

II-7 名古屋における1986年後半の小雨の異常値解析

名古屋工業大学 正 員 ○ 長尾正志、 学生員 鈴木正人

1. 研究の概要

1986年後半、夏以降の降水量は全国的に少なく、とくに中部地方では例年にない小雨傾向となった。最も顕著な小雨となったのは8～11月の4カ月間で、たとえば、愛知用水の水源牧尾ダムでは11月20日に貯水率0を記録し、常識的に水が潤沢だと考えられていた中部地方にも深刻な渇水が発生する危険性が示唆された。本研究は、降水量資料が長期にわたり信頼性の高い名古屋地方気象台の資料によって、その小雨の出現の可能性の程度と異常値検定について、ベイズモデルにより解析した結果を述べる。なお、その際正規変量という仮定を用いるので、岩井法による対数正規変換を用いる。

2. 1986年度の名古屋小雨の確率年

名古屋の1891～1986年の96年間の月別降水量を基礎資料とした。これをもとに、種々の連続した期間量を1986年値と観測期間平均の比で、比較したが、その比が8～11月で最小0.198となったので、以後この4カ月累加量を対象とする。各年値への、角屋の岩井改良法による対数正規分布の適用¹⁾結果が図-1である。全般には、理論曲線に適合しているが、最小値170mm(1986)、第2位230mm(1984)が量の少ない方へ離れている。

そこで、ある降水量を下回る確率の逆数を確率年として算出すると、以上の最小値、第2位値は、それぞれ確率年2,036、254年になり、かなり、異質なデータのようなものである。なお、8～11月内部については、必ずしも最渇水年でない。表-1は月別の小雨順位、生起した年、降水量であるが、このような渇水状態が4カ月間連続して出現した結果として、極端な水不足の事態を惹起したといえよう。

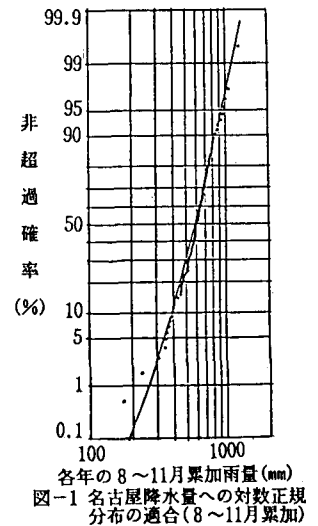


表-1 名古屋月別降水量の小雨順位、生起年および降水量値(mm)

順位	8月	9月	10月	11月	8月～11月計
1	1952 1	1962 31	1891 12	1971 10	1986* 170
2	1965 4	1940 47	1932 24	1973 10	1984 230
3	1986* 14	1922 52	1898 39	1921 11	1926 306
4	1892 23	1905 69	1977 40	1983 16	1963 340
5	1903 27	1980 69	1984 43	1947 18	1947 345
6	1909 32	1984 74	1986* 44	1943 20	1922 355
7	1897 33	1986* 80	1982 46	1926 21	1962 367
8	1961 36	1914 90	1985 51	1953 23	1961 371
9	1923 38	1943 91	1957 53	1908 24	1940 383
10	1947 45	1895 92	1911 54	1899 31	1943 389
11	1906 48	1974 101	1951 55	1986* 32	1985 393
12	1926 49	1968 102	1972 57	1905 34	1914 399

3. 異常値の検定

3.1 異常値解析へのベイズ理論の応用 以上の小雨記録を異常値とみるべきか、他のデータと同質と考えるべきかを異常値理論を用いて解析する。異常値解析は、データの棄却検定と関連して発展してきた。水文統計への応用には、角屋によってThompsonの棄却限界論を準用する方法が提案されているが、ここではより簡便で汎用性があると考えられる北川の提案したベイズ理論とAIC基準による解析法²⁾を適用した。

3.2 基礎仮定 標本総数nのうちにk個の異常値があるとする。異常値と正常値は分散が共通で、平均のみが異なる正規分布に従うとする。このことを、記号的に、正常値の集合を $I = \{i(1), i(2), \dots, i(n-k)\}$ 、異常値の集合を $J = \{j(1), j(2), \dots, j(k)\}$ とし、次式で表記する。

$$x_i \in N(\mu_0, \sigma^2) \quad (i \in I) \quad \cdots \cdots (1) \quad x_j \in N(\mu_1, \sigma^2) \quad (j \in J) \quad \cdots \cdots (2)$$

