n \frac{n!}{(n-r)!} \frac{t^{(n-r)}}{C^r} \right]_{t=e^{ab_2}} - \frac{1}{aK} \left[e^{-Ct} \sum_{r=0}^n \frac{n!}{(n-r)!} \frac{t^{(n-r)}}{C^r} \right]_{t=e^{ab_1}} \quad (9)$$

ただし、(8)、(9)式のC、J、Kは、次式で表される。

$$C = \sum_{i=1}^n e^{-ax_i}, \quad J = \int_{b_1}^{b_2} ab \exp(abn - Ce^{ab}) db,$$

$$K = \frac{-1}{aC} \left[e^{-Ct} \sum_{r=0}^{n-1} \frac{(n-1)!}{(n-r-1)!} \frac{t^{(n-r-1)}}{C^r} \right]_{t=e^{ab_2}} - \frac{1}{aC} \left[e^{-Ct} \sum_{r=0}^{n-1} \frac{(n-1)!}{(n-r-1)!} \frac{t^{(n-r-1)}}{C^r} \right]_{t=e^{ab_1}} \quad (10)$$

4. 実測資料への適用

(1) 対数正規分布 (λ : 既知、 ζ : 未知) の場合

図-2は、月降水量を用いて計算した(4)式のエントロピーと資料数との関係を図にしたものである。計算にあたり、事前分布のパラメタは、全資料を10個ずつに分け(端数は切り捨てた)その各々について ζ を計算し、その中の最小値を ζ_1 、最大値を ζ_2 として与えた。また、既知であるとした λ は、

$$\lambda = \ln \theta - \frac{1}{2} \ln \left(\frac{v}{\theta^2} + 1 \right) \quad (11)$$

とした。 θ と v は、全資料から計算した平均と分散である。この図から、資料数増加に伴いエントロピーの値は、全体的に減少していく(推定母数の信頼性が高くなる)ことがわかる。

(2) グンベル分布 (a : 既知、 b : 未知) の場合

図-3は、年最大日降水量を用いて計算した(9)式のエントロピーと資料数との関係を図にしたものである。 b_1 と b_2 は、 ζ_1 と ζ_2 の場合と同様にして、最小値を b_1 、最大値を b_2 とした。また、 a は、

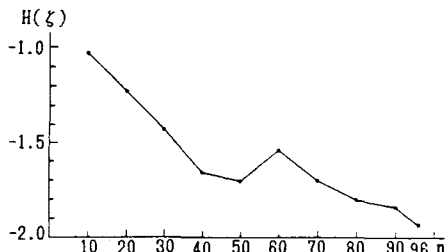
$$a = \sqrt{\frac{\pi^2}{6v}} \quad (12)$$

とした。この図から、資料数増加に伴いエントロピーの値は、減少していくことがわかる。

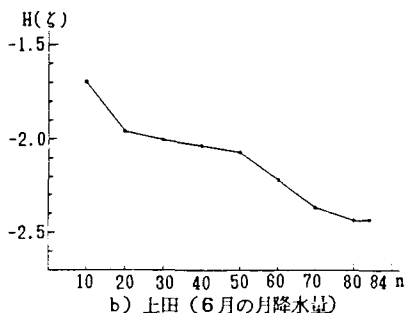
5. おわりに

今回、事前分布を一様分布として与えたが、共役事前分布よりその分布形を決定することができないかを検討するとともに、母数が2つとも未知の場合についても計算を行っていきたい。

最後に、当時、信州大学学部生であった深川君の協力に感謝する。

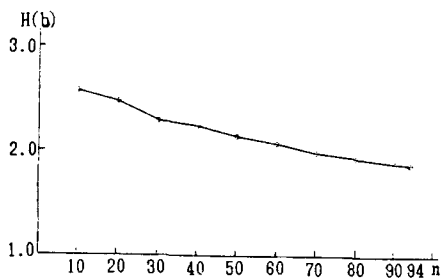


a) 長野(6月の月降水量)

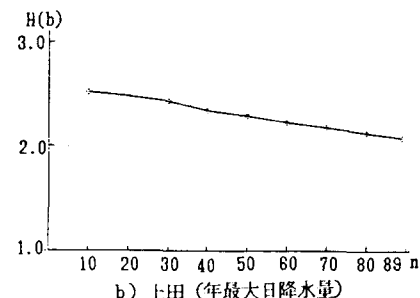


b) 上田(6月の月降水量)

図-2. 対数正規分布 (λ : 既知、 ζ : 未知) の場合の資料数とエントロピー



a) 長野(年最大日降水量)



b) 上田(年最大日降水量)

図-3. グンベル分布 (a : 既知、 b : 未知) の場合の資料数とエントロピー

1) 荒木、寒川、渡辺: 確率分布の母数の不確定性評価法、59年度中部支部研究発表会講演集、1985年。
2) 荒木、寒川、上原: 指数分布・対数正規分布の母数推定の信頼性、80年度中部支部講演集、1986年。