運輸省港湾技術研究所

港湾技術研究所報告

REPORT OF THE PORT AND HARBOUR RESEARCH INSTITUTE

MINISTRY OF TRANSPORT

VOL. 27 NO. 1 MAR. 1988

NAGASE, YOKOSUKA, JAPAN

港湾技術研究所報告 (REPORT OF P.H.R.I.)

第27巻　第1号 (Vol. 27, No. 1), 1988年3月 (Mar. 1988)

目　　次 (CONTENTS)

港湾技術研究所報告
第27巻第1号（1988. 3）

2. 極値統計におけるプロッティング公式ならびに推定値の信頼区間に関する数値的検討

合　田　良　実*

要　　旨

強風，洪水，高波などの異常事象の出現確率は極値統計資料に基づいて解析される。この解析手法については未解明の点が多いため，モンテカルロ法による大規模な数値実験を行った。対象とする分布関数は，FT-I 型分布，対数正規分布，ワイブル分布（k=0.75, 1.0, 1.4，および2.0）その他であり，標本の大きさが10～100のものを原則として各10,000回抽出し，分布関数のあてはめ，各種の再現期間に対する確率極値の推定その他の作業を行った。

分布関数のあてはめには最小2乗法を用いたが，プロッティング公式を適切に選定することによって実用上十分な精度で母数推定を行い得ることを確認した。母数推定値についてはその信頼区間を数表の形で提示した。各種再現期間に対する確率極値の信頼区間は，母分布関数が既知の場合については実験式の形でとりまとめられた。この実験式は毎年最大値資料のみならず，上位の極大値のみを抽出した部分極値資料に対しても適用される。

母分布関数が未知な場合には，データ数の少ない標本資料について分布関数のあてはめを行っても真の分布関数の推定が困難であることが統計的に例証された。このため，標本資料に対して最も適合する分布関数を採択しても真の分布関数による極値と異なる値を推定する公算が大きく，これに対する偏り補正が必要である。本報告では，全数極値資料と部分極値資料に分けて偏り補正の実験式を提示した。さらに，各種再現期間に対する確率極値の標準誤差について，母分布関数が未知な場合の実験式もとりまとめられた。

以上のほか，気象原因別の極値統計の解析結果の取り扱いや，信頼性設計法における変動係数等についても論じている。

キーワード：波浪統計，極値統計，極値分布関数，プロッティング公式，確率波高，再現期間，信頼区間，モンテカルロ法

*所　長

REPORT OF THE PORT AND
HARBOUR RESEARCH INSTITUTE
Vol. 27, No. 1 (Mar. 1988)

2. Numerical Investigations on Plotting Formulas and Confidence Intervals of Return Values in Extreme Satistics

Yoshimi GODA*

Synopsis

Environmental design conditions such as strong winds, flood discharges, and storm waves are selected based on their occurrence probabilities, which are estimated by means of extreme statistics analysis. A large scale numerical experiment by the Monte Carlo method is carried out to give answers to several unsolved problems of extreme wave statistics. Analyses are done for the FT-I type distribution, the log-normal distribution, and the Weibull distribution (k=0.75, 1.0, 1.4, and 2.0). As a standard, 10,000 samples with the size ranging from 10 to 100 are drawn from the parent distribution, the best-fitting distribution to a sample is determined, and the return values for various return periods are estimated.

The least square method is employed for fitting of a distribution function to a sample. It is confirmed that the least square method yields satisfactory results by choosing the best plotting-position formula for respective distribution functions. The confidence intervals of the estimates of the scale and location parameters are presented in tabular forms. Empirical formulas are established for the standard deviation of the return values for the case where the true distribution function is known. The formulas are applicable for both censored and uncensored samples of extreme data.

When the true distribution function is unknown, the best-fitting function to a sample does not necessarily coincide with the true one, and the probability of misfit increases as the sample size decreases. Misfit yields a bias on the estimate of return values. Empirical formulas are presented for the correction of the bias of estimated return values. Censored and uncensored data are given different formulas. The standard error of the estimate of return values is also synthesized in the form of empirical formulas for the case where the true distribution function is unknown.

The present report further discusses the method for the synthesis of different population data, the coefficient of variation of the load factors in the load and resistance factor design method, and others.

Key Words: Wave Statistics, Extreme Statistics, Extreme Distributin Function, Plotting Position Formula, Return Period, Return Value, Confidence Interval, Monte Carlo Method

* Director General.

目　次

1. まえがき

自然を相手に構築物を作る土木工学の分野では，自然の力の大きさを正しく見積ることがまず何よりも大切である。暴風，地震，高波など，考えられる最大級の事象が発生しても，構築物は安全にその機能を発揮していることが期待される。このため，こうした自然の外力の発生頻度を解析し，100年あるいは1000年に1度起きるような現象の大きさを推定する作業が行われる。この際の基礎データとなるのは，対象とする自然現象の毎年の最大値，あるいは発生した年に関係なしに取り出したすべての異常事象（ある大きさ以上とする）の極大値などである。こうしたデータを統計資料として解析するのが極値統計であり，土木工学では河川の洪水流量の解析が最も古くから行われている。極値統計は気象その他の分野でもいろいろ解析されており，Gumbel の書物[1]は古典として有名である。

極値統計の類縁に生物の寿命や製品の耐久時間を統計的に取り扱う Lifetime 統計があり，医薬品の性能判定や工場の品質管理等に活用されている。これらは極大値ではなく極小値を対象とするが，解析手法はほとんど同じといえる。この分野については最近 Lawless が専門書[2]を著している。

ひるがえって港湾・海岸・海洋構造物を考えてみると，設計外力として支配的なものの一つが波浪であり，設計波の選定には多大の努力が費やされる。しかしながら，波浪データは実測値が短期間しか得られていないものが多く，設計サイドから要求される50年，100年などの長期確率の値をどのように推定すべきかについての検討手法が確立されているとはいいがたい。実測データの不備を補うため既往の台風や低気圧などによる高波を追算する作業もしばしば行われる。しかしその場合でも，古い天気図は精度が低いため，実用に供し得るのは30～40年程度の期間にとどまる。

極値統計における最大の問題は，対象とする自然現象がどのような確率分布に従うのかが分らないことである。過去数十年程度の極値データを基に，そのデータに最も適合すると思われる確率分布を手探りで求めているのが現状である。極値統計の解析が比較的数多く行われている水文統計であっても事情は同じである[3),4)]。土木工学において自然外力に関する設計値を選定するときは，そうしたあまり信頼度の高くない確率分布を外挿し，非常に長期の時間内に起ると想定される極値を推定せざるを得ない。その場合，推定値は当然に確定した値ではありえず，その値の上下のある幅の中に存在するとしかいえない。この幅は一般に信頼区間と呼ばれている。したがって，極値統計に基づく推定値はその信頼区間を指定して結果を示さなければならないにもかかわらず，信頼区間をどのように推定するかについては十分に吟味されていない。特に，我が国においては極値推定の際の信頼区間に関する文献をあまり見ることができない。

土木工学の分野で極値統計を取り扱う際のもう一つの問題は，統計学があまりに専門化されていて，土木技術者にとって非常に分りにくいものになっていることである。特に，波浪データに関しては統計学の知識を十分にそしゃくしたとは思われない研究発表をしばしば見かける。たとえば，長期間にわたる設計波高を選定する際に一日数回の観測値のすべてを取り上げ，1～数年程度のデータに対してあてはめた確率分布を50年，100年の確率レベルにまで外挿する方式がいまだに外国では用いられている[5)]。この方法では基礎となるデータ相互間に強い相関があるため，統計学上要求されるデータの独立性の条件を満足していない。また，1～数年の全データに対する確率分布と高波の極値に対する確率分布は明らかに異なるにもかかわらず，その差異を無視している。これについては，先に著者[6)]が指摘したところであり，また近年でも Muir と El-Shaarawi[7)]が言及している。

このような状況にかんがみ，本報告は波浪データの解析を主対象として，極値統計の解析手法について統計学的に誤りでないと思われる方法についてやや教科書的に解説するとともに，理論的に不明確な部分については数値シミュレーション結果に基づいて，現状においては妥当ではないかと思われる実用的手法を提案するものである。後者については今後の研究によって修正が加えられ，あるいは別の方法によって取って替られることも起きると思われる。そうなることは極値統計の発展にとって誠に喜ぶべきことであり，それまでの間，本報告で提案する方法が設計実務に役立つことができるならば，本報告の目的は十分に達せられたといえよう。

なお，極値統計では幾つかの専門用語が使われる。本報告ではできるだけ水文統計の用語に合せたつもりであるが，著者の不注意によって正しくない用法が残っているおそれがある。したがって，水文統計等の文献を比較参照される際に，本報告での誤用について適宜修正して使って頂くようお願いする次第である。

本報告の構成は目次に示すとおりであるが，実務において波浪データの極値統計解析に携わる各位の便宜を考えて実際的な手順を付録に示したので，参考にして頂ければ幸いである。

2. 極値統計の基礎的事項

2.1 極値統計資料とその分類

(1) データの要件

まえがきで述べたように，極値統計で取り扱うデータにはいろいろなものがある。その数例を挙げると以下のとおりである。

1) 気象関係：風速，気温等
2) 水文関係：降雨量，洪水流量，渇水流量
3) 海象関係：波高，高潮位（または高潮偏差）
4) 地象関係：地震震度（あるいは加速度）
5) Lifetime：医学実験用動物の生存期間，製品・部品等の使用可能時間等

こうしたデータを極値統計として解析する際には，データに対して**独立性**（independency）と**等質性**（homogenuity）の二つの条件が強く要求される。独立性というのは，個々のデータの間に相関関係が存在しないこと（相関係数が0に近いこと）である。数時間ごとに観測された波高データの場合には，24時間後のデータとの相関係数が日本沿岸で0.3以上[6),8)]，オーストラリア東海岸で0.5以上[9)]と報告されており，こうしたデータは互いに独立であるとはいえない。第2の等質性というのは，対象とするデータがすべて同一の確率分布の集団（母集団）から抽出されたものであることを指す。同一の母集団に属することは証明困難であるので，対象とする現象の発生メカニズムを考えて判断するのが普通である。たとえば，台風による強風と低気圧による強風とでは風速の絶対値に差があることは容易に推察される。したがって，たとえ確率分布の関数形が同じであってもパラメータ* の値は異なるので，これらは異なる母集団に属する。高波についても同様である。異なる母集団からの抽出データを分離せずに混じったままで処理すると正しくない極値を推定することになる。Resio[10)] は波浪データについて解析例を示している。

(2) 期間最大値資料と極大値資料

極値統計のデータの中でも Lifetime に関するものなどは，個々のデータがそれぞれ個別に計測され，定義される。しかし，水文量や波浪データなどは時間とともに連続的に変化する量であり，何らかの方法で極値統計のデータとしての値を定義する必要がある。一つの方法は，1か月，1年などの時間単位を定め，その期間中の最大値をもって極値統計のデータとする方法である。洪水のピークがちょうど二つの期間の境目に起った場合などを除けば，こうした期間最大値はそれぞれ独立であると見なすことができる。水文統計では主として毎年最大値等を利用するところから，毎年最大値資料（annual maximum series）などと呼んでいる[11a)]。本報告では毎月の最大値を取り扱う場合も考えて，**期間最大値資料**と呼んでおく。なお，渇水流量のように最小値を対象とする極値統計も考えられるが，ここでは説明を簡単にするため最大値のみについて述べることにする。

