京都大学防災研究所年報　第30号　B-2　昭62. 4
Annuals, Disas. Prev. Res. Inst., Kyoto Univ., No. 30 B-2, 1987

水文統計解析における確率分布モデルの評価

宝　　馨・高棹　琢馬・清水　　章

EVALUATION OF PROBABILITY DISTRIBUTION MODELS IN HYDROLOGIC FREQUENCY ANALYSIS

By *Kaoru* TAKARA, *Takuma* TAKASAO and *Akira* SHIMIZU

Synopsis

The anthors emphasizes the necessity for considering the variability of the estimate of T-year event in hydrologic frequency analysis and proposes a framework for evaluating probability distribution models. The variability (or estimation error) of T-year event is used as a criterion for model evaluation as well as three goodness-of-fit criteria (SLSC, MLL, and AIC) in the framework. The jackknife and the bootstrap methods play important roles in estimating the variability. For the annual maxima of k-day precipitation ($k=1, 2, 3$) in the Lake Biwa basin, the Gumbel distribution is regarded as the best among five distribution models with two parameters; for the daily precipitation at Osaka, the SQRT-exponential-type distribution of maximum is the best among ten distributions with two or three parameters. The bootstrap method also reveals the relationship between the amount of data and the estimation error of T-year event.

1. 緒　　言

種々の水工計画にリターンピリオドあるいは確率水文量の概念が導入されて久しい。当該水文量が従う確率分布を決定（推定）することが計画の基本をなしており，水文統計学の主要なテーマの1つとなっているのである。

さて，当該水文量のデータが与えられたとき，それが従う確率分布を決定するには，大略以下のような手順を踏むのが通例である。

① データの等質性・独立性等に関して，水文学的あるいは確率統計学的観点から吟味する。[**データの吟味**]

② ヒストグラムや分布曲線を描き大体の分布形状を把握したのち，適当と思われる確率分布モデルをいくつか最終モデルの候補として選ぶ。[**候補モデルの列挙**]

③ データにそれらのモデルをあてはめる。この際，何らかの方法で母数（パラメタ）を推定する。[**母数推定**]

④ モデルの良否を何らかの規準により比較検討し，最も良いと思われる分布モデルを選ぶ。[**モデル評価**]

この一連の手順は必ずしも確立されたものではなく，次のような未解決の問題を残している。

i）手順①については，観測方法の改変，現象生起原因の異同，地球規模・流域規模での気象学的あるいは水文学的条件の経年変化といった点に関する吟味が必ずしも十分ではない。厳密にデータを取り扱おうとすれば，データ数が激減したり，厳密に取り扱うための理論がなかったりする。すなわち，データの質と量の問題，データ処理技術の問題であり，これにはまだしばらく時日を要するものと思われる。

ii) 手順②については，これまでに提案され使用されてきた確率分布モデルが果たして十分であるかという問題がある。当該水文量に対して，既存のモデルのどれもが適合しないかも知れない。こうした場合には新たなモデルを開発するか経験分布を採用しなければならない。

iii) 手順③の母数推定には，確率紙を用いた図式推定法と，最尤法・積率法などに代表される解析的方法があり，方法ごとに母数あるいは確率水文量の推定値に大きな差異の生じることがある。どの方法によるべきか明らかでない。

iv) 手順④については，適合させた確率分布モデルの評価規準が確立されていない。従来はデータとモデルの適合度をモデルの評価規準としてきた。図式推定法の場合は，プロットしたデータが直線上に並んでいるかどうか，目視による一致性 (visual consistency) によって評価することが多かった。解析的方法によって母数を推定した場合には，ヒストグラムと確率密度関数のあてはまり具合を，目視による判断あるいはカイ二乗検定や Kolmogorov-Smirnov 検定のような適合度検定手法によって調べたり，データとモデルの誤差の二乗和や，尤度などの適合度規準を用いたりすることもある。このように適合度によりモデルを評価しようとする場合，適合度の優れたモデルがただ1つだけ定まるのであればあまり問題はない。ところが，当該水文量に対して同程度の適合度を示すモデルが複数個存在し，それらが異なる確率水文量を与えることがある。また，規準ごとに最良とみなされるモデルが異なることがある。こうした場合，どのモデルを最終的に採用すればよいのかについて，明確な提示がなされてこなかった。すなわち，適合度はモデル評価の規準として不十分であり，何らかの新たな評価規準が必要である。

本研究では，上記 iv) に焦点をあて，確率水文量の変動性をモデル評価の規準とすることを提案する。ここで筆者らが強調したいのは，データの蓄積が進んでも確率水文量の推定値が大きく変動しないような確率分布モデルが望ましいということである。なぜならば，データの蓄積により確率水文量が大きく変動すると，その都度水工計画の大幅な見直しを要請されることになるからである。

こうした観点に立って，本論文では

（Ⅰ）データとモデルの適合度だけでなく，確率水文量の変動性をも評価規準として，水文統計解析における確率分布モデルの評価の手順を確立すること；

（Ⅱ）データ数と確率水文量の推定精度との関連を定量的に明らかにすること；

を研究目的とした。

確率水文量の変動性，推定精度を調べるために，2つのリサンプリング手法——jackknife 法と bootstrap 法——を適用する。リサンプリング手法とは，簡単に言うと，現在手元にある1組のデータセット（標本）から，部分的にデータを抽出したり，繰り返しを許して元の標本と同じデータ個数だけ抽出したりという操作を反復して，多数個のデータセットを作り出し，元の標本から得られる統計量の偏倚を補正したり，統計量の推定誤差を推定したりする手法である。こうして生成される多数のデータセットの取り扱いは，近年，計算機の著しい発達により容易に実現可能となったので，統計学の分野でも CIS (computer intensive statistics) と呼ばれる一分野を形成している。

次章以後の本論文の構成は以下のようである。まず，**2.** では適合度がモデル評価の規準として必ずしも十分ではないことを例示する。**3.** では，確率水文量の変動性を確率分布モデルの評価規準とすることの意義を述べ，モデル評価の手順を提示する。さらに，jackknife 法と bootstrap 法の概要を述べ，モデル評価手順への適用方法を示す。**4.** では，この評価手順を琵琶湖流域の年最大 k 日降水量 (k=1, 2, 3) と大阪の年最大日降水量に適用した結果を示す。最後に **5.** では，データ年数と確率水文量の変動性との関係を bootstrap 法で調べた結果を述べる。