ここに、 $N(\mu, \sigma^2)$ は平均 μ 、分散 σ^2 の正規分布を意味する。ただし、 k が2つ以上の場合 μ_i はあい異なるものとする。標本 x が、上記の分布に従うとき、母集団から抽出される n 個の独立観測値の同時分布は、未知パラメータを総合的に θ で表現して、次式で与えられる。

$$f(x|\theta) = (2\pi\sigma^2)^{-n/2} [\exp \{-(2\sigma^2)^{-1} \sum_{i \in I} (x_i - \mu_0)^2 + \sum_{j \in J} (x_j - \mu_j)^2\}] \quad \text{---(3)}$$

なお、 $\theta = (\mu_0, \mu_j (j \in J), \sigma^2)$ である。したがって、 θ についての最尤解は、対数尤度 $L(\theta) = \log f(x|\theta)$ を θ の任意の要素で偏微分して0とおくことによって、次式で求められる。

$$\hat{\mu}_0 = \sum_{i \in I} x_i / (n - k) \quad \text{---(4)}, \quad \hat{\mu}_j = x_j \quad (j \in J) \quad \text{---(5)}, \quad \hat{\sigma}^2 = \sum_{i \in I} (x_i - \mu_0)^2 / n \quad \text{---(6)}$$

これを(3)式に代入した最大対数尤度 $L(\theta)$ より、赤池の情報量基準AICをえるが、ここではそれを小標本にも適用しうるように修正したAIC'と、それに対応した条件付最大対数尤度 $L(x|j)$ を示す。

$$AIC' = n \log 2\pi \hat{\sigma}^2 + n + \{2n/(n - k + 3)\} (k + 2) \quad \text{---(7)}, \quad L(x|j) = \exp(-AIC'/2) \quad \text{---(8)}$$

3.3 事前確率と事後確率

異常値の個数についての事前情報はなしとし、異常値の個数についての事前確率に、等確率 $p(0) = p(1) = \dots = \alpha$ を仮定する。つぎに、 k 個の異常値がある場合、 k 個の組み合わせ方が ${}_nC_k$ 、その各々に対して k 個の異常値の分布の割り当て方が $k!$ 通りある。したがって、 k 個の異常値を含む推定のモデルの事前確率は、 $p(J) = \alpha / {}_nC_k \cdot k! \propto ({}_nC_k \cdot k!)^{-1} \quad \text{---(9)}$

事前確率が判ると、(8)式のモデル尤度を用いて事後確率が $p(J|x) \propto L(J|x) \cdot p(J)$ と表されるから、結局、事後確率は次式で計算できる。

$$p(J|x) = [L(J|x) \cdot p(J)] / \sum_{j \in J} [L(j|x) \cdot p(j)] \quad \text{---(10)}$$

3.4 適用計算

表-2 事後確率が0.001以上となる異常値の組み合わせ
(異常値の最大個数は5以下と仮定)

先述のデータについて、異常値が精々5個以下とし、対数正規変量とした解析を行った。まず、事後確率が0.01以上となる異常値の組み合わせを表-2に示す。1986年値のみを異常値とするモデルの事後確率が0.386と最大である。なお、異常値なしとするモデル、1986年と1984年の組み合わせをモデルとする事後確率はそれぞれ0.297、0.109である。

つぎに、ある観測値を異常値に含むすべての組み合わせについての事前確率の和として、周辺事前確率を計算した結果が表-3である。これより、1986年の170mmの確率は0.649と群を抜いて大きく、この観測値の異常さがよく反映されている。また、最大値1243mm、第2位最小値230mmの確率は、0.202、0.158でこの程度であれば、正規分布から多少離れている程度といえる。

結局、1986年後半の小雨は、他とは異質なデータといえそうである。従来から、中部地方は水資源的には他地域に比して、かなり恵まれた条件にあり、それほど深刻な水不足はないものといわれてきたが、今回のような異常な渇水状態も起こりうるわけであるから、そうした異常渇水事態への対応も勘案しておくべきだということが示唆されたといえよう。

順位	事後確率	異常値の組み合わせ
1	0.386	170
2	0.297	異常値なし
3	0.109	170 230
4	0.100	170 1,243
5	0.051	1,243
6	0.038	170 230 1,243
7	0.005	170 230 306
8	0.005	170 1,049 1,243
9	0.002	170 230 1,049 1,243
10	0.002	170 230 306 1,243
11	0.002	1,049 1,243

表-3 0.001以上となる周辺事後確率に対応する観測値

順位	周辺事後確率	観測値
1	0.649	170 (Min 1)
2	0.202	1,243 (Max 1)
3	0.158	230 (Min 2)
4	0.011	1,049 (Max 2)
5	0.008	306 (Min 3)

参考文献 1) 角屋陸：対数正規分布の適用範囲・定数について、農土研別刷3号，1962，p.16

2) 北川源四郎：異常値解析ベイズモデル，数理科学，No.213，1981.3，サイエンス社，pp.62-66