もう一つの方法は，台風や低気圧等の気象じょう乱によって高波が発生し，発達し，やがて減衰する過程を一つの事象と見なし，その過程中の最大値を，データとして取り上げるものである。水文統計における部分的水文資料（partial-duration series）の中の非毎年超過値（nonannual exceedance series）に相当すると思われるものである。著者は以前にこれを「極値時系列」と呼んだが[6)]，その名称では定義がまぎらわしいおそれがあるので，ここでは単純に**極大値資料**と呼ぶことにする。この極大値資料の場合，引き続いて発生する二つの台風・低気圧による高波の間には相関があるとは考えにくい（ゼロでないにしても非常に弱い）ので，極大値資料もデータの独立性を満足しており，極値統計の対象として使うことができる。

なお，期間最大値資料，極大値資料のいずれも等質性についての吟味が必要であり，台風・低気圧等の気象原因別に極値統計資料を用意しなければならない。

また，極値統計に関しては，同一の原データ，たとえば長期間にわたる波浪，降雨量等の記録から抽出されたものであっても，期間最大値資料と極大値資料とでは確率分布形状が異なることがある。たとえば，単位期間中に発生する極大値の数が十分に大きく，それらが指数分布に属するときは，期間最大値資料は後述の二重指数分布で近似できることが多い。

極大値資料の解析においては，極大値の発生頻度も重要なパラメータである。今，K年間に対象とする事象がN_T個発生したとすると，1年当りの平均発生回数は

$$\lambda = N_T/K \qquad (1)$$

である。本報告ではこれを**平均発生率**と呼んでおく。

(3) 全数極値資料と部分極値資料

極値統計の資料としては，対象とすべきデータをすべて拾い上げているものばかりでなく，対象とすべきものの一部が脱落したままのものも少なくない。製造部品の寿命テストなどでは破損等が起きた順にその耐久時間を記録していくが，テスト部品の中にはいつまでも破損しないで機能するものも出てくる。すべての供試体が破損するまでテストを続けるのは不経済なので，普通はある

* 統計の分野では統計量の分布関数のパラメータを母数ということが多い。

時間で試験を打ち切り，機能し続けた供試体は寿命N時間以上と記録する。こうしたデータは censored data と呼ばれており，全数テストのものは uncensored data といわれる。

自然現象の極値統計においても似たようなデータの部分的取得が行われる。波浪データについて極大値資料を準備するときであれば，あまり大きくない高波のデータは選別の手間を省くため除外し，ある波高レベル以上のものに注意を向けるのが普通である。特に，過去の高波を波浪推算法によって推定する場合は推算作業の経費を節約するため，対象とする気象じょう乱の数を制限する。Lifetime データの場合がいわば「頭切り」であるのに対し，波浪データでは「足切」りが行われることになる。

一方，期間最大値資料の場合には各単位期間ごとにそれぞれ1個の最大値データを抽出するのであるから，全数データが容易に取得できる。すなわち uncensored data である。

こうしたデータ採択の差異は極値統計の手法にも影響を及ぼす。この差異を明示する術語は固まっていないようであるので，本報告では uncensored，censored に対応させてそれぞれ**全数極値資料**，**部分極値資料**と呼び分けることにする。そして，**データ採択率**νを次のように定義する。

$$\nu = N/N_T \tag{2}$$

ただし，Nは極値資料として採択したデータの個数，N_Tは対象とする期間に発生した当該事象の個数であり，式(1)に用いたものと同じである。全数極値資料にあっては$\nu=1$，部分極値資料にあっては$\nu<1$である。

前項の分類と合せると，極値統計資料は次のように分類される。

$$\text{極値統計資料}\begin{cases}\text{期間最大値資料}\begin{cases}\text{全数極値資料}\\\text{部分極値資料}\end{cases}\\\text{極大値資料}\begin{cases}\text{全数極値資料}\\\text{部分極値資料}\end{cases}\end{cases}$$

期間最大値資料は通常は全数極値資料として解析されるけれども，水文統計でいう「非毎年最大値資料」のように特定の値以下の年のデータを除外した部分極値資料のこともあり得る。一方，極大値資料は大半が部分極値資料であるが，台風に起因する事象などの場合には全数データを取得することも可能である。

2.2 分布関数とその特性

(1) 分布関数の定義

極値統計は，得られた資料に対して特定の確率分布をあてはめ，その分布関数を使って50年，100年などの長期間に1回以上出現すると予想される事象の大きさやその推定値の信頼区間その他を推定することを主要な役割としている。

ここで分布関数というのは，統計的に分布する量を大きさの順に並べてその累積度数を求め，全体個数に対する比率で表したものであり，その統計量が特定の値を超えない確率を表す。すなわち，統計量をXで表すと

$$F(x) = \mathrm{Prob.}[X \leqq x] \tag{3}$$

したがって，分布関数は$x=-\infty$から$x=\infty$の間において0から1に向って単調に増加する（局部的には一定でもよい）関数である。

分布関数をxについて微分したものは確率密度関数$f(x)$である。統計量tがxから$x+dx$の間に分布する確率は$f(x)dx$で表される。

与えられた統計資料についての分布関数は累積度数分布に基づいて直ちに計算されるけれども，我々が知りたいのは標本として入手した統計資料が所属する母集団の分布関数である。厳密な意味での母集団の特性は永遠に未知であるが，我々はそれをできる限り高い精度で推定しようと努力する。分布関数については，数学的取り扱いの難易や種々の分布形状への適合の柔軟性などを考慮していろいろな関数形を仮定し，標本としての統計資料にあてはめて適合度を検討する。

極値統計において利用される分布関数については，Gumbel[1]，Lawless[2]，神田・藤田[11]の書物などにいろいろ解説されておりまた波浪の極値統計の紹介論文[7),12)]も記述されている。以下においては，波浪データの解析用として本報告で取り上げた分布関数について述べる。

(2) FT-I 型分布（Fisher-Tippett Type I）

これは二重指数分布であり，また Gumbel[1]がこれについて詳しく論述し，水文資料等への応用を解説したところから Gumbel 分布とも称されている。分布関数および確率密度関数は次のとおりである。

$$F(x) = \exp\{-\exp[-(x-B)/A]\} \tag{4}$$

$$f(x) = (1/A)\exp\{-(x-B)/A - \exp[-(x-B)/A]\} \tag{5}$$

ここに，パラメータAは尺度母数（scale parameter），Bは位置母数（location parameter）と呼ばれる。なお，文献によっては$a=1/A$を尺度母数として使っており，またAとBを逆に使う例もある。

この分布について平均値を計算すると*

$$\mu = E[x] = \int_{-\infty}^{\infty} x f(x)\,dx = B + A\gamma \tag{6}$$

ここに，γはオイラー定数（0.5772…）である。標準偏

* この積分は Laplace 変換型の定積分公式[13]を利用して求める。

差σについてはGumbel[1)]（5.2.5節）が積率母関数を用いて計算しており，その結果は次のように書き表される。

$$\sigma=\pi A/\sqrt{6} \quad (7)$$

(3) 2母数対数正規分布

これはxの自然対数 $X=\ln x$ が正規分布に従うような分布であり，次式で記述される。

$$f(x)=\frac{1}{\sqrt{2\pi}\,xA}\exp\left[-\frac{(\ln x-B)^2}{2A^2}\right] \quad (8)$$

$$F(x)=\int_{-\infty}^{x} f(t)dt \quad (9)$$

ここでの位置母数Bは $X=\ln x$ の平均値μ_Xに等しく，尺度母数AはXの標準偏差σ_Xに等しい。なお，ここに示したものは2母数対数正規分布といわれるもので，さらにデータへの適合度を良くするために母数を増した3母数対数正規分布といわれる分布も使われる[11b)]が，ここでは取り上げない。

この分布に対してxのr次の積率を計算すると[11c)]，

$$\nu_r=\int_0^{\infty} x^r f(x)dx$$
$$=\exp\left[\frac{1}{2}r^2A^2+rB\right] \quad (10)$$

したがって，xの平均および標準偏差は

$$\mu=E[x]=\exp[B+A^2/2] \quad (11)$$

$$\sigma=\{E[x^2]-E[x]^2\}^{1/2}=\mu[\exp(A^2)-1]^{1/2}$$
$$=\exp[B+A^2/2][\exp(A^2)-1]^{1/2} \quad (12)$$

なお，式(11)，(12)は母集団についての期待値であるので，有限個数の標本について $X=\ln x$ の平均値および標準偏差を求めてそれぞれBおよびAと置いて式(11)，(12)の右辺に代入しても，xについての標本平均$\bar{x}$および標準偏差σ_xとは若干の差異を生じるのが普通である。

(4) ワイブル分布

これはWeibull[14)]が材料の破断強度の統計的解析に際して導入したものといわれており，波浪の極値統計ではPetruaskas と Aagaard[15)] が種々の分布形状に対する適合の柔軟性を例示し，我が国ではこれを著者[16)]が紹介して以来しばしば用いられている。分布関数および確率密度関数は次のとおりである。

$$F(x)=1-\exp\left[-\left(\frac{x-B}{A}\right)^k\right] \quad (13)$$

$$f(x)=\frac{k}{A}\left(\frac{x-B}{A}\right)^{k-1}\exp\left[-\left(\frac{x-B}{A}\right)^k\right] \quad (14)$$

ここに，kは形状母数（shape parameter）である。また，xはBよりも大きいことの条件が付せられている。

このワイブル分布は下限値を有する Fisher-Tippett のIII型分布である。Lifetimeに関する統計では上限値を有するFT-III型分布がしばしば用いられ，それもワイブル分布と呼ばれているので注意する必要がある。

ワイブル分布の母集団平均および同標準偏差は式(14)の確率密度関数を用いてxの1次および2次の積率を計算することにより，次のように求められる。

$$\mu=B+A\Gamma(1+1/k) \quad (15)$$

$$\sigma=A\,[\Gamma(1+2/k)-\Gamma^2(1+1/k)]^{1/2} \quad (16)$$

ただし，$\Gamma(\ )$はガンマ関数である。なお，文献[7),12)]では式(15)のガンマ関数の引数の符号が逆になっているが，誤りである。

(5) 分布関数の形状の比較

以上の分布関数の形状を確率密度関数の形で比較したのが図-1,2である。図-1はFT-I型分布，対数正規分布，およびワイブル分布（$k=2.0$）について母数A，B

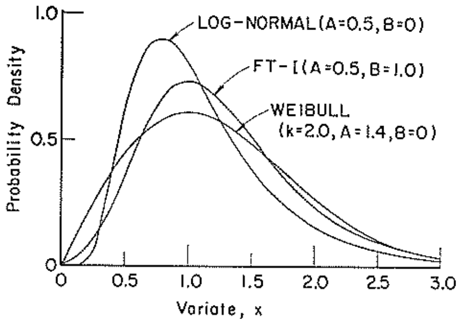


図-1 確率密度分布関数の形状(1)

