

II-8

多項分布モデルによる多地点間日降水量頻度分布の同一性の検討

岐阜工業高等専門学校 正員 鈴木正人 名古屋工業大学 正員 長尾正志

1. はじめに

著者らは、水文量の統計的取扱いにおいて仮定されてきた定常性・均質性の前提を見直すことを目的として、一地点における各年の日降水量頻度分布を多項分布でモデル化し、その同一性をAICにより判定するとともに、同一性が認められるいくつかの集団にグループ分けする手法を提案してきた。本研究はこの手法を、異なる地点間における日降水量の頻度分布の同一性の検討に適用したものである。

2. 多項分布モデルによる日降水量頻度分布の同一性判定手法

複数の頻度分布が与えられた場合、多項分布モデルを基礎にそれらの頻度分布の同一性をAICにより判定する手法が提案されている¹⁾。これを利用して一地点における年単位の頻度分布を対象とした適用手法についてはすでに報告している²⁾ので、ここでは、多地点間の頻度分布の同一性の判定手法について概述する。

以下の記号を用い各地点の日降水量頻度分布を多項分布でモデル化して、その同一性を判定する。

$n(i_1, i_2)$: 各地点毎の水文量の頻度分布 (i_1 : 地点, i_2 : 頻度分布の階級, $i_1 = 1, 2, \dots, c_1$; $i_2 = 1, 2, \dots, c_2$)

$n(i_1)$: 地点 i_1 の資料数, $p(i_2 | i_1)$: 地点 i_1 で i_2 階級の水文量が生起する確率

いま、ある資料 $\{n(i_1, i_2)\}$ が得られる確率 ($= P[\{n(i_1, i_2)\} | \{p(i_2 | i_1)\}]$) は、 c_2 個の項を持つ、 c_1 個の多項分布の積で与えられる。そこで $p(i_2 | i_1)$ をパラメータとみなし、パラメータに依存しない項を除いたときの対数尤度は次式で表現される。

$$L[p(i_2 | i_1)] = \sum_{i_1} \sum_{i_2} n(i_1, i_2) \ln p(i_2 | i_1) \quad (1)$$

パラメータ $p(i_2 | i_1)$ の与えかたによりモデルを表現し、分布の同一性の判定を行う。すなわち、各地点が M 個 ($1 \leq M \leq c_1$) の分布に従うとするモデル MODEL (M) のパラメータをつぎのように表現する。

MODEL (M) : $p(i_2 | i_1) = \theta(i_2 | m_j)$ ($j=1, 2, \dots, M$) 自由パラメータ数: $M \times (c_2 - 1)$ 個

$$\theta(i_2 | m_j) = \frac{\sum_{i_1 \in m_j} n(i_1, i_2)}{\sum_{i_1 \in m_j} n(i_1)}, \quad m_j: \text{分割されたグループ} \quad (2)$$

上式のMODEL (M) で $M=1$ とすると、各地点が同一の分布に従うモデル、 $M=c_1$ とすると各地点がそれぞれ異なる分布に従うモデルになる。各モデルの対数尤度と自由パラメータ数よりAICを計算しAICが最小なモデルを採択することで、分布の同一性の判定とともに適したグループ分けが得られる。

3. 適用計算

木曽川水系の笠置（岐阜県恵那郡、標高220m）、三川（岐阜県加茂郡、標高220m）、上麻生（岐阜県加茂郡、標高132m）の3地点の1958年～1990年までの33年間の日降水量を対象とした。3地点の位置関係の概略を図-1に示す。笠置は木曽川沿いに、上麻生は木曽川の支流飛騨川沿いに、三川は飛騨川の支流赤川沿いに位置しており、水平距離はそれぞれ、笠置・三川間:16km、笠置・上麻生間:23km、上麻生・三川間:8kmである。各年を一括した単位とし、日降水量（mm表示）の平均、分散、最大値（年最大日降水量）を求め、資料年数にわたり平均を取ったものを表-1に示す。基礎統計量の平均値をみると、平均、分散ともに、上麻生と三川の値が近く両地点の距離は近いこともあり降水特性が類似していることが推察される。

