信頼性を付与した確率水文量の評価

信州大学工学部　学生員○草刈智一　信州大学工学部　正会員　寒川典昭
信州大学工学部　正会員　荒木正夫　信州大学工学部　　　　　高橋浩一

1. はじめに

大きなリターンピリオドに対する確率水文量はかなり不安定であることが，従来から経験的に知られている。我々[1),2),3)]はこれまで，その原因の一つであるデータの特性による不確定さについて，ベイズ論的立場から得られた事後確率分布のエントロピーを用いて評価してきた。

本稿では，上述の研究を基礎として，データのもつ情報量と，そのデータから得られる確率水文量の変動性をデータ数の増加と関連づけて考えるとともに，事後確率分布の危険側の許容範囲の上限で母数を推定することにより，情報量と信頼性を付与した確率水文量を提案する。なお，ここでは母集団としてガンベル分布を使用し，それぞれ一方の母数のみが未知の場合について考察する。

2. データのもつ情報量

未知とした母数を確率変数とし，得られた事後確率分布からエントロピーを計算するとともに，データのもつ情報量を求める。

1) 尺度母数ａを未知，位置母数ｂを既知とした場合

ａとは無関係に一定で既知のｂをもって，ｎ個の確率変数$\tilde{X}(n)$が互いに独立に(1)式のガンベル分布に従うものとする。このとき，確率変数ａの事前確率分布に(2)式の一様分布を仮定すると，$\tilde{X}(n)=x(n)$が与えられた後の事後確率分布は(3)式となる。従って，ａの事後確率分布のエントロピーは(4)式で表される。

2) ａを既知，ｂを未知とした場合

1)と同様に，確率変数ｂの事前確率分布に(5)式の一様分布を仮定すると，$\tilde{X}(n)=x(n)$が与えられた後のｂの事後確率分布は(6)式となる。従って，ｂの事後確率分布のエントロピーは(7)式で表される。

更に，1)及び 2)の場合においてｎ個のデータを得た後の情報量$I(x(n))$は，事前の不確定さを$H(a\ \mathrm{or}\ b \mid x(0))$，事後の不確定さを$H(a\ \mathrm{or}\ b \mid x(n))$とすると(8)式から計算される。

3. 確率水文量の信頼性

ａは小さい程，ｂは大きい程，危険側の確率水文量を与える。従って，ａを未知とした場合は事後確率分布の非超過確率，ｂを未知とした場合は超過確率がそれぞれｔになるように母数を与えた。この母数に対して，(9)式を用いてＴ年確率水文量を算定すると，その信頼性は100×(1-t)%であると考えられる。

4. 実データへの適用と考察

実データとして長野の年最大日降水量を用いた。また，データ数は観測された年代の古い順にｎ＝10,20,・・・,Nとした。ここで，全データを用いてｔに対応するＴ年確率水文量が算定されたとき，筆者らは「情報量がＩ(nat)のＴ年確率水文量は100×(1-t)%の信頼性をもってｘ(mm/day)以内におさまる」という表現を用いることを提案する。

1) ａを未知，ｂを既知とした場合

既知としたｂは全データを用いて積率法により求め，51.4とした。また，事前情報であるa_1，a_2には，0.02，0.15を与えた。図－１は，データ数を10個ずつ増加させて，情報量と90%の信頼性をもつ確率水文量の変動を示したものである。この図から情報量が1.5(nat)程度以降から安定した確率水文量を与えていることがわかる。また，上述の表現を用いれば「情報量が1.71(nat)の100年確率水文量は90%の信頼性をもって

123.5(mm/day)以内におさまる」と言える。

2)　a を既知，b を未知とした場合

既知とした a 及び，事前情報 b_1，b_2 は 1)と同様の方法で求め，それぞれ 0.062，40.0，80.0を与えた。図－2は，データ数を10個ずつ増加させて，情報量と90%の信頼性をもつ確率水文量の変動を示したものである。この図から全体的に確率水文量は安定しているが，特に情報量が1.3(nat)程度以降からほとんど確率水文量に変動がみられないことがわかる。更に，1)の場合と同様「情報量が1.82(nat)の100年確率水文量は90%の信頼性をもって118.6(mm/day)以内におさまる」と言える。また，1)の場合より 2)の場合の方が一貫して安定した値を示しているため，ガンベル分布から計算される100年確率水文量は a に依存するところが大きいと推察される。