2. 適合度によるモデル評価の限界

本章では，適合度がモデル評価の規準として必ずしも十分ではないことを示す。**Fig. 1** は，琵琶湖流域

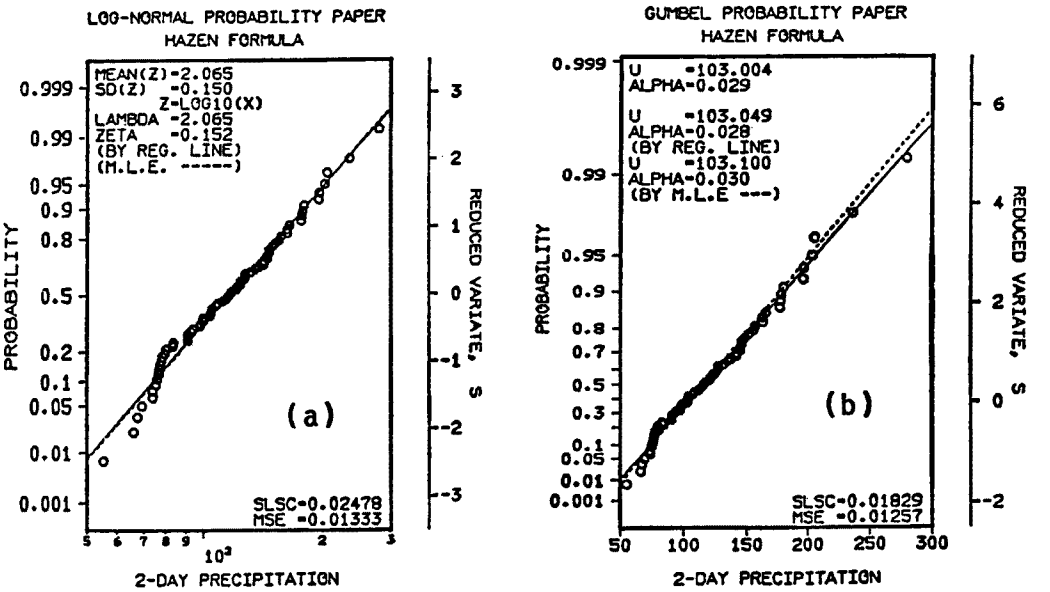

Fig. 1 Probability plotting of the annual maxima of 2-day precipitation (in mm) of the Lake Biwa basin, for the period 1912-1981 (70 years), obtained by the Hazen formula ((a) log-normal distribution and (b) Gumbel distribution).

の年最大2日降水量（1912～1981年の70年間）を，対数正規確率紙と Gumbel 確率紙にプロットしたもの (Hazen プロット）を示している。図中，実線は，図式推定法（標準変量について最小二乗法を適用）であてはめた分布直線，破線は最尤法によるそれである。どちらの分布に対してもプロット点はほぼ直線に並んでおり，適合度は良好である。

筆者らは，次式のような適合度の評価規準 (SLSC, standard least-squares criterion) を提案している[1]。

$$\mathrm{SLSC}=\frac{\sqrt{\xi^2_{min}}}{|s_{0.99}-s_{0.01}|} \quad \cdots\cdots (1)$$

ここに，$s_{0.99}$, $s_{0.01}$：それぞれ非超過確率 0.99, 0.01 に対応する当該分布の標準変量，ξ^2：順序統計量データ y_i に対応する標準変量 s_i と適当な確率 q_i に対応する標準変量 r_i との差の二乗和の平均で，

$$\xi^2=\frac{1}{n}\sum_{i=1}^{n}(s_i-r_i)^2 \quad \cdots\cdots (2)$$

である（n はデータの個数）。確率 q_i はいわゆるプロッティング・ポジション公式として Hazen 公式を用いると，$q_i=(i-0.5)/n$ である。ξ^2_{min} は (2) 式を最小化した値で，上述の図式推定法はこの場合に相当する。

SLSC は ξ^2_{min} を (1) 式の形に標準化したものであり，その値によって，異なる確率分布モデルのデータへの適合度の比較ができる。SLSC の値が小さい程よく適合していることになる。また，SLSC は，適合度の相対的評価の規準としてだけでなく絶対的評価にも有用である。SLSC≒0.02 であれば，**Fig. 1** に示すように良い適合度を示す。SLSC>0.03 であれば他の分布を試みるべきである[1]。すなわち，確率紙にデータをプロットしたり，データのヒストグラムと確率密度関数のグラフを描いたりすることなく，その値を計算するだけで適合度の良否が判定できるという利点をもつ。

Table 1 (a) には，琵琶湖流域年最大 k 日降水量 (k=1, 2, 3) について，上述の図式推定法（標準変量に関する最小二乗法）によって，2母数対数正規分布と Gumbel 分布をあてはめた場合の SLSC 値を示した。SLSC によれば k=1, 2 の場合は Gumbel 分布が，k=3 の場合は2母数対数正規分布の方がよく適合していると言える。

次に，最小二乗法ではなく最尤法によってあてはめることを考えてみよう。変量 X の確率密度関数を

Table 1 Comparison of two-parameter log-normal and Gumbel distributions by SLSC, MLL, and AIC for the annual maxima of k-day precipitation of the Lake Biwa basin, for the period 1912-1981 (70 years)

k	Probability distribution	(a) Least-squares method	(b) Maximum likelihood method		
		SLSC	SLSC	MLL	AIC
1	Log-normal	0.02459	0.02476*	−334.557	673.113
	Gumbel	0.02240*	0.02890	−334.387*	672.774*
2	Log-normal	0.02482	0.02498	−357.342*	718.683*
	Gumbel	0.01829*	0.02136*	−357.549	719.098
3	Log-normal	0.01872*	0.01890*	−366.850*	737.701*
	Gumbel	0.02214	0.02389	−367.031	738.061

* denotes better one.

$f(x;\theta)$ とし (θ: 母数のベクトル)，観測データを $x_1, x_2, \ldots, x_n$ とするとき，

$$L(x_1, x_2, \ldots, x_n; \theta)=\prod_{i=1}^{n} f(x_i; \theta) \quad \cdots\cdots (3)$$

とおき，この $L(x_1, x_2, \ldots, x_n; \theta)$ を θ の関数とみなして尤度関数という。尤度関数を最大にするような母数 $\hat{\theta}$ を求めるのが最尤推定法で，この $\hat{\theta}$ を θ の最尤推定値という。通常は尤度関数の対数をとったもの（対数尤度，log L）を最大化する。こうして得た最大対数尤度（MLL, maximum log-likelihood）は，次式で与えられ，適合度を表わす指標となる。

$$\mathrm{MLL}=\sum_{i=1}^{n} \log f(x_i; \hat{\theta}) \quad \cdots\cdots (4)$$

同じデータに対して，異なる確率分布モデルを最尤法であてはめたとき，MLL の値の大きいモデルが適合度がよいということになる。

さて，一般に，母数の個数が増えると適合度がよくなることに留意しなければならない。すなわち，母数の個数が増えると，SLSC は小さくなっていくし，MLL は大きくなっていく。したがって，SLSC や MLL などをモデルの良否を判定する評価規準とした場合，母数の個数のより多いモデルが“良いモデル”であると評価されることになる。モデルを評価する際には，モデルの簡潔さ（式形が複雑でなく計算が簡単なこと，母数の個数が多すぎないこと）も，適合度と同じく基本的な要件である。そこで，赤池[2]は，母数の個数をも考慮することのできる情報量規準（AIC, Akaike's information criterion）と呼ばれる評価規準を提案している。

$$\mathrm{AIC}=-2\ \log\ (\text{最大尤度})+2p=-2\mathrm{MLL}+2p \quad \cdots\cdots (5)$$