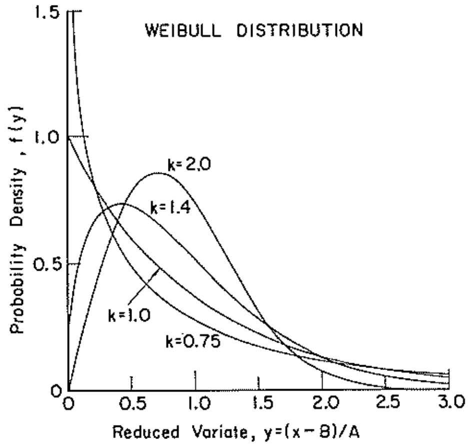


図-2 確率密度分布関数の形状(2)

を適宜選定して比較したものである。これらの分布の最大値 $f_{\max}$ およびそのときの変数値すなわちモード（並み数または最頻値）x_{mode} は次のように与えられる。

1) FT-I 型分布

$$\left.\begin{aligned} x_{\mathrm{mode}} &= B \\ f_{\max} &= e^{-1}/A \end{aligned}\right\} \quad (17)$$

2) 対数正規分布

$$\left.\begin{aligned} x_{\mathrm{mode}} &= \exp[B-A^2] \\ f_{\max} &= \frac{1}{\sqrt{2\pi}\ A}\exp\left[\frac{1}{2}A^2-B\right] \end{aligned}\right\} \quad (18)$$

3) ワイブル分布

$$\left.\begin{aligned} x_{\mathrm{mode}} &= A(1-1/k)^{1/k}+B \\ f_{\max} &= \frac{k}{A}\left(1-\frac{1}{k}\right)^{1-1/k}\exp\left[-\left(1-\frac{1}{k}\right)\right] \end{aligned}\right\} \quad (19)$$

ただし，$k>1$

図-1を見ると分るように，モードより左側では分布形状にやや違いがあり，FT-I 型では $x<0$ の領域にまで裾を引いている。しかし，モードの右側では分布形状がかなり類似している。このため，5.で述べるように極値統計の標本資料に対してこれらの分布をあてはめると，いずれの分布もよく適合することが多く，適合度に差別をつけるのがむずかしい。なお，水文統計で用いられるピアソンIII型分布や対数ピアソンIII型分布は，神田・藤田[11d)]が例示している図表から判断すると図-1の分布形状に類似しており，差別化がむずかしいと思われる。

これに対して，図-2はワイブル分布について $k=0.75\sim2.0$ を比較したもので，横軸には次式で定義される基準化変量 y を用いている。

$$y=(x-B)/A \quad (20)$$

図-2で明らかなように，ワイブル分布は $k\leqq1$ の場合が単調減少関数となるのに対し，$k>1$ の場合は式(19)で与えられる単一のピークを持つ。したがって，ワイブル分布は k の値を変えることによっていろいろな分布形状で表現できることになる。また，k の値が小さくなるにつれて y の値の大きいところまで裾を長く引くようになる。なお，江藤・室田[17)]が提案している平方根指数分布は，ワイブル分布の $k=0.5$ のものよりも裾を長く引く形をとる。

2.3 分布関数の母数の推定法

極値統計の標本資料に対して分布関数をあてはめることは，分布関数の母数 A, B, k の値を推定することを意味する。この推定法としては次のような方法が用いられる。

1) 図式推定法
2) 最小2乗法
3) 積率法
4) 最尤法

第1の図式推定法は，母関数から抽出されたデータが図上で直線上にプロットされるようにくふうされた特殊な確率紙，たとえば対数正規確率紙などを使うものである。そして，確率紙上にプロットされたデータに対して目視によって直線をあてはめ，その勾配と切片から母数 A，Bを推定するものである。

第2の最小2乗法は図式推定法における直線のあてはめを数値計算によって行うようにしたものである。計算機の発達していなかった頃は計算が面倒であるため，もっぱら図式推定法に頼っていたけれども，現在ではそうした制約がないため広く使われている。

第3の積率法は，前節で述べた各分布関数の平均 μ および標準偏差 σ をそれぞれ標本資料の平均 $\bar{x}$ および標準偏差 σ_x に等しいと仮定し，これによって母数 A, B を推定するものである。ワイブル分布の形状母数 k については標本資料のひずみ度を用いて推定することが可能であるけれども，実務上はあらかじめ幾つかの k の値を設定し，それぞれについて母数A，Bを推定した上で最も適合度の高い k の値を選定すればよい。なお，FT-I 型分布について Gumbel[1)]（6.2.3節）は標本の大きさ，すなわちデータ個数を考慮して母数A，Bを推定することを推奨している。しかし，最近の数値シミュレーション（モンテカルロ法）の結果[18),19)]によれば，Gumbel の推奨する方法では極値の推定結果に大きな偏り（バイアス bias）を伴うので，標本の平均および標準偏差を母集団値に等置すべきであるとされている。

第4の最尤法は与えられたN個のランダムなデータに対して定義される次の尤度関数の値を最大にするように母数を推定するものである。

$$L(\boldsymbol{\theta})=\prod_{i=1}^{N}f(x_i\,;\boldsymbol{\theta}) \quad (21)$$

ここに，$\boldsymbol{\theta}=(\theta_1,\cdots,\theta_l)$ は確率密度関数 $f(x\,;\boldsymbol{\theta})$ に含まれる母数をベクトル的に表示したものである。式(21)は乗積の形を取っていて尤度最大化の計算が面倒であるため，一般には $L(\boldsymbol{\theta})$ の代りに $\ln L(\boldsymbol{\theta})$ の最大化を図る。$\ln L(\boldsymbol{\theta})$ の最大値が得られる条件は次の尤度方程式が満足されるときである。

$$\frac{\partial \ln L}{\partial \theta_i}=0 \qquad : i=1,\cdots,l \quad (22)$$

式(22)の尤度方程式は対数正規分布などでは解析的に解くことができるが，一般的には非線型な連立方程式の数値解として求めなければならない。種々の分布関数に

対する尤度方程式は神田・藤田[11c]に記載されている。

以上の母数推定法はそれぞれ一長一短があり，優劣を決めることがむずかしい。我が国の水文統計では積率法が早くから用いられている。統計の分野では電子計算機の普及につれて最尤法を使う例が増えているようで，水文統計でもそうした傾向が見られる。最小2乗法は単純であるだけに洗練さを欠くと思われるのか専門家の間では評価が低い（たとえば Lawless[2a]）。しかし，本報告では以下のような理由で最小2乗法を用いて母数推定を行うことにする。

1) 最小2乗法は母数の推定精度が低いといわれるが，Carter と Challenor[19] の数値シミュレーション結果を見るかぎり他法に比べて有意な差異は認められない。
2) 最小2乗法では各データに割り当てるべき非超過確率の決定法が未解決とされているが，これについては Cunnane[20] の検討結果が出ており，さらに必要に応じて数値的検討を加えればよい。
3) 波浪データのように足切りされた部分極値資料の場合，積率法では対応が困難である。最尤法では原理的には対処可能であるが数値計算が一層面倒になる。これに対して最小2乗法では部分極値資料でも容易に取り扱うことができる。
3) 最尤法は理論的検討が多々進められていて母数推定値の信頼区間等も計算できる。しかし，計算が複雑であり，現時点では一般の技術者が利用することがむずかしい。これに対して最小2乗法は計算が簡単であって計算過程を容易に把握することができる。

したがって，以下においては分布関数のあてはめをすべて最小2乗法によって行うこととして議論を進める。

2.4 再現期間と再現確率統計量

先に述べたように，波浪データ等の極値統計解析は長い間に1度以上起るような異常値を推定することを目的としている。こうした推定においては再現期間（return period）の概念が用いられる。再現期間とは，対象とする事象のうちある特定の値 x_u を超えるものが平均して1回起きる時間間隔をいう。Borgman[23] はこの再現期間が次のようにして極値の分布関数に結び付けられることを説明している。

今，簡単のために毎年最大値資料を対象とし，その分布関数が $F(x)$ で表されるとする。事象の値 x が x_u を超えない確率は分布関数の定義によって $F(x_u)$ であり，等しいか超える確率は $1-F(x_u)$ である。最初に $x\geq x_u$ となってから $x<x_u$ の年が続き，n 年目にようやく $x\geq x_u$ となったとするとその事象の確率 P_n は

$$P_n=F^{n-1}(1-F) \tag{23}$$

で与えられる。$x\geq x_u$ の事象は $n=1$ から $n=\infty$ のどの時点でも起り得るから，その期待値を計算するとこれは定義によって再現期間 R に等しい。すなわち

$$E[n]=R=\sum_{n=1}^{\infty}nP_n=(1-F)\sum_{n=1}^{\infty}nF^{n-1}$$

$$\therefore\quad R=\frac{1}{1-F(x_u)} \tag{24}$$

極大値資料によって分布関数 $F_*(x)$ が求められている場合については，式(1)の平均発生率 λ を導入し，

$$R=\frac{1}{\lambda[1-F_*(x_u)]} \tag{25}$$

として関係づけられる。これは極大値の発生時間単位が平均して $1/\lambda$ 年であり，再現期間 R 年内にはこの単位時間が λR 個含まれると考えれば式(24)から直ちに導かれる。

極値統計では特定の値 x_u の再現期間 R を求めるだけでなく，逆に R を指定したときの極値の値 x_R も推定したい。このためには式(24)，(25)を書き換えた次式を解いて求めればよい。

$$F(x_R)=1-1/R \tag{26}$$

$$F_*(x_R)=1-1/\lambda R \tag{27}$$

具体的計算法については **4.3** で述べる。

再現期間 R 年に対応する極値は，R 年確率水文量，R 年確率波高などと呼ばれる。英語では return value の語が使われる例がある[22]。ここでは総称として R 年確率統計量あるいは再現確率統計量の語をとりあえず用いておく。この R 年確率統計量はその定義から明らかなように，その値を超える事象が平均して R 年に1回起きるような値である。すなわち，R 年に1回起きる事象の下限値である。

R 年確率統計量を推定するに当って毎年最大値資料を用いるべきかあるいは極大値資料を用いるべきかは議論の分れるところである。米国においても1940年代には洪水流量の解析に関していろいろ討論されたようであるが，1940年に Langbein[24] が再現期間5年以上であれば両者の差は無視できると結論づけ，Chow[25] もこれを支持している。ここでは，やや異なるアプローチによって同様の結論が得られることを示しておく。