表－1　式の一覧

$$f(x)= a\exp\{-a(x-b)-e^{-a(x-b)}\} \tag{1}$$

$$f(a)= \frac{1}{a_2-a_1} \tag{2}$$

$$f(a\mid x(n))= K_a\, a^n \exp\{aA-\sum_{i=1}^{n} e^{-a(x_i-b)}\} \tag{3}$$

ここで，$K_a= \dfrac{1}{\int_{a_1}^{a_2} a^n \exp\{aA-\sum_{i=1}^{n} e^{-a(x_i-b)}\}\, da}$，

$A= -n(\bar{x}-b)$　とする。

$$\begin{aligned} H(a\mid x(n)) &= -\int_{a_1}^{a_2} f(a\mid x(n))\cdot \ln f(a\mid x(n))\, da \\ &= -\ln K_a \\ &\quad - K_a n \int_{a_1}^{a_2} a^n \ln a \exp\{aA-\sum_{i=1}^{n} e^{-a(x_i-b)}\}\, da \\ &\quad - K_a A \int_{a_1}^{a_2} a^{n+1} \exp\{aA-\sum_{i=1}^{n} e^{-a(x_i-b)}\}\, da \\ &\quad + K_a \int_{a_1}^{a_2} \sum_{i=1}^{n} e^{-a(x_i-b)} \exp\{aA-\sum_{i=1}^{n} e^{-a(x_i-b)}\}\, da \end{aligned} \tag{4}$$

$$f(b)= \frac{1}{b_2-b_1} \tag{5}$$

$$f(b\mid x(n))= K_b \exp\{anb-Be^{ab}\} \tag{6}$$

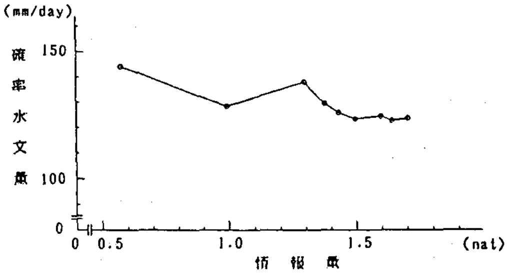
ここで，$K_b= \dfrac{1}{\int_{b_1}^{b_2} \exp(abn-Be^{ab})\, db}$，

$B= \sum_{i=1}^{n} e^{-ax_i}$　とする。

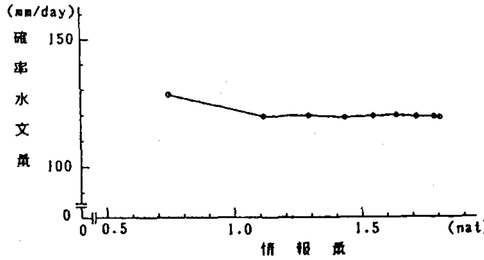
$$\begin{aligned} H(b\mid x(n)) &= -\int_{b_1}^{b_2} f(b\mid x(n))\cdot \ln f(b\mid x(n))\, db \\ &= -\ln K_b \\ &\quad - K_b n \int_{b_1}^{b_2} ab \exp\{anb-Be^{ab}\}\, db \\ &\quad + K_b B \int_{b_1}^{b_2} \exp\{ab(n+1)-Be^{ab}\}\, db \end{aligned} \tag{7}$$

$$I(x(n))=H(a \text{ or } b\mid x(0))-H(a \text{ or } b\mid x(n)) \tag{8}$$

$$X_{100}= b+4.6/a \tag{9}$$



図－1　情報量と確率水文量（a：未知，b：既知）



図－2　情報量と確率水文量（a：既知，b：未知）

5.　おわりに

本稿では，データのもつ情報量とそのデータから得られる確率水文量の変動性を，データ数の増加と関連づけて議論した。更に，設定した超過あるいは非超過確率に対応する母数の値を事後確率分布から与え，情報量と信頼性を付与した確率水文量を提案した。今後は，この議論を，a，b を同時に未知とした場合へ拡張したい。

1)荒木，寒川，上原，草刈：非正規母集団の推定母数の信頼性（その３），61年度中部支部講演集，1987年。
2)草刈，荒木，寒川，上原：ガンベル分布の推定母数の信頼性，第42回土木学会年講，1987年。
3)荒木，寒川，草刈，田辺：推定母数の信頼性評価の諸特性，62年度中部支部講演集，1988年。