ここに，p は母数の個数である。母数が増えると，(5) 式第2項は大きくなるが，適合度はよくなるので第1項は小さくなる。このトレードオフ関係の中で，AIC を最小とするようなモデルが最も“良い”モデルである，とするのが赤池の考え方である。AIC の誘導過程や応用例は坂元らの成書[3]に詳しい。

先程と同じデータに同じ2つの確率分布モデルを最尤法であてはめた場合の SLSC, MLL, AIC の値を **Table 1 (b)** に示す。ここで，SLSC は，順序統計量 y_i を最尤推定値 $\hat{\theta}$ を用いて変換した標準変量 s_i から (2) 式によって ξ^2 を求め，この ξ^2 を (1) 式の ξ^2_{min} に代入して得たものである。当然のことながら，**Table 1 (a)** の SLSC 値よりも大きな値となっている。特に，k=1 の場合，Gumbel 分布の SLSC 値が相対的に悪くなっている。また，これら2つの分布はともに母数の個数 $p=2$ であるから，この例では AIC は MLL と等価である。**Table 1** より言えることは，適合度評価が規準により異なるということである。

Table 2 Comparison of ten distributions by SLSC, MLL, and AIC for the annual maxima of daily precipitation at Osaka, for the period 1889-1980 (92 years), fitted by the maximum likelihood method

Prob. distribution		SLSC	MLL	AIC
Normal	(2p)	0.07937	−450.151	904.301
Log-normal	(3p)	0.01666①	−432.818①	871.636②
Log-normal	(2p)	0.02996	−434.914	873.828
Pearson Ⅲ	(3p)	0.03765	−432.902②	871.803③
Pearson Ⅲ	(2p)	0.06116	−438.172	880.344
Log-Pearson Ⅲ	(3p)	0.01749②	−432.910③	871.819
SQRT-ET-max	(2p)	0.02423	−433.091	870.182①
Gumbel	(2p)	0.04769	−434.414	872.829
Log-Gumbel	(3p)	0.01859③	−433.167	872.335
Log-Gumbel	(2p)	0.03496	−434.531	873.062

① denotes the best distribution for each criterion;
② and ③ the second and the third, respectively.

最尤法を用いた場合，k=1 では SLSC によれば対数正規分布が，MLL（あるいは AIC）によれば Gumbel 分布がよりよい適合を示している。k=2 では全く逆の結果となっている。また，k=1 において，最小二乗法の場合 SLSC によれば Gumbel 分布がよいが，最尤法の場合に SLSC を適用すれば対数正規分布がよい。このように，同じデータに対して，規準ごとに適合度評価の結果が異なったり，同じ規準でも母数推定法によってやはり結果が異なることが生じるのである。

大阪の年最大日降水量（1889～1980年の92年間）のデータ[4]に，2～3個の母数をもつ10種の分布を最尤法であてはめ，SLSC，MLL，AIC を求めた結果を **Table 2** に示す。取り扱った分布は，正規分布（2母数），対数正規分布（2母数，3母数），Pearson Ⅲ 型分布（2母数，3母数），対数 Pearson Ⅲ 型分布（3母数），平方根指数型最大値分布（2母数，表中では SQRT-ET-max と記した）[5]，Gumbel 分布（2母数），対数 Gumbel 分布（2母数，3母数）である。これらすべての分布モデルに対して最尤法を用い，その求解には改訂準ニュートン法による多変数関数の極小化プログラム DMINF1[6] を利用した。これらのモデルをすべて最適化するのに要する計算時間は，京大大型計算機センターの FACOM M-382 で約6 sec，M-780 で約3 sec であった。今日，最尤法は基本的なサブルーチンを利用して瞬時に実現できる。

Table 2 の中で丸つき数字①②③は，各評価規準について適合度のよい順位を表している。3母数対数正規分布はどの評価規準についてもベスト3に入っており，SLSC，MLL についてはともに第①位，AIC については第②位である。その他の3母数の分布（Pearson Ⅲ 型，対数 Pearson Ⅲ 型，対数 Gumbel）はどれかの規準でベスト3に入っている。2母数の分布では，平方根指数型最大値分布は AIC が第①位であり，SLSC，MLL についてはともに第4位である。SLSC=0.02423 であるから，まずまずの適合度である。他の2母数の分布はどれもベスト4にも入らずかなり適合度が悪い。これらの分布の適合度を，データのヒストグラムとの比較で図示したのが **Fig. 2** である（紙数の都合で，適合度の悪い正規分布，2母数 Pearson Ⅲ 型分布，2母数対数 Gumbel 分布は割愛した）。**Fig. 2** を見ても，いくつかの確率分布モデルの適合度はほぼ同程度に見える。また，**Table 1** と同様，**Table 2** においても，良い評価を得るモデルが規準ごとに異なっている。大津，彦根の極値データについても同様であった[7]。

結局，モデルの良否を適合度によって評価するのは非常にむずかしいと言える。適合度は，“良いモデル”のための必要条件であり，モデル群を screening する（適合度の悪いモデルをふるい落とす）規準として有用であるが，同程度によい適合度を示す2つ以上のモデルを評価するには別の規準が必要となるわけである。