極大値の年間発生回数を r で表すと，これは年によって変動するが，発生回数の確率分布が次のポアソン分布で近似できると仮定する。

$$P_r=e^{-\lambda}\lambda^r/r!\qquad : r=0,1,2,\cdots \tag{28}$$

極大値 x の分布関数は $F_*(x)$ とする。今，1年間に極大値が r 個発生し，その値を $x_1, x_2, \cdots, x_r$ とする。この

中での最大値を ξ で表すと

$$\xi=\max\{x_1, x_2, \cdots, x_r\} \quad : r=0, 1, \cdots \tag{29}$$

である。1年間に r 個発生した中の最大値 ξ が x を超えない条件は $x_1, x_2, \cdots, x_r$ のいずれもが x を超えないことと同等であるので，その確率は $[F_*(x)]^r$ で表示される。個数 r は 0 ～∞の任意の値を取り得るから，年間最大値 ξ の非超過確率は次のように求められる。

$$\text{Prob.}[\xi<x]=\sum_{r=0}^{\infty}P_r[F_*(x)]^r \tag{30}$$

この非超過確率はすなわち毎年最大値の分布関数 $F(x)$ に等しい。式(30)に式(28)を代入すると

$$\begin{aligned}F(x)&=\sum_{r=0}^{\infty}\frac{e^{-\lambda}}{r!}[\lambda F_*(x)]^r\\&=\exp\{-\lambda[1-F_*(x)]\}\end{aligned} \tag{31}$$

R年確率統計量 x_R に対する非超過確率は式(26)によって $F(x_R)=1-1/R$ で与えられるので，この条件を上式に代入すると $F_*(x_R)$ が次のようになる。

$$F_*(x_R)=1+\frac{1}{\lambda}\ln\left(1-\frac{1}{R}\right) \tag{32}$$

ここでRが1よりも十分に大きいものとして自然対数のべき級数展開を使うと次の近似式が得られる。

$$F_*(x_R)\fallingdotseq 1-\frac{1}{\lambda(R-1/2)} \tag{33}$$

この結果は式(27)と比べて再現期間が 1/2 年だけ短くなっている。逆にいえば，極大値資料の分布関数を使い，R年間に総計 λR 個の極大値が出現するものとしてR年確率統計量を推定する式(27)の方式の方が，毎年最大値資料を使う方式よりも 1/2 年に相当する分だけ大き目の値を算出することになる。この差異は安全側の誤差であり，しかもRが大きくなれば無視できる量である。したがって，毎年最大値資料と極大値資料の選択の問題は，再現期間RあるいはR年確率統計量 x_R の算出の上からは定まらないことになる。ただし，分布関数の母数や x_R の推定値の信頼度を高める上ではデータ個数が多いほど望ましく，この点では極大値資料の方が有利である。これについては **6.2** で論じる。

3. プロッティング公式の選択

3.1 順序統計量とプロッティング公式

(1) 順序統計量の確率密度関数

極値統計資料に対して最小2乗法によって分布関数のあてはめを行うときには，各データに対してそれぞれ何らかの非超過確率を割り付ける必要がある。これは図式推定法の場合に各データを確率紙上のどの位置にプロットするかということと同じ問題であり，プロッティング・ポジション（plotting position）の問題といわれている。

非超過確率の割り付けに当っては，まず取得された極値データを大きさの順番に並べ替える作業が行われる。統計の分野では，最小値を1番目とし，昇順に番号を付したデータを順序統計量と呼び，i 番目のデータを $x_{(i)}$ と表す。すなわち，

$$x_{(1)}\leq x_{(2)}\leq\cdots\leq x_{(i)}\leq\cdots\leq x_{(N)} \tag{34}$$

波浪データの解析では最大値を $x_{(1)}$ とし，降順に番号を付すことが多いので，こうした場合に対しては添字mを使うことにする。すなわち，

$$x_{(1)}\geq x_{(2)}\cdots\geq x_{(m)}\geq\cdots\geq x_{(N)} \tag{35}$$

昇順あるいは降順の順序統計量にどのような非超過確率を割り付けるべきかというのが統計学的に見たプロッティング・ポジションの問題である。このプロッティング・ポジションは，母集団の分布関数の特性によって規定される順序統計量の確率密度関数を考慮して定めるべき問題である。この順序統計量については，たとえば Rao の書物（奥野他訳）[26] や Borgman[23] では次のように解説されている。

今，分布関数 $F(x)$ を持つ確率変数の母集団から互いに独立なN個の観測値を抽出して一つの標本を作り，この観測値を昇順に並べ替えて順序統計量 $x_{(i)}$ の標本を得たと考える。この i 番目の順序統計量が任意の値 x を超えない確率 $P_{(i)}$ は，x_1～x_N のN個の観測値のうち i 個以上のデータが x を超えない確率に等しい。ところで，N個中 r 個のデータが x を超えない確率は

$$P_{r,N}(x)=\frac{N!}{r!(N-r)!}[F(x)]^r[1-F(x)]^{n-r} \tag{36}$$

である。したがって $x_{(i)}$ が x を超えない確率 $P_{(i)}$ は次式で与えられる。

$$\begin{aligned}P_{(i)}[x]&=\text{Prob.}[x_{(i)}\leq x]\\&=\sum_{r=i}^{N}\frac{N!}{r!(N-r)!}[F(x)]^r[1-F(x)]^{n-r}\end{aligned} \tag{37}$$

この級数和は不完全ベータ関数を用いて次のように書き換えることができる。

$$P_{(i)}[x]=\frac{N!}{i!(N-i)!}\int_0^{F(x)}t^{i-1}(1-t)^{N-i}dt \tag{38}$$

さらに，$x_{(i)}$ の確率密度関数は式(38)を x について微分することによって次のように求められる。

$$\begin{aligned}\varphi_{(i)}(x)=&\frac{N!}{(i-1)!(N-i)!}\\&\times[F(x)]^{i-1}[1-F(x)]^{N-i}f(x)\end{aligned} \tag{39}$$

このように順序統計量 $x_{(i)}$ の確率密度関数は，母集団の分布関数ならびに標本の大きさNと順位 i によって規定される。たとえば，確率変数が [0,1] の範囲に一様に分布するものを考えてみると，

$$\left.\begin{array}{ll} f(x)=1 & :0\leq x\leq 1 \\ F(x)=x & :0\leq x\leq 1 \end{array}\right\} \quad (40)$$

この確率変数の母集団から抽出された順位統計量 $x_{(i)}$ の確率密度関数 $\varphi_{(i)}(x)$ は式(39)により

$$\varphi_{(i)}(x)=\frac{N!}{(i-1)!(N-i)!}x^{i-1}(1-x)^{N-i} \quad (41)$$

確率密度関数が求められれば $x_{(i)}$ の期待値，標準偏差，ひずみ度などの統計量は容易に計算することができる。たとえば，期待値すなわち母集団平均は次のようになる。

$$E[x_{(i)}]=\int_0^1 x\varphi_{(i)}(x)\,dx=\frac{i}{N+1} \quad (42)$$

同様にして $x_{(i)}$ の2次，3次の積率を計算し，標準偏差およびひずみ度を求めた結果が以下である。

$$\sigma[x_{(i)}]=\frac{1}{N+1}\sqrt{\frac{i(N-i+1)}{N+2}} \quad (43)$$

$$\sqrt{\beta_1}=\frac{2(N-2i+1)}{N+3}\sqrt{\frac{N+2}{i(N-i+1)}} \quad (44)$$

このうち式(42)，(43)の結果は Gumbel[1] が示している結果（2.1.3節）と同じである。

ここで求めた [0,1] の区間で一様に分布する確率変数の代表例は，任意の確率変数の非超過確率Fそのものである。すなわち，どのような分布関数の母集団であっても，その中からランダムにN個抽出したときの最小値 $x_{(1)}$ が母集団において取る非超過確率 F_1 の平均は $1/(N+1)$，最大値 $x_{(N)}$ が母集団において取る非超過確率 F_N の平均は $N/(N+1)$ となる。ただし，式(44)から明らかなように F_1 および F_N は強いひずみ度を持った非対称分布である。このことは式(41)で F_1 および F_N の確率密度関数が次のように導かれることからも明らかである。

$$\left.\begin{array}{l} \varphi_{(i)}(F_1)=N(1-F_1)^{N-1} \\ \varphi_{(i)}(F_N)=NF_N{}^{N-1} \end{array}\right\} \quad (45)$$

(2) プロッティング・ポジションの考え方

順序統計量 $x_{(i)}$ が式(39)で与えられる確率密度関数を持つということは，一つの母集団から大きさがNの標本を多数抽出した場合，$x_{(1)}$～$x_{(N)}$ のそれぞれの値が標本ごとに変化し，その分布状態が式(39)で表されることを意味している。このことを例示するため，FT-I 型分布($A=1$，$B=0$)からランダムに10個のデータを抽出し，

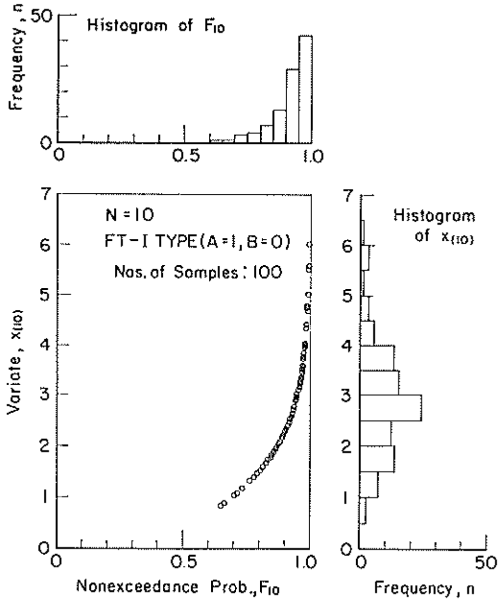


図-3　標本中の最大値の分布の例

これを100回繰り返し，各標本中の最大値 $x_{(10)}$ の分布を調べてみた結果が図-3である。左下の図の横軸は母集団における $x_{(10)}$の非超過確率 F_{10}である。当然のことながら $x_{(10)}$ と F_{10} とは式(4)の関係で結ばれており，一本の曲線上にプロットされている。上側の図は非超過確率 F_{10} の度数分布であり，右側の図は最大値 $x_{(10)}$ の度数分布である。F_{10} の度数分布は式(45)，$x_{(10)}$ の度数分布は式(39)の確率密度関数の形状にほぼ一致するはずである。

図-3で分るように，最大値は $x_{(10)}=0.84$～6.04 の範囲に広く分布している。標本数を増すと分布幅はさらに広がる。母集団が既知の場合には順序統計量 $x_{(i)}$ の挙動を的確に予測できるけれども，実際の問題ではただ1例入手できた標本から母集団の特性を推定しなければならない。したがって，$x_{(10)}=0.84$の標本に対しても，$x_{(10)}=6.04$の標本に対しても同一の手順によって推定作業を進めることになり，いずれの $x_{(10)}$ に対しても同一の非超過確率の推定値 $\hat{F}_{10}$ を与えなければならない。

標本中の順序統計量 $x_{(i)}$に対して割り付けるべき非超過確率 $\hat{F}_i$ の値については，少なくとも次の4通りの考え方[27]がある。

1) 母集団における F_i の平均 $E[F_i]$ を用いる。
2) 母集団における $x_{(i)}$ の平均 $E[x_{(i)}]$ に対応する確率 $F[E(x_{(i)})]$ を用いる。
3) 順序統計量 $x_{(i)}$ の分布の中央値（median）に対応

