

確率限界法検定に基づく極値水文量の統計的予測に関する研究

中央大学 学生会員 ○清水 啓太
中央大学 フェロー会員 山田 正

1. はじめに

河川・水文分野における極値統計学は主に降水量や河川流量といった水文量の年最大・最小値(極値水文量)が従う極値分布を推定する際に用いられる。また、一般に推定した極値分布を用いて、ある確率年に相当する極値水文量(以下、確率水文量とする)を求めることを水文頻度解析という。ここで確率年とはある基準値を超える値が出現してから、再びその基準値を超える観測値が出現するまでの平均期間のことをいう。この基準値が確率水文量である。水文頻度解析には主に2つのステップがある。これを次に示す。1)対象とする極値水文量の観測値を使用し、経験分布を作成する。ここで、経験分布とは観測値を小さい順に並び替え、それぞれの観測値に応じた累積確率を与えることで定まる階段関数である。2)経験分布に適合する確率分布を推定する。しかし、推定した確率分布は経験分布に完全には適合しない。また、2つの分布の差は非常に大きくなることがある。すなわち、確率年が同じであっても、推定値と観測値が大幅にずれることがある。また、記録的豪雨の観測値と確率分布による推定値の差は非常に大きくなることがあり、このような値は、しばしば想定外であるとみなされてきた。しかし、近年の洪水や土砂災害の頻発、地球温暖化に対する適応策を考えなければならない現状において、想定外とされる規模の豪雨を予測することは重要となる。したがって、本研究では想定外とみなされるような豪雨を統計的手法に基づき予測する理論的枠組みを示すことを目的としている。

2. 研究手法

極値水文量の統計的予測を行うために本研究では水文頻度解析に予測区間を導入し、従来は決定論的に定められてきた確率水文量を「極値分布が示す値を平均値とし、分布をする量」という概念に拡張した。ここで、予測区間とは新たに観測や実験を行った際、予測される値の範囲のことである。本研究では予測区間を森口¹⁾が提案した確率限界法検定に基づき導出した。

予測区間と仮説検定には極めて強い関係性が存在する。確率分布に対する仮説検定を例として挙げ、この関係性を以下に示す。両側確率5%で仮定した確率分布に対する仮説検定を行う際、次のような手順が存在す

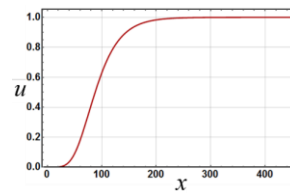


図-1(a) ある確率分布の累積分布関数

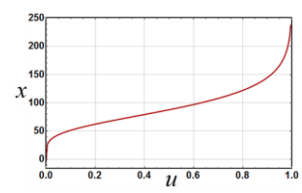


図-1(b) ある確率分布の表現関数

る。1)仮定した確率分布に関する予測区間を作成する。ここで、両側確率5%で検定を行う際、対応する予測区間を95%予測区間という。2)95%予測区間の外部に新たな観測値が存在する場合、仮定した確率分布は有意であるとみなされる。対して、予測区間の内部に新たな観測値が存在するならば仮定した確率分布は棄却されない。

確率限界法検定の検定手法及び特徴を以下に示す。確率限界法検定では仮定した確率分布から求めた確率表現関数の両側に確率限界線を作成し、両線により定まる値の範囲を外れる値が得られた際、有意な差があると判定する。また、特徴としては1)検定を行う際、確率表現関数(以下、表現関数とする)を用いること、2)仮定した確率分布の両端部に関して検定力が特に高いこと(検定力とは誤った仮説を棄却する確率のことである)、3)ノンパラメトリック検定であることなどが挙げられる。表現関数とは累積分布関数の逆関数であると定義されている¹⁾。表現関数は確率限界法検定において、予測区間の導出を簡単化する役割を持つ。ここで、累積分布関数の関数型を(1)式に、表現関数の関数型を(2)式に示す。

$$F(x) = u \quad (1)$$

$$\chi(u) = x \quad (2)$$

ここに、 $\chi(u)$:表現関数、 x :確率変数の実現値、 $F(x)$:累積分布関数、 u :累積確率である。また、これらの関数の関数型を図-1(a)及び図-1(b)に示す。

次に、予測区間の導出に関して、確率限界法検定を採用するに至った理由を以下に示す。1)確率限界法検定に基づく予測区間は極めて大きな極値水文量に対して幅が狭くなるため。ここで、予測区間の幅が狭くなると、その分、予測の精度は向上する。この性質は確率限界法検定が仮定した確率分布の両端部に対して検定力

キーワード：確率限界法検定、確率表現関数、統計的予測

連絡先 〒112-8551 東京都文京区春日 1-13-27 中央大学理工学部 TEL:03-3817-1805

E-mail:shimizu@civil.chuo-u.ac.jp

が高いため成立する。2)確率限界線の概念を応用することで予測区間を導出することができるため。

確率限界法検定に基づき導出した 95%予測区間を図-2 に示す。この図から、両側確率 5%で確率限界法検定を行うとき、95%予測区間の外部に新たに観測値がプロットされた場合、仮定した Gumbel 型極値分布は有意であるとみなされる。