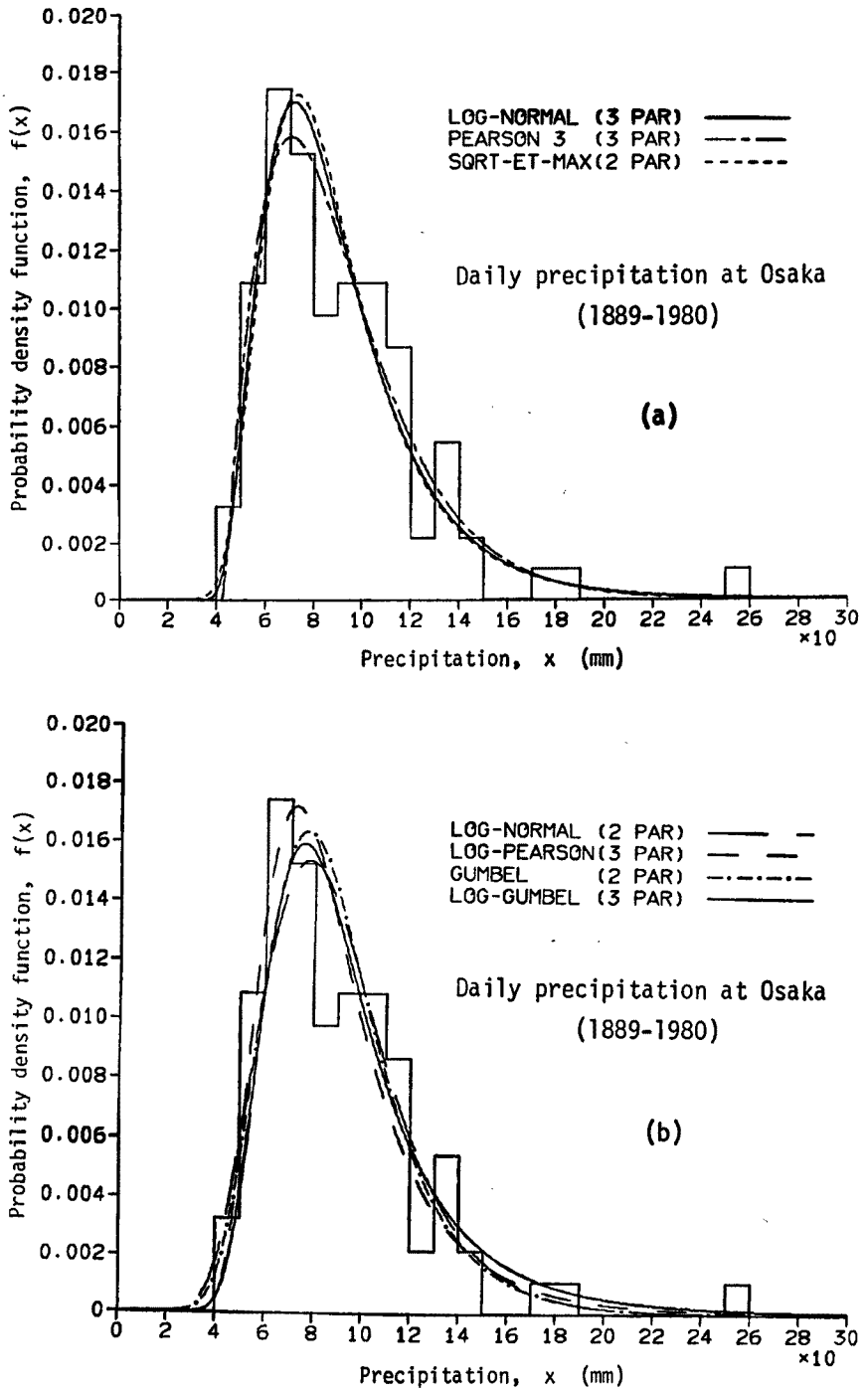

Fig. 2 Distributions fitted by the maximum likelihood method for the annual maxima of daily precipitation (in mm) at Osaka, for the period 1889-1980 (92 years).

1. で述べたような理由によって，筆者らは確率水文量の変動性をモデル評価の規準とするのがよいのではないかと考える。次章では，確率水文量の変動性を規準としたモデル評価の手順を提示し，その適用例を示す。

3. 確率水文量の変動性を規準とした確率分布モデルの評価

3.1 提案するモデル評価の手順

年最大値などの極値水文量を取り扱う場合，極値データの分布全体の適合度も重要であるが，分布の裾の部分の形状や適合度がより重要視される。というのは，分布の裾，すなわち非超過確率の大きい部分（渇水などの場合のように小さな値を対象とするときは，超過確率の大きい部分）のちょっとした形状の違いによって確率水文量の値がかなり異なってくるからであり，実際の種々の水工計画の立案はこの確率水文量の値を基礎としてなされるからである。したがって，1. でも述べたように，データの蓄積が進んでも（言い換えると，データの組み合わせが異なっても）確率水文量の推定値が大きく変動しないような確率分布が望ましい。

そこで，次のようなモデル評価の手順を提案する。

① データの吟味（前述）。
② 候補モデルの列挙（前述；なるべく多くのモデルを挙げる）。
③ 母数推定（n 個のデータ全部を用い，最尤法による）。
④ 適合度の比較評価によるモデル群の screening（SLSC, MLL, AIC などを用いて，適合度の悪いモデルを除外する）。
⑤ リサンプリング手法による確率水文量の変動性の検討（④で除外されなかったモデルすべてに対して，jackknife 法と bootstrap 法を適用する）。
⑥ 最終モデルの決定（確率水文量の変動の最も小さいモデルを選ぶ）。

3.2 リサンプリング手法概説

リサンプリング手法とは，現在手元にあるデータに対して何度もサンプリングを繰り返す（resampling）方法であり，統計量の推定値の偏りの補正，推定誤差の推定などを行うための手法である。本研究では，jackknife 法と bootstrap 法を用いる。これらの方法は，従来の統計的方法よりずっと理解し易くしかも適用範囲が広いと言われており，近年，計算機の発達とともに急速に進歩しつつある。

jackknife 法は，統計量の偏倚を推定・補正するために Quenouille が考案したノンパラメトリックな方法で，1949年に発表された。その後（1958年），Tukey によって偏倚だけでなく分散（すなわち，統計量に含まれる誤差の程度）をも推定できることが指摘された。1974年に，Miller[8] によって jackknife 法に関する詳細なレビューがなされている。

bootstrap 法は，1977年に Efron が発表した方法で，jackknife 法と同様，統計量に含まれる誤差の程度を推定するものである[9]。ただし，偏倚の補正はできず，また，jackknife 法より多数回の繰り返し計算を必要とする。

水文学の分野でのこれらの手法の適用例としては以下のようなものがある。

Bardsley (1977)[10]は，極値データに対して，3つの極値分布のうちどれを選ぶかという問題に jackknife 法を適用している。Tung and Mays (1981)[11]は，流量資料の少ない地点に対して，洪水頻度分析を行う際に，対数 Pearson Ⅲ 型分布で用いる歪係数を，標本歪みと Water Resources Council が作成した歪みマップ（skew map）を用いて得られる歪みの分散から推定する問題に jackknife 法と bootstrap 法を適用している。データ数の多寡にかかわらず，また，マップの精度の良否にかかわらず，jackknife 法が頻度係数を精度よく推定しうることを示している。Cover and Unny (1986)[12]は，年流量時系列に ARMA モデルをあてはめ，パラメタ推定誤差をこれらのリサンプリング手法を用いて議論している。

わが国においては，水文学関係で jackknife 法や bootstrap 法を用いた研究はこれまでにないようであるが，一般のデータ解析の分野で徐々に用いられはじめているようである[13]。

jackknife 法や bootstrap 法だけでなく，その応用や他のリサンプリング手法についても，Efron の成書

に詳しい記述がある[9]。

3.3 Jackknife 法[9]

(1) 偏倚の推定

ある未知の確率分布 F に従う互いに独立な n 個の確率変数を $X_i(i=1, 2, \ldots, n)$ とし，これを

$$X_1, X_2, \ldots, X_n \overset{iid}{\sim} F$$

と記すことにする。$X_1=x_1, X_2=x_2, \ldots, X_n=x_n$ なる観測値を得て，これらの観測値から母集団の性質を表す量 θ（平均値，歪係数，相関，分位値など）の推定値（統計量）$\hat{\theta}$ が計算できる。すなわち，

$$\hat{\theta}=\hat{\theta}(x_1, x_2, \ldots, x_n)$$

θ は分布 F によって決まっているので $\theta=\theta(F)$ と記すと，統計量 $\hat{\theta}$ を求める際には観測値 x_i に何らかの経験分布 $\hat{F}$ をあてはめることになる。観測値 $x_1, x_2, \ldots, x_n$ はそれぞれ $1/n$ の確率で生起したものと考えるのが自然であり，したがって $\hat{F}$ は以下のような離散型一様分布であるとする。

$$\hat{F}: mass\frac{1}{n} at\ x_1, x_2, \ldots, x_n$$

$\hat{F}$ は一般に F と異なるので $\hat{\theta}$ に偏倚が生じることになる。この偏倚の大きさを知りたいのである。そこで，統計量の偏倚 $Bias$ を次式で定義する。

$$Bias\equiv E_F\theta(\hat{F})-\theta(F) \quad \cdots\cdots (6)$$

ここに，E_F は $X_1, X_2, \ldots, X_n \overset{iid}{\sim} F$ における期待値を示す。いま，n 個のデータ $x_i(i=1, 2, \ldots, n)$ から i 番目のデータを除外した $n-1$ 個のデータが，ある経験分布 $\hat{F}_{(i)}$ に従うとして再び計算した統計量を $\hat{\theta}_{(i)}$ とする。すなわち，

$$\begin{aligned}\hat{\theta}_{(i)}&=\theta(\hat{F}_{(i)})\\&=\hat{\theta}(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \quad \cdots\cdots (7)\end{aligned}$$