する確率を用いる。

4) 順序統計量 $x_{(i)}$ の分布の最頻値 (mode) に対応する確率を用いる。

図-3の例では標本数が 100 と少ないので2)～4)の差はあまり明瞭でないが，FT-I 型分布における $x_{(N)}$ の平均値，中央値および最頻値はそれぞれ次のように与えられる。ただし，式(20)の基準化変量 y を使って表す。

$$\left.\begin{aligned} y_{(N)\mathrm{mean}} &= \ln N+\gamma \\ y_{(N)\mathrm{median}} &= \ln N-\ln(\ln 2) \\ y_{(N)\mathrm{mode}} &= \ln N \end{aligned}\right\} \qquad (46)$$

$N=10$の場合について計算するとそれぞれ2.880，2.669および 2.303 となる。これに対応する非超過確率はそれぞれ0.9454，0.9330，および0.9048である。これに対して，第1の考え方によれば F_{10} の平均は式 (42) により $E[F_{10}]=0.9091$ である。

この例で明らかなようにプロッティング・ポジションの確率値は考え方によって相当に変化する。 Gumbel[1] は標本中の最大値の再現期間は標本の大きさNに漸近すべきであるという直感的前提の下に，第1の $E[F_i]$を用いることを推奨した。これによるプロッティング公式は次節に述べるワイブル公式（我が国ではトーマス公式と呼ばれることがある）である。この Gumbel の説に対して Cunnane[20] は厳しい批判を加え，プロッティング公式は第2の $F[E(x_{(i)})]$ によるべきであるとした。この第2の考え方は先に Kimball[27] が推奨したものであり，Petruaskas と Aagaard[15] によるワイブル分布に対するプロッティング公式の決定でも使われている。また，Carter と Challenor[19] が FT-I 型分布に対する母数推定法を数値的に比較検討したときもこの第2の考え方を採用している。

プロッティング・ポジションを与える公式は以上の4通りの考え方に基づくものだけでなく，別の方式で導かれるものもある。プロッティング公式の選択に関して，Kimball[27] は極値統計データを確率紙上にプロットする目的に応じて判断するのが良いとし，その目的として次の三つを挙げている。

1) 標本のデータが特定の分布関数を有する母集団に属するか否かを判断する。

2) 標本データの標準偏差を求める簡便法として利用する（正規確率紙上の勾配から算出)。

3) あてはめた分布関数の外挿によってR年確率統計量を推定する。

Cunnane[20]はさらにこのR年確率計量の推定に際しては，推定値の偏り (bias) がないことと推定値の分散が最小であることの二つをプロッティング公式の選定基準としている。この Cunnane の選定基準は妥当なものであり，本報告においてもこの基準を採択する。

3.2 既往のプロッティング公式

(1) ワイブル (Weibull) 公式

$$\hat{F}_i=\frac{i}{N+1}, \qquad \hat{F}_m=1-\frac{m}{N+1} \qquad (47)$$

ここに，i は昇順，mは降順の順位序数である。

これは式(42)で導いたように$0\leq x\leq 1$で一様に分布する確率変数の順序統計量の期待値である。本報告でも例示するように，このプロッティング公式を極値分布に適用すると推定値の偏りが著しく，不適切な結果を生む。このことは Carter と Challenor[19] による FT-I 型分布の数値的検討でも結論づけられている。

(2) ハーゼン (Hazen) 公式

$$\hat{F}_i=\frac{i-1/2}{N}, \qquad \hat{F}_m=1-\frac{m-1/2}{N} \qquad (48)$$

これはプロッティング公式として最も早いもので，Cunnane[20] の引用によれば 1914 年の提唱である。$i=1\sim N$のすべてのデータをプロットできるよう最も単純な形を選んだものと思われる。理論的根拠を欠くものであるので，本報告では比較検討の対象から除外する。

(3) ブロム (Blom) 公式

$$\hat{F}_i=\frac{i-3/8}{N+1/4}, \qquad \hat{F}_m=1-\frac{m-3/8}{N+1/4} \qquad (49)$$

これは Blom[28] が正規分布に対するプロッティング公式として推奨したものであるが，誘導法等については未調査である。Kimball[27] は標本が小さい場合に対する検討ではあるが，正規分布に対してはこのブロム公式が適当であるとしている。なお，対数正規分布に対しても正規分布と同様に適用可能である。

(4) グリンゴルテン (Gringorten) 公式

$$\hat{F}_i=\frac{i-0.44}{N+0.12}, \qquad \hat{F}_m=1-\frac{m-0.44}{N+0.12} \qquad (50)$$

これは Gringorten[29] が FT-I 型分布に対するプロッティング公式として導いたものである。考え方としては $F[E(x_{(i)})]$ を用いているが，順序統計量の全体に対してではなく，最大値 $x_{(N)}$ に対してのみ適用して導いている。まず，プロッティング公式の一般形として

$$\hat{F}_i=\frac{i-\alpha}{N+1-2\alpha} \qquad (51)$$

を仮定する（このようにすると昇順，降順の両者とも同形式の公式とすることができる)。$i=N$に対する順序統計量の平均値は先に式(46)で求めてあるので，それを式(4)の分布関数に代入すると

$$\hat{F}_N=\exp\{-\exp[-(\ln N+\gamma)]\} \qquad (52)$$

この結果を式(51)に代入し，$N=10(10)100$ の範囲で α の値を求めることによって，$N \geq 20$の範囲において小数点2桁の精度で $\alpha=0.44$の結果を得たものである。

(5) ペトルアスカス・アーガード公式（P&A 公式）

$$\hat{F}_i=\frac{i-1+(\alpha'+\beta)}{N+\beta}, \quad \hat{F}_m=1-\frac{m-\alpha'}{N+\beta} \qquad (53)$$

ここに，

$$\left.\begin{aligned}\alpha'&=0.30+0.18/k\\ \beta&=0.21+0.32/k\end{aligned}\right\} \qquad (54)$$

この公式は Petruaskas と Aagaard[15] が式(13)のワイブル分布の順序統計量に対して導いたものであり，k はワイブル分布の形状母数である。この公式の場合は昇順に対するものと降順に対するものとで分子の定数が異なる。

Petruaskas と Aagaard はまず規準化した順序統計量 $y_{(m)}$ の確率密度関数を式(39)で求め，$y_{(m)}$ の期待値をその1次の積率の計算によって次のように導いている。

$$E[y_{(m)}]=\frac{N!}{(m-1)!}\Gamma\left(1+\frac{1}{k}\right)\times\sum_{r=0}^{N-m}\frac{(-1)^r}{r!(N-m-r)!(m+r)^{1+1/k}} \qquad (55)$$

この計算においては2項展開による項別積分を行っている。上式は原論文の表現とやや異なるが，実質的には同一である。この期待値に対する非超過確率を解析的に与えることは不可能なので，Petruaskas と Aagaard はプロッティング公式として式(53)の形を仮定した上で次の2乗残差が最小になるように幾つかの形状母数 k 値に対して α と β を選定した。

$$E_r=\sum_{m=1}^{N}\left\{F[E(y_{(m)})]-1+\frac{m-\alpha}{N+\beta}\right\}^2 \qquad (56)$$

そして，α および β の値を k に対してプロットして関数形を式(54)のように定めたものと考えられる。

(6) バーネット（Barnett）の方式

$$\left.\begin{aligned}\hat{F}_i&=\exp[-\exp(-y_{(i)})]\\ y_{(i)}&=3(\ln 2)\left(\frac{2i}{N+1}-1\right)+\gamma\end{aligned}\right\} \qquad (57)$$

この公式はその形から推測されるように FT-I 型分布に対して Barnett[30] が導いたもので，順序統計量に対して分布関数をあてはめたときの2乗残差が最小になるように定めたとしているが，理論的誘導は難解である。Carter と Challenor[19] の数値シミュレーション結果ではグリンゴルテン公式よりも良好な結果を示している。ただし，数値を代入してみると分るように $x_{(i)}$ と $y_{(i)}$ の関係は直線ではなく，逆S字状になるため，$x_{(i)}$ と $y_{(i)}$ の相関係数の値が低い。このため，標本のデータが FT-I 型分布に適合しているか否かを検討する目的には不向きである。

(7) カーター・チャレノアの方式

$$\hat{F}_i=F[E(y_{(i)})]$$

$$\left.\begin{aligned}E[y_{(i)}]=&\frac{N!}{(i-1)!}\sum_{r=0}^{N-i}\frac{(-1)^r}{i!(N-i-r)!}\\ &\times\frac{\gamma+\ln(i+r)}{i+r}\end{aligned}\right\} \qquad (58)$$

この方式も FT-I 型分布に対するもので，Carter と Challenor[19] が式(39)の確率密度関数の公式を用いて $y_{(i)}$ の期待値を計算した結果である。

(8) その他

Cunnane[20] は以上の他にも幾つかのプロッティング公式を引用しており，自分も式(51)の形式で $\alpha=2/5$ と置くとどのような分布関数に対してもほぼ妥当な結果を与えるとしている。ただし，いずれも簡便な近似式であり，ブロム公式やグリンゴルテン公式と大差ないので本報告では検討の対象から除外する。

3.3 順序統計量の数値シミュレーションに基づくプロッティング公式の誘導

プロッティング・ポジションを順序統計量の期待値に基づいて求めるには，ワイブル分布であれば式(55)，FT-I 型であれば式(58)の級数計算を行わなければならない。標本の各データについてこのような計算を行うのは面倒であり，式(51)あるいは(53)のような簡易式が導くことができるならば便利である。式(53)，(54)のペトルアスカス・アーガード公式はそうした公式化の1例である。そこで，本報告では極値統計に関する数値的検討の第1段階として，順序統計量についてモンテカルロ法による数値シミュレーションを実施した。

計算手順は次のとおりである。

1) [0,1] の区間における一様乱数 ζ を N 個発生させる。
2) 対象とする分布関数について，その値がそれぞれ $\zeta_1, \cdots \zeta_N$ に等しいときの規準化変量 $y_1, \cdots y_N$ の値を分布関数の逆関数 $F^{-1}(\zeta_i)$ として求める。
3) y_i を大きさの順に並べ替え，順序統計量 $y_{(i)}$ を得る。
4) 以上の操作を多数回繰り返し，$y_{(i)}$ の平均，標準偏差，ひずみ度などの統計量を計算する。
5) 標本の大きさ N を幾つか変え，N による $y_{(i)}$ の統計量の変化を調べる。

先に提示した図-3はこうした数値シミュレーションの

表-1 順序統計量の最大値に関する数値シミュレーション結果と理論値との比較
(FT-I 型分布：標本数10,000)