3. 予測区間を導入した水文頻度解析の概要

95%予測区間を記入した確率紙を図-3 に示す。この図から、将来、観測される極値水文量（本研究では年最大日降水量）のうち 95%は図示した予測区間の内部にプロットされることがわかる。また、確率年が大きくなるほど予測区間が広がっており、ある確率年に着目すると、確率水文量が取りうる値の範囲を知ることができる。例えば、95%予測区間を考慮すると 100 年確率水文量は約 216.2 [mm/day]を平均値にもち、およそ 189.1 [mm/day]から 319.5 [mm/day]までの範囲に分布していることがわかる。すなわち、当該地域において、100 年確率水文量は 100 [mm/day]程度の幅を持つことが理論的に示された。このように、確率水文量の値に幅をもたせることで設計基準等を考える際、過去の観測値から推定した確率分布が示す値に加えて、予測区間の理論直線が示す値も考慮に入れることができるようになった。

今日の水文頻度解析では、推定した確率分布を大きく外れる観測値が新たに得られた際、その値に合うような確率分布を再度推定するということが行われる場合がある。しかし、予測区間を導入することにより、そのような値は十分に起こりうることでありと解釈できるようになった。例を挙げると、300[mm/day]の年最大日降水量が新たに生起したと仮定する。このとき Gumbel 型極値分布から確率年を推定すると 1000 年以上となるが、予測区間を考慮することで当該値の確率年は 100 年程度であることがわかる。すなわち、当該値は 100 年確率程度の年最大日降水量のうち、発生頻度が平均値よりも低い値であると考えることができる。

次に、観測値の総数と予測区間の関係を示す。図-4(a)は石狩川水系幾寅観測所における 36 年間分の年最大日降水量[mm/day]を用いて導出した予測区間を記入した確率紙である。対して、図-4(b)は地上観測地点（東京）における 142 年間分の年最大日降水量を用いて導出した予測区間を記入した確率紙である。両図には 50%, 60%, 70%, 80%, 90%, 95%, 99%予測区間を記入した。両図から観測値の総数が多いほど、予測区間が狭まり、将来の極値水文量に対する予測の精度が向

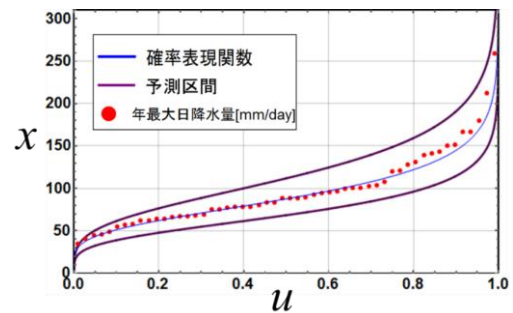


図-2 利根川水系八斗島観測所における 54 年間分の年最大日降水量 [mm/day] 及び、これらの値から推定した Gumbel 型極値分布の表現関数、95%予測区間

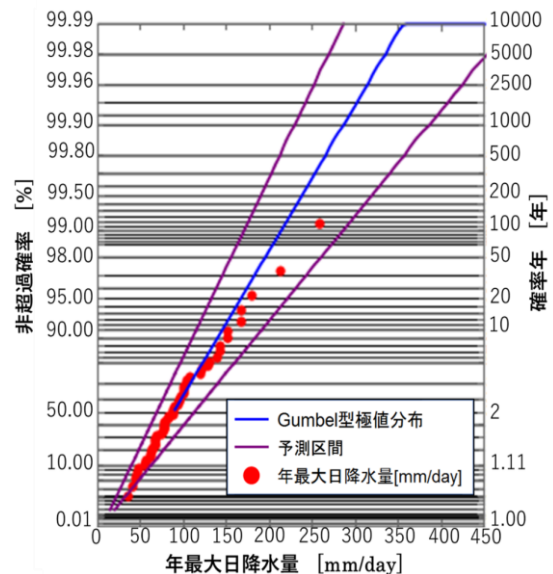


図-3 利根川水系八斗島観測所における 54 年間分の年最大日降水量 [mm/day] 及びこれらの値から推定した Gumbel 型極値分布、95%予測区間を記入した確率紙

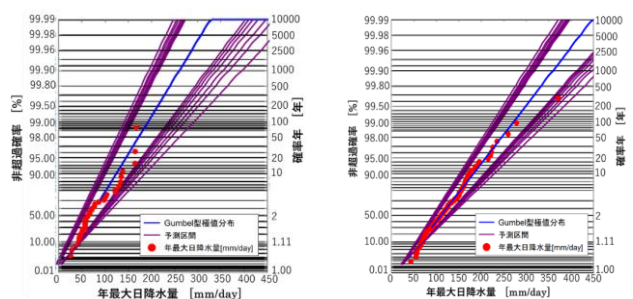


図-4(a) 36 個の観測値から作成した予測区間
図-4(b) 142 個の観測値から作成した予測区間
上することが示された。

4. まとめ

本研究では水文頻度解析に確率限界法検定に基づく予測区間を導入することで、確率水文量の概念を拡張し、極値分布から大きく外れる水文量を評価する理論的枠組みを示した。

参考文献

- 森口繁一：確率表現関数の検定について—Kolmogorov-Smirnov 検定を考え直す—，日本統計学会，第 25 巻，pp. 233-244，1995