$\hat{F}_{(i)}$ は，先程と同様

$$\hat{F}_{(i)}: mass\frac{1}{n-1} at\ x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$$

こうして得た n 個の統計量 $\hat{\theta}_{(i)}$ の平均を

$$\hat{\theta}_{(\cdot)}=\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{(i)} \quad \cdots\cdots (8)$$

とする。Quenoulli が定義した偏倚の推定値 $\widehat{BIAS}$ は次式で与えられる。

$$\widehat{BIAS}\equiv(n-1)(\hat{\theta}_{(\cdot)}-\hat{\theta}) \quad \cdots\cdots (9)$$

$\hat{\theta}$ からこの偏倚の推定値 $\widehat{BIAS}$ を取り除いて，それを $\tilde{\theta}$ と記すことにすると，

$$\tilde{\theta}\equiv\hat{\theta}-\widehat{BIAS}=n\hat{\theta}-(n-1)\hat{\theta}_{(\cdot)} \quad \cdots\cdots (10)$$

となる。これを θ の **jackknife 推定値**と呼ぶ。

以下に，(10)式の jackknife 推定値 $\tilde{\theta}$ が偏倚を補正したものとなることを説明する。まず，E_n を標本数 n の場合の $\hat{\theta}$ の期待値と定義する。すなわち，

$$E_n\equiv E_F\hat{\theta}(X_1, X_2, \ldots, X_n) \quad \cdots\cdots (11)$$

いま，E_n の θ からの偏倚が $1/n$ に比例すると仮定すれば，

$$E_n=\theta+a/n \quad \cdots\cdots (12)$$

ただし，a は未知の比例定数である。データ数が $n-1$ の場合は，

$$E_F\hat{\theta}_{(\cdot)}=E_{n-1}=\theta+a/(n-1) \quad \cdots\cdots(13)$$

(12)，(13)式から a を消去して θ を求めると，

$$\begin{aligned}\theta&=nE_n-(n-1)E_{n-1}\\&=nE_F\hat{\theta}-(n-1)E_F\hat{\theta}_{(\cdot)}\\&=E_F\{n\hat{\theta}-(n-1)\hat{\theta}_{(\cdot)}\}\\&=E_F\tilde{\theta} \quad \cdots\cdots(14)\end{aligned}$$

となり，jackknife 推定量 $\tilde{\theta}=n\hat{\theta}-(n-1)\hat{\theta}_{(\cdot)}$ の期待値は θ に一致する。これが偏倚を補正する原理である。

より厳密にいうと，jackknife 推定量は $1/n$ のオーダーの偏倚を除いた推定量である。(12)式の代わりに

$$E_n=\theta+\frac{a_1(F)}{n}+\frac{a_2(F)}{n^2}+\cdots \quad \cdots\cdots(15)$$

（ただし，関数 $a_1(F)$, $a_2(F)$, ... は n に依存しない）とすると，$E_F\hat{\theta}_{(\cdot)}$ は，

$$E_F\hat{\theta}_{(\cdot)}=E_{n-1}=\theta+\frac{a_1(F)}{n-1}+\frac{a_2(F)}{(n-1)^2}+\cdots$$

となるので，$E_F\tilde{\theta}$ は(10)式より，

$$\begin{aligned}E_F\tilde{\theta}&=E_F\{n\hat{\theta}-(n-1)\hat{\theta}_{(\cdot)}\}\\&=nE_F\hat{\theta}-(n-1)E_F\hat{\theta}_{(\cdot)}\\&=nE_n-(n-1)E_{n-1}\\&=\theta-\frac{a_2(F)}{n(n-1)}+a_3(F)\left(\frac{1}{n^2}-\frac{1}{(n-1)^2}\right)+\cdots\end{aligned}$$

となり，この場合，$O(1/n)$ の偏倚が除かれたが厳密には θ と一致しない。すなわち，jackknife 推定量は $O(1/n^2)$ の偏倚をもつ。

(2) 分散の推定

$X_1, X_2, \ldots, X_n \overset{iid}{\sim} F$ なる確率変数から得られる統計量 $\hat{\theta}$ の分散 Var を

$$Var\equiv E_F(\hat{\theta}(X_1, X_2, \ldots, X_n)-E_F\hat{\theta})^2 \quad \cdots\cdots(16)$$

としよう。Tukey は，$\hat{\theta}_{(i)}$ によって Var のノンパラメトリックな推定値 $\widehat{VAR}$ も評価できることを示した。すなわち，

$$\widehat{VAR}=\frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)}-\hat{\theta}_{(\cdot)})^2 \quad \cdots\cdots(17)$$

ただし，$\hat{\theta}_{(\cdot)}=\sum_{i=1}^{n}\hat{\theta}_{(i)}/n$ である。

(3) まとめ

jackknife 法の手順をまとめると以下のようになる。

i) n 個のデータを用いて対象とする統計量 $\hat{\theta}$ を求める。

ii) n 個のデータから，まず1番目のデータを除外した $n-1$ 個のデータを用いて，推定値 $\hat{\theta}_{(1)}$ を求める。次に2番目のデータを除外して $\hat{\theta}_{(2)}$ を求める。この作業を n 番目のデータまで繰り返し，$\hat{\theta}_{(1)}$, $\hat{\theta}_{(2)}$, ..., $\hat{\theta}_{(n)}$ を求める。

iii) 次式により $\hat{\theta}_{(i)}(i=1, \ldots, n)$ の平均 $\hat{\theta}_{(\cdot)}$ を求める。

$$\hat{\theta}_{(\cdot)}=\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{(i)}$$

iv) θ の jackknife 推定値 $\tilde{\theta}$ とその分散 $\widehat{VAR}$ はそれぞれ,

$$\tilde{\theta}=n\hat{\theta}-(n-1)\hat{\theta}_{(\cdot)}$$

$$\widehat{VAR}=\frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)}-\hat{\theta}_{(\cdot)})^2$$