標本の大きさ N	非超過確率 F_N		規準化順序統計量 $y_{(N)}$		$F[\bar{y}_{(N)}]$
	平均	標準偏差	平均	標準偏差	
10	0.9085(0.9091)	0.0824(0.0830)	2.868(2.880)	1.283(1.283)	0.9447
20	0.9519(0.9524)	0.0455(0.0454)	3.562(3.573)	1.301(1.283)	0.9720
30	0.9677(0.9677)	0.0317(0.0312)	3.981(3.978)	1.288(1.283)	0.9815
40	0.9758(0.9756)	0.0237(0.0238)	4.282(4.266)	1.301(1.283)	0.9863
50	0.9805(0.9804)	0.0191(0.0192)	4.493(4.489)	1.300(1.283)	0.9889
60	0.9837(0.9836)	0.0160(0.0161)	4.689(4.672)	1.310(1.283)	0.9908
80	0.9876(0.9876)	0.0123(0.0122)	4.959(4.959)	1.292(1.283(	0.9930
100	0.9902(0.9901)	0.0098(0.0098)	5.197(5.812)	1.291(1.283)	0.9945

注：括弧内は理論値（式. 42, 43, および46）

1例である。ただし，100回程度の繰り返しではまだ標本数が不足であるので，順序統計量に関してはすべて10,000回の繰り返し計算を行った。表-1は分布関数がFT-I 型の母集団から抽出した標本について，各標本中の最大値の大きさ $y_{(N)}$ およびその非超過確率 $\zeta_{(N)}=F_N$ の平均と標準偏差の値を示したものである。括弧内の数値は理論値であり，F_N については式(42)と(43)，$y_{(N)}$ の平均値は式(46)による。$y_{(N)}$ の標準偏差は分布関数自身の標準偏差と等しく $\pi/\sqrt{6}$ であることが証明できる。

表-1 の結果を見ると，$y_{(N)}$ の標準偏差が理論値よりも最大2%ほど大きいケースが見られるものの，$y_{(N)}$ の平均値は理論値との差が±0.4%以下，F_N については理論値との一致がさらに良好であって，数値シミュレーションが適正に実施されたことを示している。

なお，表-1 の最右欄は $y_{(N)}$ の平均値に対応する分布関数の値，すなわち $F[E(y_{(N)})]$ に相当するものであり，プロッティング公式を導く基礎となるものである。

表-1にはN個中の最大値のみを出しているが，$i=1$～Nについて $F[\bar{y}_{(i)}]$ の値を求めて順位序数 i に対してプロットすると図-4のような結果が得られる。この例では $F[\bar{y}_{(i)}]$ と i とはほとんど1本の直線上に乗っているが，他の分布関数の場合についても直線式で近似することが可能であった。すなわち，一般に

$$F[\bar{y}_{(i)}]\fallingdotseq ai+b \tag{59}$$

ここで，$F[\bar{y}_{(i)}]=\hat{F}_i$ と置き，プロッティング公式として次式を仮定する。

$$\hat{F}_i=\frac{i-\alpha}{N+\beta} \tag{60}$$

そうすると，式(59)の形であてはめた係数 a，b を用いてプロッティング公式の定数 α，β を次のように求めることができる。

$$\alpha=-b/a, \qquad \beta=1/a-N \tag{61}$$

図-4の $N=20$ の例では最小2乗法による推定値が $a=0.04948$，$b=-0.01563$ であるので，$\alpha=0.316$, $\beta=0.210$ の結果が得られる。なお，式(59)の係数を最小2乗法で定めることは，Petruaska と Aagaard が行った式(56)の2乗残差の最小化と同等である。

以上の方法によって $N=10$～100の範囲で FT-I 型分布および正規分布についてプロッティング公式の定数 α と β を求めた結果が表-2である。標本の大きさによって若干の変動は見られるが，この程度であれば平均値を採用しても良いと思われるので，順序統計量の平均値に基づくプロッティング公式を次のように定めることができる。

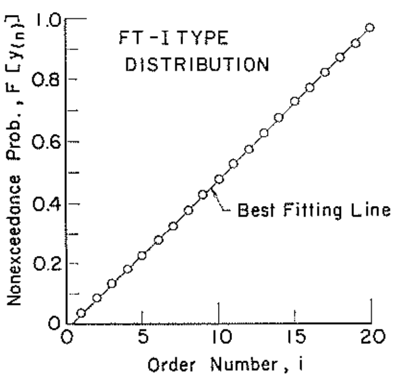


図-4 順序統計量の期待値の確率のシミュレーション結果の例

表-2 プロッティング公式の定数の算出結果

標本の大きさ N	FT-I 分布		正規・対数正規分布	
	α	β	α	β
10	0.296	0.239	0.380	0.259
20	0.316	0.210	0.400	0.233
30	0.290	0.202	0.374	0.225
40	0.302	0.161	0.386	0.182
50	0.307	0.198	0.404	0.178
60	0.309	0.149	0.413	0.172
80	0.332	0.140	0.389	0.235
100	0.299	0.193	0.408	0.223
平　均	0.308	0.179	0.394	0.213

表-3 ワイブル分布に対するプロッティング公式の定数の平均値

形状母数 k	順序統計量による値		P&A 公式の値	
	α'	β	α'	β
0.50	0.653	0.848	(0.66)	(0.85)
0.60	0.596	0.736	(0.60)	(0.74)
0.75	0.537	0.648	0.54	0.64
0.85	0.502	0.584	0.51	0.59
1.00	0.480	0.533	0.48	0.53
1.10	0.481	0.471	0.46	0.50
1.25	0.457	0.442	0.44	0.47
1.50	0.416	0.402	0.42	0.42
2.00	0.403	0.344	0.39	0.37
2.50	0.361	0.351	(0.37)	(0.34)
0.00	0.370	0.315	(0.36)	(0.32)

注：P&A 公式で括弧内は式 (54) による計算値。

FT-I 型分布：

$$\hat{F}_i=\frac{i-0.31}{N+0.18} \tag{62}$$

正規分布および対数正規分布：

$$\hat{F}_i=\frac{i-0.39}{N+0.21} \tag{63}$$

既往の公式と比べると，FT-I 型分布についてはグリンゴルテン公式との違いが目立つが，正規分布についてはブロム公式との差異が小さい。したがって，以下の検討では式 (63) は取り上げず，ブロム公式を使用する。なお，対数正規分布では極値データを $\ln x$ の形に変換した上で正規分布にあてはめるので，プロッティング公式は正規分布に対するものと同一のものを使うべきである。

ワイブル分布については $k=0.5\sim3.0$ の分布形状11種類を取り上げ，$N=10\sim100$ について同様の作業を行った。これによって得られた α と β の結果を $N=10\sim100$ のデータの平均として示したのが表-3である。また，全データを図示したのが図-5，6である。ただし，ペトルアスカス・アーガード公式との比較のため降順の序数 m に関する定数 α' として表示してある。標本の大きさ N による変動はかなり大きいが，特定の傾向は認められない。形状母数と定数 α'，β との関係は，大局的には Petruaskas と Aagaard が提示した式(54)によってほぼ表示できる。なお，図中の破線は次節に述べる再現確率統計量の数値シミュレーションにおいて偏り (bias) をできるだけ小さくするように修正した値を示している。

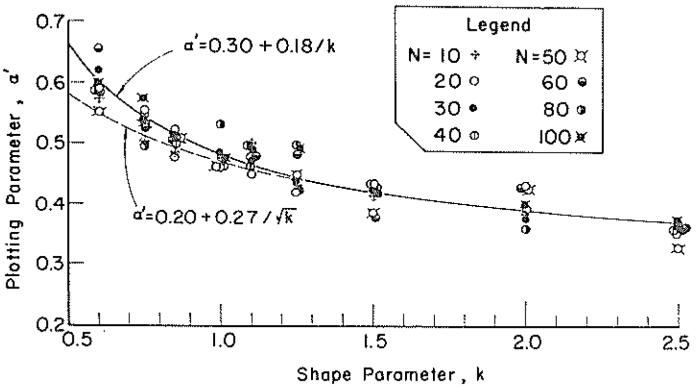


図-5 ワイブル分布のプロッティング公式のパラメータ α'

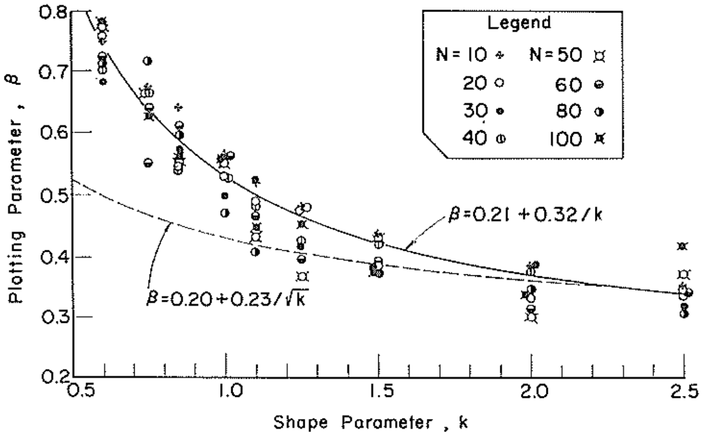


図-6　ワイブル分布のプロッティング公式のパラメータ β

3.4 再現確率統計量の数値シミュレーションによるプロッティング公式の比較

(1) 概　要

プロッティング公式の適否は，Cunnane[20]が述べた推定値の不偏倚性およびその分散の最小性によって判断される*。Carter と Challenor[19] は FT-I 型分布を対象として，最小2乗法，積率法，最尤法，ジャックナイフ法その他の分布関数適合法の優劣をモンテカルロ法を用いて検討している。標本の大きさとしては$N=10$，20，および40を対象とし，各条件に対して500回ずつの繰り返し計算である。しかし，この試行回数では標本数として不十分と考えられるので，本報告では表-4の組み合せについて10,000回ずつのシミュレーションを実施した。

数値的検討の具体的方法は次のとおりである。

1) [0,1] の一様乱数 ζ_i をN個発生させ，非超過確率が ζ_i に等しい確率変数 x_i を次式により計算する。

$$x_i=AF^{-1}(\zeta_i|k)+B \tag{64}$$

ここに，A，B，kは対象とする分布関数の母数としてあらかじめ設定した定数であり，F^{-1} は逆関数を表す。

2) x_i を大きさの順に並べ替えて順序統計量 $x_{(i)}$ を作成する。

3) あらかじめ選定したプロッティング公式により，$x_{(i)}$ のプロッティング確率 $\hat{F}_i$ を計算する。

4) $\hat{F}_i$ に対応する基準化変量 $y_{(i)}$ を次式で求める。

$$y_{(i)}=F^{-1}(\hat{F}_i) \tag{65}$$

5) $x_{(i)}$ と $y_{(i)}$ の間には直線関係が存在するはずであるから，最小2乗法によって定数の推定値 $\hat{A}$ および $\hat{B}$ を求める。すなわち，

$$x_{(i)}=\hat{A}\,y_{(i)}+\hat{B} \tag{66}$$

6) 上で得られた $\hat{A}$ および $\hat{B}$ を分布関数に代入し，所定の再現期間Rに対応するR年確率統計量の推定値 $\hat{x}_R$ を後出の **4.3** の式 (73)～(76) によって求める。

7) 対象とする分布関数におけるR年確率統計量の真値 x_R は既知であるので，$\hat{x}_R/x_R$ の比率等を計算する。

8) 以上のシミュレーション作業を10,000回繰り返し，$\hat{x}_R/x_R$等の平均値，標準偏差値等を算出する。

なお，標本の大きさは$N=10$，14，20，30，40，60，および100としたが，FT-I 型については$N=200$および300についても検討した。この数値シミュレーションは

* 神田・藤田[11]はこの二つを不偏性および有効性と呼んでいる。

表-4　再現確率統計量の数値シミュレーションの組み合せ

プロッティング公式	分布関数		
	FT-I型	対数正規	ワイブル7種類
ワイブル公式	○	○	○
ブロム公式	—	○	—
グリンゴルテン公式	○	—	—
新公式（式. 62)	○	—	—
バーネット方式	○	—	—
P&A 公式	—	—	○
修正 P&A 公式	—	—	○

注：標本の大きさは$N=10$～100.