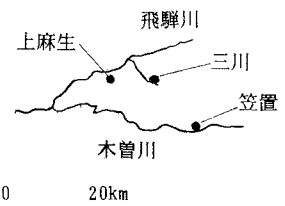


図-1 対象地点の概念図

表-1 各年基礎統計量の平均値

	平均	分散	最大
上麻生	5.90	238.12	142.59
笠置	5.34	192.93	123.05
三川	5.98	227.44	136.06

3. 1 各地点における各年頻度分布の同一性の検討 地点間の頻度分布の同一性を検討する前に、各地点において、各年の頻度分布

の同一性の検討、並びにグループ分けをする。方法は、各地点ごとに33年分の各年頻度分布を求め、平均、分散、最大の統計量を指標に昇順で順位づけし、その順位に従って1～n位、n+1位～33位（n=1～32）というように2分割、3分割を繰り返し、各分割のなかでA I Cが最小な分割方法を求めることで適したグループ分けを行う。なお、階級数は全資料を一括して取り扱うモデルを前提として、スタージェスの式より16個、階級幅は25mmとした。各指標ごとの分割結果を表-2に示す。上麻生では指標によるA I Cの差が僅かで、特に平均と最大を指標とした場合のA I Cの差は0.08しかない。これは、両指標による順位付けの結果がほぼ類似していることから、どちらも同程度に良いモデルと考えられる。他の地点では、平均を指標とした場合と他のモデルのA I Cの差が1以上あることから、平均を指標とした2分割が適当なモデルと判定できる。

3. 2 地点間の頻度分布の同一性の検討 上述の手段により得られた各地点において分割された頻度分布を要素として、地点間の頻度分布

の同一性を検討する。具体的には、3地点とも平均を指標とした2分割が適当であったので、 $3 \times 2 = 6$ つの頻度分布を2つのグループに分けたモデルを考え、各モデルのA I Cを計算する。表-3に主なグループ分けとそのA I Cを示す。表中で例

えば「上小」というのは上麻生において2分割されたうち平均が小さいグループの頻度分布を意味する。また、仮定したモデルのA I Cの順位をA I Cの下に併記してある。A I Cが最も小さいのは、3地点をまとめて平均値の大きいグループと、小さいグループに分けたモデルAであり、同じグループに属する頻度分布は同一と考えられる。これは、分布形状では地点間の相違よりも各年の間の平均値の大小による相違の方が大きいことを意味する。また、モデルAに続いてA I Cが小さいモデルB、C、Dは、二地点の平均値が小さい頻度分布と他の分布とのグループ分けとなっているので、平均値が小さい頻度分布が他の分布と比べて形状が異なる度合いが大きいと思われる。なお、全ての頻度分布を同一としたモデルFのA I Cは、一つの地点と他の二地点をグループ分けしたモデルG、H、IのA I Cより小さいことから、地点間の頻度分布形状の同一性は容認されるものと判断される。

4. おわりに

今回の計算では、対象地点間の相互距離が近いこともあり地点間での頻度分布形状の同一性が認められた。今後は適用地域をひろげ、より多くの地点について適用していきたい。

最後に、建設省木曽川上流工事事務所には、貴重な資料を御提供いただいた。また、資料整理、計算等において前岐阜工業高等専門学校生、藤澤幸広君の助力を得た。ここに記し謝意を表する。

参考文献 1) 坂元慶行・石黒真木夫・北川源四郎：情報量統計学、共立出版、1983、pp. 74～77 2) 鈴木正人・長尾正志：多項分布モデルによる日降水量頻度分布の同一性の検討とグループ分け、水工学論文集37巻、1993、pp. 33～38

表-2 各指標による最適分割結果とA I C

	平均	分散	最大	1分割
上麻生	1～17, 18～33位 8053.79	1～24, 25～33位 8054.85	1～17, 18～33位 8053.87	1～33位 8063.32
笠置	1～13, 14～33位 7234.23	1～19, 20～33位 7239.19	1～19, 20～33位 7239.59	7251.51
三川	1～15, 16～33位 8329.14	1～20, 21～33位 8342.20	1～28, 29～33位 8334.19	8335.91

※上段：分割結果、下段：A I C

表-3 各地点での最適分割を要素としたモデルのA I C

	グループ1	グループ2	A I C
A	上大 笠大 三大	上小 笠小 三小	23546.72 (1位)
B	上大 笠大 三大, 三小	上小 笠小	23575.36 (2位)
C	上大 笠大, 笠小 三大	上小 三小	23583.58 (3位)
D	上大, 上小 笠大 三大	笠小 三小	23585.02 (4位)
E	上大, 上小 笠大 三大, 三小	笠小	23599.32 (5位)
F	上大, 上小 笠大, 笠小 三大, 三小		23627.24 (10位)
G	上大, 上小 三大, 三小	笠大, 笠小	23634.12 (14位)
H	上大, 上小 笠大, 笠小	三大, 三小	23635.04 (15位)
I	笠大, 笠小 三大, 三小	上大, 上小	23644.80 (17位)