で与えられる。

3.4 Bootstrap 法[9)]

$X_1, X_2, \ldots, X_n \overset{iid}{\sim} F$ なる n 個の確率変数で定義される統計量 $\hat{\theta}(X_1, X_2, \ldots, X_n)$ が与えられたとき, $\hat{\theta}$ の標準偏差 Sd を次のように書くことにしよう。

$$Sd=\sigma(F, n, \hat{\theta})=\sigma(F) \quad \cdots\cdots(18)$$

これは, データ数 n と統計量の形式 $\hat{\theta}(\cdot, \cdot, \cdot, \cdot)$ が与えられれば, $\hat{\theta}$ の標準偏差 Sd は未知の確率分布 F の関数になることを表している。

標準偏差 Sd の bootstrap 推定値 $\widehat{SD}$ は, 単に $F=\hat{F}$ として得られる。すなわち,

$$\widehat{SD}=\sigma(\hat{F}) \quad \cdots\cdots(19)$$

$\hat{F}$ を前節と同じく離散型一様分布とすると, $\hat{F}$ は F のノンパラメトリックな最尤解となるから, $\widehat{SD}$ は Sd のノンパラメトリックな最尤推定値と言うことができる。

一般に $\sigma(F)$ は陽に書き下せないので, (19)式の $\widehat{SD}$ の計算を実行するには Monte Carlo 法を用いる必要がある。bootstrap 法の手順は次のようである。

i) F に経験分布 $\hat{F}$ をあてはめる。ただし, $\hat{F}$ は

$$\hat{F}: mass\frac{1}{n} at\ x_1, x_2, \ldots, x_n$$

ii) $\hat{F}$ から bootstrap 標本 X_i^* (ただし, $X_1^*, X_2^*, \ldots, X_n^* \overset{iid}{\sim} \hat{F}$) を抽出し, 統計量 $\hat{\theta}^{*1}$ を計算する。平たく言うと, n 個のデータ $x_1, x_2, \ldots, x_n$ から無作為に反復を許して n 個抽出し, その抽出したデータから $\hat{\theta}^{*1}$ を求める。

iii) ii) の作業を独立に多数回 (B 回) 繰り返し, $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B}$ を求める。

iv) 統計量 $\hat{\theta}$ の標準偏差 $\widehat{SD}$ は, 結局, 次式で与えられる。

$$\widehat{SD}=\left[\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}^{*b}-\hat{\theta}^{*\cdot})^2\right]^{1/2} \quad \cdots\cdots(20)$$

ここに, $\hat{\theta}^{*\cdot}$ は $\hat{\theta}^{*i}(i=1, 2, \ldots, B)$ の平均で

$$\hat{\theta}^{*\cdot}=\frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^{*b}$$

である。(20) 式において繰り返し回数 B を ∞ とすれば(20)式は (19) 式と一致することがわかるが, 実際にはどのくらいの回数が必要であるか不明である。データ数が少ない時には B を大きくしすぎないように注意しなければならない。**5.** に示す検討の途上で, $n=20\sim70$ に対して $B=100$, 1000, 10000 を試みたところ, ほとんどの場合, $B=1000$ と 10000 の結果は大差なく, $B=100$ の結果が若干それとは異なっていた。$n=20$ の場合だけは, $B=100$ と 1000 の結果が似通っていて, $B=10000$ のときとは少し差があった。したがって, **5.** では $B=1000$ の結果を示している。

3.5 モデル評価手順へのリサンプリング手法の適用

確率水文量の変動性 (推定誤差) を種々の分布について求めるのは一般には容易でない。しかし, jackknife 法や bootstrap 法を用いれば, 比較的容易にそれが実現できる。

3.1 の⑤の手順において jackknife 法を適用する場合，1つの分布モデルに対して，$n-1$ 個のデータからなる n 組のデータセットそれぞれについて1回ずつ計 n 回最尤推定を行うことになる。この n 回のあてはめの1回ごとに得られる T 年確率水文量が(7)式の $\hat{\theta}_{(i)}$ に対応する。したがって，(8)式から $\hat{\theta}_{(\cdot)}$ が求まり，$\hat{\theta}$ は **3.1** の③の手順で予め求めておけばよいから，偏倚を補正した T 年確率水文量 $\tilde{\theta}$ が(10)式により求まり，その変動性（分散の推定値 $\widehat{VAR}$）が(17)式により与えられることになる。こうして得た $\sqrt{\widehat{VAR}}$ が確率水文量の変動性を示す。

bootstrap 法を適用する場合は，**3.1** の⑤において，n 個のデータからなる元の標本から繰り返しを許して n 個抽出した bootstrap 標本を B 組作り出し，この B 組すべてに最尤法により分布のあてはめを行う。こうして，B 個の T 年確率水文量 $\hat{\theta}^{*b}(b=1, 2, \ldots, B)$ を求め，(20)式より $\widehat{SD}$ を得る。この $\widehat{SD}$ が確率水文量の変動性を与える。

4. モデル評価の結果と考察

4.1 琵琶湖流域の年最大k日降水量への適用

琵琶湖流域の年最大 k 日降水量（k=1, 2, 3）に，2母数の5つの分布（正規分布，対数正規分布，指数分布，Gumbel 分布，対数 Gumbel 分布）をあてはめたところ，正規，指数，対数 Gumbel の3分布は適合度が悪かった[1]。また，**2.** で見たように，対数正規分布と Gumbel 分布は，適合度規準では優劣がつけ難かった。

これら2つの分布に jackknife 法を適用した結果を **Table 3** に示す。この表から次のようなことが言える。

(1) k=1, 2, 3 のすべての場合について，確率水文量の変動性の小さいのは Gumbel 分布の方である。したがって，本研究で提示した評価手順によれば Gumbel 分布の方が対数正規分布よりも優れていると言える。

(2) **Table 1** の k=3 の欄を見ると，適合度評価では対数正規分布の方が全ての規準についてよい評価を得ていたが，これはその評価を覆す結果である。

(3) **Table 3** の k=1, 2, 3 の各欄の下段には70個のデータを全て用いて得た確率水文量 $\hat{\theta}$ を，上段には jackknife 法により偏倚を補正した $\tilde{\theta}$ を記している。対数正規分布の場合，最尤法によって全データにあてはめて得た $\hat{\theta}$ をほとんど補正する必要がないことがわかる。Gumbel 分布の場合，最尤法に

Table 3　Variability of T-year event of the k-day precipitation (in mm) of the Lake Biwa basin, for the period 1912-1981 (70 years), obtained by the jackknife method

k	Prob. distribution		Log-normal distribution			Gumbel distribution		
	Return period, T (year)		50	100	200	50	100	200
1	Jackknife	$\tilde{\theta}$	172.8(12.7)	189.1(15.1)	205.3(17.5)	172.8(11.1)	189.7(12.7)	206.6(14.3)
	All data	$\hat{\theta}$	172.9	189.2	205.5	171.3	187.9	204.5
2	Jackknife	$\tilde{\theta}$	236.1(16.7)	259.4(19.9)	282.8(23.2)	239.6(13.8)	263.6(15.8)	287.4(17.9)
	All data	$\hat{\theta}$	236.3	259.6	283.1	234.1	257.5	280.8
3	Jackknife	$\tilde{\theta}$	269.8(19.0)	296.7(22.6)	323.6(26.3)	269.3(15.8)	296.4(18.1)	323.5(20.5)
	All data	$\hat{\theta}$	269.9	296.9	323.9	267.3	294.2	320.9

( ) denotes $\sqrt{\widehat{VAR}}$, the standard deviation of the T-year event estimated.