プロッティング公式の優劣を比較すると同時に，母数AおよびBならびにR年確率統計量 x_R の推定値の信頼区間のデータを取得するために実施したものである。信頼区間については4.で述べるので，本節ではプロッティング公式の比較結果のみを紹介する。

(2) FT-I 型分布

ここでの検討では$R=(2\sim100)N$の再現期間に対する再現確率統計量を算出した。$\hat{x}_R$ の推定誤差は，再現期間が長くなるにつれて増大し，一方，標本の大きさNが増すにつれて減少する。ここでは，実際上の上限と想定される$R=10N$の再現期間における結果を図-7，8に示す。

図-7 は（$\hat{x}_R/x_R-1$）の平均値を百分率で表したもので，R年確率統計量の推定値が平均的に真値から偏っている度合を表す。図-8は（$\hat{x}_R/x_R-1$）の RMS 値（自乗平均方平根値）を百分率で表示している。なお，分布関数の定数としては次の値を使用した。

$$\left.\begin{array}{l}\text{FT-I 型分布：} A=1.0,\ B=5.0\\ \text{対数正規分布：} A=0.16118,\ B=1.49668\\ \text{ワイブル分布}(k=0.75\sim2.0)\text{：} A=1.0,\ B=5.0\end{array}\right\} \quad (67)$$

ここに示した平均値の偏り量および RMS 値の絶対量はこうした定数の大きさによって影響されるけれども，プロッティング公式ごとの相対差は変らないはずである。

まず，図-7を見るとワイブル公式を使うとR年確率統計量が平均的に大きく正の方向（真値よりも大）へ偏ることが明らかである。グリンゴルテン公式とバーネット方式は偏り量がほとんど0であり，この点では優れたプロッティング公式である。一方，順序統計量の平均値に基づいて定めた式(62)の新公式は，標本中のデータ数が少ないときは負の偏り量を示し，あまり適切なプロッティング公式とはいえない。

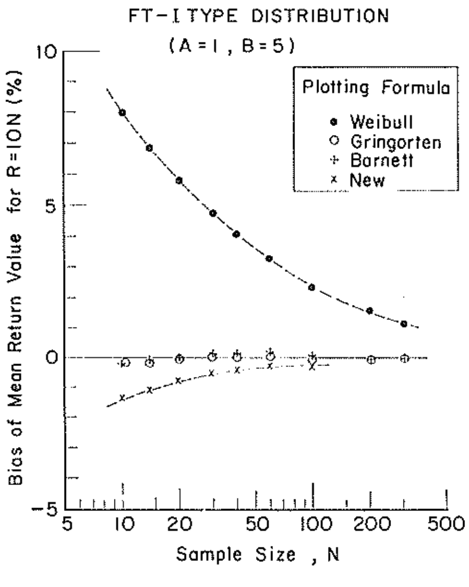


図-7 FT-I 型分布における再現確率統計量の推定値の偏り量

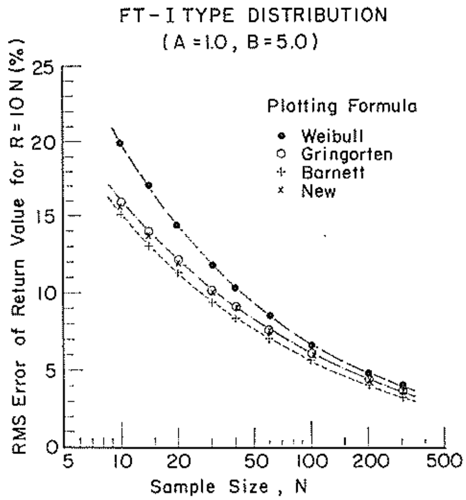


図-8 FT-I 型分布における再現確率統計量の RMS 誤差

次に，（$\hat{x}_R/x_R-1$）の RMS 値を見ると，平均値の偏り量の場合と同様にワイブル公式を用いたときが最大の RMS 値を示し，推定値のばらつきが大きい。この RMS 値は

$$\mathrm{RMS}[X]=\{\mathrm{Var}[X]+E[X]^2\}^{1/2} \qquad (68)$$

と変量Xの分散に加えて平均値のずれも取り込んで計算しているので，標準偏差を用いると若干小さくなるけれども，ワイブル公式を用いたときのばらつきが大きいことに変りはない。バーネット方式はグリンゴルテン公式よりも RSM 値が小さく，FT-I 型分布の母数推定の精度が良いことを示している。ただし，ここには提示していないがバーネット方式では $x_{(i)}$ と $y_{(i)}$ との相関係数が低く出る傾向がある。極値統計の母分布関数が FT-I 型であると断定できる場合にはバーネット方式でプロッティング・ポジションを計算し，最小2乗法によって母数AおよびBを推定する方法が最も実用的と考えられる。しかし，対象とする極値統計の母分布関数が未知であり，幾つかの分布関数をあてはめた結果に基づいて最適合関数を選定するような場合（これが通常のケースである）には，相関係数の値が低く出ることは不利であ

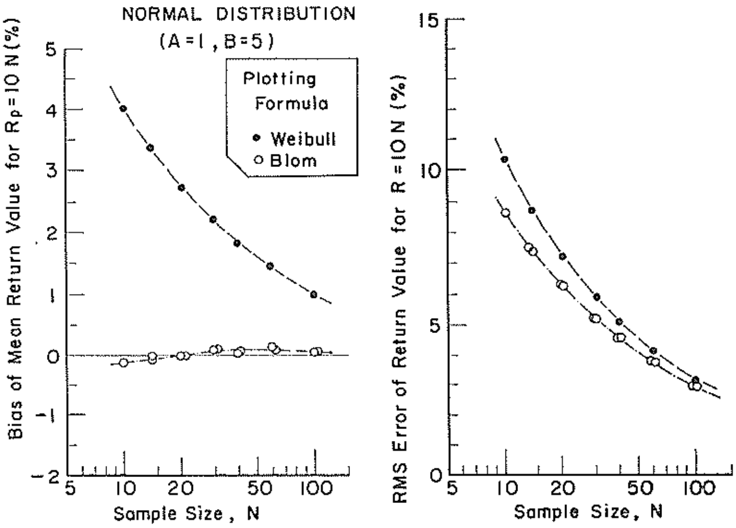


図-9 正規分布における再現確率統計量の推定値の偏り量と RMS 誤差

る。そのような場合にはグリンゴルテン公式を用いるのが適切と考えられる。

式(62)の新公式はグリンゴルテン公式よりも RMS 値が僅かながら小さく出ている。しかしながら，図-7に示した平均値の偏り量が負となる短所を勘案すると，この程度の RMS 値の差ではグリンゴルテン公式よりも優るということはできない。したがって，以下の検討ではFT-I 型分布関数のあてはめにおいてはグリンゴルテンによるプロッティング公式を用いることとした。

(3) 正規分布および対数正規分布

図-9は正規分布から抽出した標本に対してワイブル公式およびブロム公式を適用してR年確率統計量を求めた結果である。図の左側が（$\hat{x}_R/x_R-1$）の平均値の偏り量，右側はその RMS 値である。ブロム公式を用いたときのデータはたまたまシミュレーションを2回行った（各 10,000 回の繰り返し）ので，$N\geq14$はデータが2個ずつ得られている。

二つのプロッティング公式を比べると，ワイブル公式の不具合さが明瞭であり，正規分布および対数正規分布に対してはブロム公式を使うべきであると結論される。なお，対数正規分布に対してプロッティング・ポジションをワイブル公式で定めると再現確率統計量の推定値に正の偏り量が発生する事例は，Earle と Baird[31] の数値シミュレーション結果に見ることができる。この研究では，波浪の観測・推算値の誤差および標本データの統計的変動性が再現確率統計量の推定結果に及ぼす影響を調べるため，母分布関数として対数正規分布を仮定し，異常高波浪海域，高波浪海域，および低波浪海域に対応して定数AおよびBの値を3通りに設定した。標本の大きさは$N=5$, 10, 20, および40とし，各条件に対して1000回ずつの繰り返しシミュレーションを行って再現期間5～100年に対する確率波高を求めた。この結果，観測誤差が0の場合でも推定確率波高の平均値として1～15%も真値より大きい値を得，原著者は「少なくとも安全側の誤差である」と結論づけている。

しかし，Earle と Baird[31]で注意すべき点はプロッティング・ポジションをワイブル公式で計算していることである。著者は二人の使ったA，Bの定数値を用いてシミュレーション結果を追試験してみた*。その結果，ワイブル公式を使うとほとんど同一規模の正の偏り量および RMS 誤差を得られるけれども，同じデータに対してブロム公式を適用すると再現確率統計量の偏り量がほとんど0となることが確認された。このことは，正規分布および対数正規分布に対するワイブル公式の不適格性を例証するものであるとともに，Earle と Baird の結論が不正確であることを意味する。

(4) ワイブル分布

ワイブル分布については，形状母数kを0.75～2.0の範囲でいろいろ変えてプロッティング公式を比較した。なお，このkの範囲は Petruaskas と Aagaard[15] が波浪の極値統計解析において用いたものである。**図-10, 11, 12**はこのうち$k=0.75$，1.0および2.0のケースについて