よれば若干小さめの確率水文量 $\hat{\theta}$ を得，jackknife 法により +2～+7mm 程度補正されている。

(4) k=2 の場合を除き，対数正規分布と Gumbel 分布が与える確率水文量の jackknife 推定量に大きな差はない。

Bootstrap 法を適用した場合も，モデル評価の結果は同じであった。これについては **5.** で述べる。

4.2 大阪の年最大日降水量への適用

2. で見たように，10種の分布を最尤法であてはめたところ，3つの適合度規準 (SLSC, MLL, AIC) のどれに対してもベスト3に入らない分布が5つあった (**Table 2**)。それらは除外して，残りの5つの分布に

Table 4 The estimates of the T-year event of daily precipitation at Osaka, for the period 1889-1980 (92 years) and the standard deviations (in mm), obtained by the jackknife method

Prob. distribution	Return period, T (years)		
	50	100	200
SQRT-ET-max (2p)	*180.46 (11.25)**	203.56 (13.58)	227.85 (16.08)
Log-normal (3p)	179.94 (17.83)	201.66 (23.69)	224.17 (30.44)
Pearson Ⅲ (3p)	172.82 (14.67)	189.34 (17.34)	205.51 (20.36)
Log-Pearson Ⅲ (3p)	181.99 (19.56)	205.69 (27.06)	230.87 (36.13)
Log-Gumbel (3p)	182.95 (21.07)	207.62 (30.35)	233.86 (42.03)

* denotes the jackknife estimate of T-year event (in mm);

** the standard deviation $\sqrt{\widehat{VAR}}$.

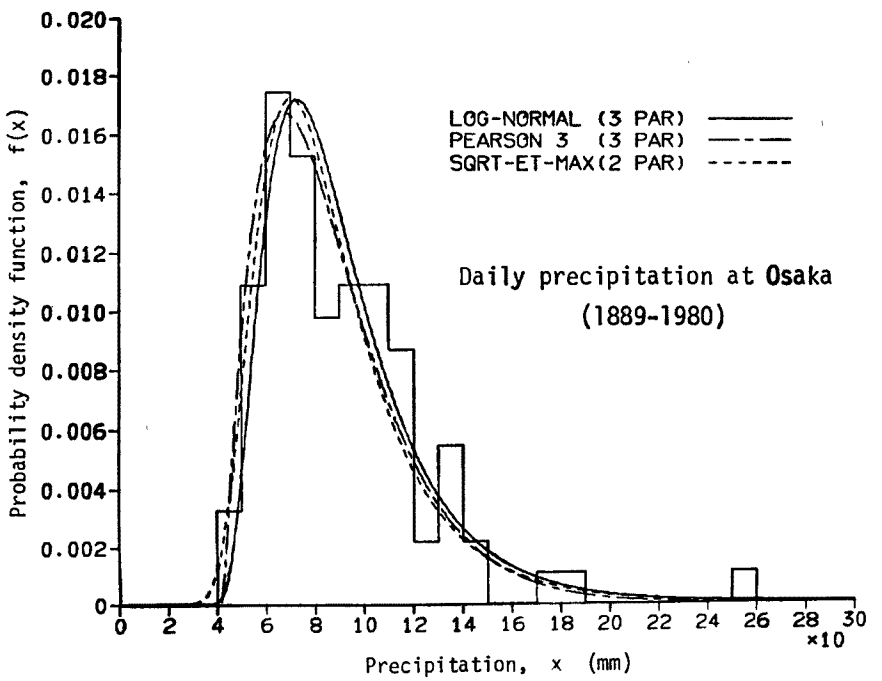

Fig. 3 Distributions whose biases were corrected by the jackknife method for the annual maxima of daily precipitation (in mm) at Osaka, for the period 1889-1980 (92 years).

ついて jackknife 法を適用した。結果を **Table 4** に示す。この表から，以下のことが言える。

(1) 確率水文量の変動は，平方根指数型最大値分布が最も小さく，これが最も良いモデルであると評価できる。

(2) 3母数 Pearson Ⅲ 型分布の確率水文量の変動も残りの3つの分布と比べるとかなり小さい。しかしながら，確率水文量の jackknife 推定値がかなり小さく，この分布の200年確率水文量は，平方根指数型最大値分布の100年確率に相当する。他の分布と比べてみて危険側に過ぎるようである。

Fig. 3 には，平方根指数型最大値分布，3母数 Pearson Ⅲ 型分布，3母数対数正規分布について，jackknife 法による偏倚補正後の密度関数の形状を示した。補正前の図 (**Fig. 2 (a)**) と比べると興味深い。平方根指数型最大値分布と3母数 Pearson Ⅲ 型分布は補正による形状の変化が大きく，ヒストグラムの形状に近づいたように見える（ピーク付近の変化が特にそうである)。対数正規分布の変化が小さいのは前節と同じ傾向である。Peason Ⅲ 型分布は他の2つに比べて右に歪んでおり，小さな確率水文量を与えている **Table 4** の結果がこの図から確認できる。

5. データの個数と確率水文量の変動性

琵琶湖流域の年最大 k 日降水量 (k=1, 2, 3) に対して，2母数対数正規分布と Gumbel 分布の両方について bootstrap 法を適用した。この際，n 個の原データから繰り返しを許して m 個抽出 ($m \leqq n$) することにより bootstrap 標本を構成し ($m < n$ の場合を部分抽出法と呼んでおく)，m の値を変化させて，デ

Table 5　Change of variability of T-year event of the k-day precipitation (in mm) of the Lake Biwa basin with m, the number of resampled data (B=1000 iteration)

k	Methods	m	Log-normal distribution			Gumbel distribution		
			50	100	200	50	100	200
1	Partial resampling with repetitions	20	172.8(24.2)	189.2(28.7)	205.6(33.4)	169.3(20.9)	185.6(23.9)	201.7(27.0)
		30	172.4(19.0)	188.6(22.5)	204.9(26.2)	169.8(16.6)	186.2(19.0)	202.5(21.4)
		40	172.0(16.4)	188.2(19.5)	204.4(22.7)	169.9(14.4)	186.3(16.4)	202.7(18.5)
		50	171.8(15.1)	188.0(17.9)	204.1(20.8)	170.1(13.2)	186.5(15.2)	202.9(17.1)
		60	171.7(13.8)	187.8(16.3)	203.9(19.0)	170.1(11.9)	186.6(13.6)	203.0(15.3)
	Bootstrap	70	171.6(12.5)	187.7(14.8)	203.8(17.2)	170.1(10.9)	186.6(12.5)	203.0(14.1)
2	Partial resampling with repetitions	20	236.5(32.1)	260.0(38.2)	283.6(44.7)	231.9(27.5)	254.8(31.3)	277.6(35.2)
		30	236.4(25.4)	259.8(30.2)	283.3(35.3)	232.9(21.8)	256.0(24.9)	279.1(28.0)
		40	235.7(22.0)	259.0(26.2)	282.3(30.5)	233.0(19.0)	256.2(21.7)	279.3(24.4)
		50	235.8(20.3)	259.0(24.0)	282.3(28.0)	233.4(17.5)	256.6(20.0)	279.7(22.4)
		60	235.5(18.1)	258.8(21.4)	282.0(25.0)	233.3(15.5)	256.5(17.7)	279.7(19.9)
	Bootstrap	70	235.2(16.5)	258.3(19.6)	281.5(22.8)	233.3(14.4)	256.5(16.4)	279.7(18.4)
3	Partial resampling with repetitions	20	269.4(36.0)	296.3(42.8)	323.4(50.0)	264.0(31.2)	290.1(35.6)	316.1(40.1)
		30	269.9(28.3)	296.8(33.6)	323.9(39.3)	265.7(24.7)	292.1(28.2)	318.5(31.7)
		40	269.3(24.6)	296.1(29.2)	322.9(34.1)	265.9(21.4)	292.4(24.5)	318.8(27.5)
		50	269.4(22.3)	296.3(26.4)	323.2(30.8)	266.6(19.4)	293.2(22.2)	319.8(24.9)
		60	269.2(20.3)	296.0(24.1)	322.9(28.2)	266.5(17.6)	293.1(20.1)	319.7(22.6)
	Bootstrap	70	268.7(18.7)	295.4(22.2)	322.2(25.9)	266.3(16.4)	293.0(18.8)	319.5(21.1)

ータの個数と確率水文量の変動性との関係を調べた。bootstrap 標本の組数 B=100, 1000, 10000 を試みたが，**3.4** の末尾に記した理由により B=1000 の場合の結果を **Table 5** に示した。B=100, 10000 の場合も大略同様の結果である。

まず，$m=n=70$ の場合（純粋な bootstrap 法），確率水文量の変動性は，jackknife 法を適用したときと同様，Gumbel 分布の方が k=1, 2, 3 のどれに対しても小さい。jackknife 法において得られた $\sqrt{\widehat{VAR}}$ (**Table 3**) と bootstrap 法で得られた $\widehat{SD}$ とを比べると全て近い値を示していることがわかる。このことから，**3.1** のモデル評価の手順⑤においては jackknife 法を用いればよく，bootstrap 法を併用するまでもなさそうに思われる。

次に，$m<n=70$ の場合（部分抽出法）を見てみよう。m=20, 30, 40, 50, 60 の5ケースを考えた。この場合も k=1, 2, 3 のいずれについても Gumbel 分布の確率水文量の変動性の方が対数正規のそれより小さく，このデータに対する Gumbel 分布の優位性を示している。

さて，bootstrap 標本のデータ数 m の変化に伴なう確率水文量の変動性（推定誤差）の変化を見てみよう。m が増加するにつれて推定誤差は減少する傾向を示し，その値は n=70 個の標本による推定誤差に対して，m=60 で約1割増，m=50 であれば約2割増，m=20 では約2倍になることがわかる。この傾向はリターンピリオド T=30, 50, 100, 200 年のいずれについても同様であった。m=30 から 40 を境として，すなわち資料年数が30年から40年以上になると，確率水文量の変動が小さくなる（安定する）ように見える。この結果は，長野県下28か所の年最大日降水量資料に Gumbel 分布をあてはめた寒川ら[14]の検討とも符合する。**Table 5** によれば，Gumbel 分布の場合，$m\geqq$30 程度で，標準偏差が確率水文量の10%程度以内に収まり，m=70 では7%程度にまで低下することがわかる。

6. 結　　語

当該水文量が従う確率分布を決定する手順が未だに確立されていないのは，モデル評価のための有効な評価規準が見当たらなかったことが大きな要因の1つであると言える。そこで，本研究では，直観的にも理解し易く，実用的にも重要な問題であるところの確率水文量の変動性を規準として新しいモデル評価の手順を提案した。

得られた成果は以下の通りである。

(1) 従来モデル評価は，データとモデルの適合度の主観的判定に委ねられる場合が多かった。モデル評価の規準として適合度は必ずしも十分ではないことを，実例をもって示した。

(2) 確率水文量の変動性を規準としたモデル評価の手順を具現し，多数（10種以上）の確率分布モデルを一挙に評価できるようにした。

(3) 確率水文量の変動性の推定にリサンプリング手法（jackknife 法と bootstrap 法）を応用し，その変動性とデータ個数の関係を定量的に評価した結果，琵琶湖流域平均降水量の極値データの場合，Gumbel 分布が良く，資料年数 $m\geqq$30 程度で標準偏差が確率水文量の10%程度以内に収まり，m=70 では7%程度にまで低下することがわかった。

データの蓄積がかなり進み，また，高速に大量の統計処理が可能となった今日，本研究で提示した確率分布モデルの評価法は時宜にかなったものと考える。今後さらなる適用を通じてその有用性を検証してゆきたい。

参 考 文 献

1) 高棹琢馬・宝　馨・清水　章：琵琶湖流域水文データの基礎的分析，京大防災研年報，第29号 B-2, 1986, pp. 157-171.

2) Akaike, H.: A new look at the statistical model identification, *IEEE Trans. Autom. Contr.*, AC-19, 1974, pp. 716-723.
3) 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.
4) 神田　徹・藤田睦博：水文学——確率論的手法とその応用——，技報堂出版，1982, p. 49.
5) 江藤剛治・室田　明・米谷恒春・木下武雄：大雨の頻度，土木学会論文報告集，第369号/II-5, 1986, pp. 165-174.
6) 富士通（株）：FACOM FORTRAN SSL Ⅱ 使用手引書，1980, pp. 403-406.
7) 宝　　馨・清水　章・高棹琢馬：降水量データに対する確率分布の適合度について，昭和62年度関西支部年次学術講演会講演概要，土木学会関西支部，1987, Ⅱ-58.
8) Miller, R. G.: The jackknife—a review, *Biometrica*, Vol. 61, No. 1, 1974, pp. 1-15.
9) Efron, B.: The Jackknife, the Bootstrap and Other Resampling Plans, *SIAM Monograph* No. 38, 1982.
10) Bardsley, W. E.: A test for distinguishing between extreme value distributions, *Journal of Hydrology*, Vol. 34, 1977, pp. 377-381.
11) Tung, Y.-K. and L.W. Mays: Generalized skew coefficients for flood frequency analysis, *Water Resources Bulletin*, Vol. 17, No. 2, 1981, pp. 262-269.
12) Cover, K.A. and T.E. Unny: Application of computer intensive statistics to parameter uncertainty in streamflow systhesis, *Water Resources Bulletin*, Vol. 22, No. 3, 1986, pp. 495-507.
13) 奥村晴彦：パソコンによるデータ解析入門——数理とプログラミング実習——，技術評論社，1986, pp. 215-224.
14) 寒川典昭・荒木正夫・渡辺輝彦：確率分布の推定母数の不確定性評価法，土木学会論文集，第375号/Ⅱ-6, 1986, pp. 133-141.