* 式(67)の対数正規分布に対する母数の値は Earle と Baird の与えた低波浪海域に対する値を換算したものである。

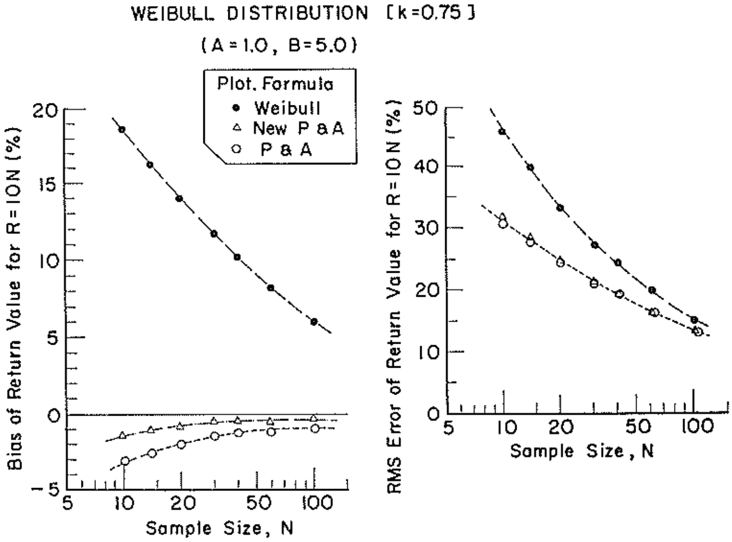


図-10 ワイブル分布（$k=0.75$）における再現確率統計量の推定値の偏り量と RMS 誤差

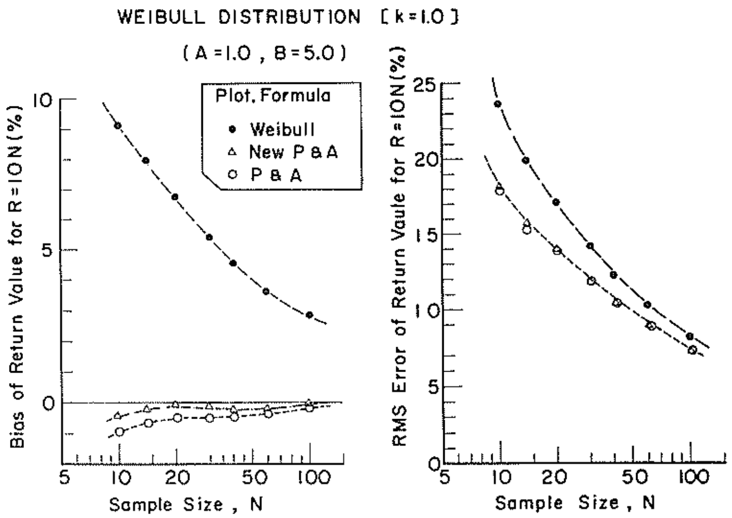


図-11 ワイブル分布（$k=1.0$）における再現確率統計量の推定値の偏り量と RMS 誤差

の結果を示したものである。

FT-I 型分布や正規分布の場合と同様に，ワイブル公式は再現確率統計量の偏り量が大きく，過大な推定値を与える傾向が著しい。特に形状母数が小さいときは偏り量が顕著であり，$k=0.75$，$N=10$ では 19%に近い値を示す。これに対して P&A 公式は偏り量が負となる傾向がある。この P&A 公式は順序統計量の期待値に基づいて導かれたもの（図-5，6 参照）にもかかわらず，こうした不具合なところが現われるのはやや奇異に感じられる。しかし，FT-I 型分布について図-7 で示したように，順序統計量の期待値に基づいた式(62)のプロッティング公式よりもバーネット公式の方が成績が良い例もある。このことは，順序統計量の期待値に基づくプロッティング・ポジション $F[E(x_{(i)})]$ は，**3.1**(2)に述べた他の方法によるものよりも優れた結果をもたらすけれども最善のものとは必らずしもいえず，他にさらに優れたものがあり得ることを示唆している。

ワイブル分布についてはこのような観点から P&A 公

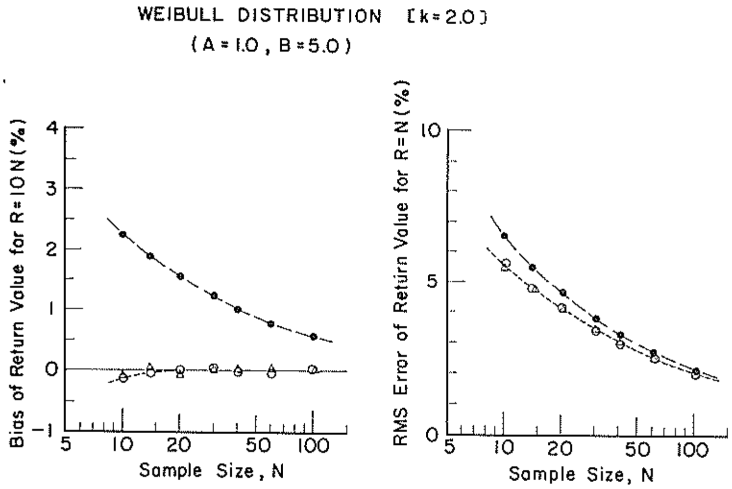


図-12 ワイブル分布 ($k=2.0$) における再現確率統計量の推定値の偏り量と RMS 誤差

式を若干修正して再現確率統計量の負の偏り量を減らすことをくふうした。二，三の試行錯誤の結果，本報告では式(53)の定数 α', β として次のような値を用いることとした。

$$\left.\begin{aligned}\alpha' &= 0.20+0.27/\sqrt{k}\\ \beta &= 0.20+0.23/\sqrt{k}\end{aligned}\right\} \tag{69}$$

なお，昇順の場合は次のような形になる。

$$\left.\begin{aligned}\hat{F}_i &= \frac{i-\alpha}{N+\beta}\\ \alpha &= 0.60-0.50/\sqrt{k}\end{aligned}\right\} \tag{70}$$

β は式(69)と同一である。

図-10～12 には式 (69) の定数を用いた結果を新 P&A 公式として記入してある。$k=0.75$ の場合は負の偏り量が P&A 公式の半分以下に減少しており，$k=1.0$ でも若干の減少が見られる。$k=2.0$ の場合はもともとの偏り量が小さいこともあって有意な差は出ていない。なお，式(69)による値と式(54)による値とは $k\fallingdotseq 2.2$ でほぼ一致する。RMS 値に関してはほとんど差がないが，$k=0.75$ では新公式の方が僅かながら RMS 値を増大させる傾向が出ている。再現確率統計量の偏り量をさらに減少させることは不可能ではないが，式(69)の値自体が 図-5, 6 に示したように順序統計量の期待値から導いた値のほぼ下限値に相当することや，あまり強い修正を行うと再現期間の短いところで偏り量が大きくなるなどの欠点が出ることなどを勘案し，図-10～12 の結果で実用上は満足できるものと判断した。

3.5 部分極値資料に対するプロッティング公式

先に 2.1 で述べたように，極値統計のデータには対象とすべき資料のすべてを収録した全数極値資料と，対象とすべきものの一部しか収録していない部分極値資料の2種類が存在する。高波などのデータでは後者の場合が多い。こうした部分極値資料に対するプロッティング公式としては二つの考え方がある。一つは Petruaskas と Aagaard[15] によるもので，あらかじめ設定した限界値以下のデータの存在を無視し，得られた部分極値資料をあたかも全数極値資料であるかのように見なしてプロッティング・ポジションを計算する方法である。すなわち，プロッティング公式における N としては，得られた標本の大きさをそのまま使うものである。

もう一つは Muir と El-Shaarawi[7] が提示しているもので，N として対象期間中に発生したと考えられる極値の総数 N_T を使い，プロッティング・ポジション $\hat{F}_i$ は実際に収録できた上位 N 個についてのみ算出する方法である。すなわち，

$$\left.\begin{aligned}\hat{F}_i &= \frac{i-\alpha}{N_T+\beta} &&: i=N_T-N+1, \cdots, N_T\\ \hat{F}_m &= 1-\frac{m-\alpha'}{N_T+\beta} &&: m=1,\ 2, \cdots, N\end{aligned}\right\} \tag{71}$$

前者の方法は単純であって使いやすいため，我が国でも著者がこれを紹介[16]して以来，波浪の極値解析では広く用いられている。しかし，後者の方法と比べると分布関数の形状をひずめて推定するという難点がある。いま図-13 に示したような確率密度関数の形を考えてみると高波資料のように足切りをしたデータでは $f(x)$ の左側の一部分が切り落されたデータを入手したことに相当する。Petruaskas と Aagaard の方法では，設定された限

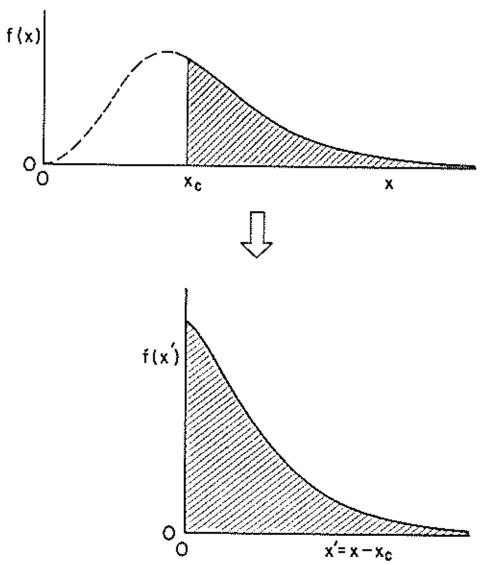


図-13 部分極値資料における確率密度分布関数の形状

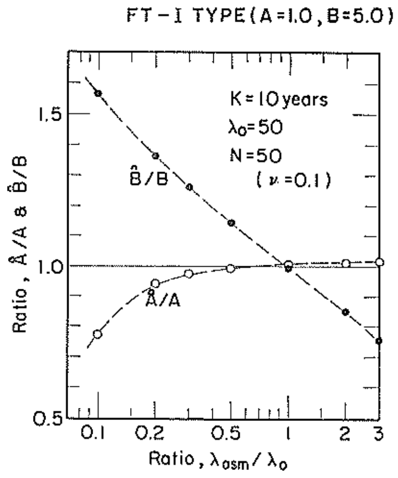


図-14 平均発生率の推定誤差による母数推定値の変化

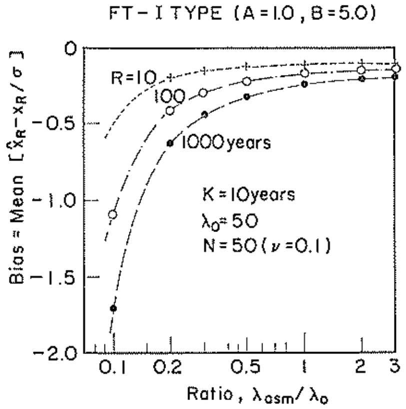


図-15 平均発生率の推定誤差による再現確率統計量の推定値の偏り量の変化

界波高 x_c 以上のデータをあたかもそれが全体であるかのように取り扱うので，図の上部の確率密度関数は図の下部のような形状と見なされることになる。$k\leq1$ のワイブル分布は図-2に示したようにもともと右下りの形をしているので，図-13 のような処理を行っても母分布関数に近い形を保持するけれども，図-1のグループの場合には元の形とはかけ離れてしまい，したがって母関数の復元が困難である。

後者の式(71)による方法の場合にはこのような分布関数の変形の問題は生じない。しかし，極値の総数 N_T をどのようにして正しく把握するかという問題が新しく生じる。期間最大値資料の場合には対象となる期間が固定されているので，その個数は自明である。これに対して極大値資料の場合，特に自然現象に係わる極値の問題では極値の対象となる事態をどのように定義するかによって極値の総数 N_T も異なってくる。

N_T の推定問題は，別の見方からすればどれほどの正確さが必要かということである。このことを調べるため，次のような条件で数値シミュレーションを行い，N_T の推定精度によって分布関数の母数および再現確率統計量の推定値が変化する状況を検討した。

1) 母分布関数：FT-I 型分布
2) 母数　　　：$A=1.0,\quad B=5.0$
3) 対象期間　：$K=10$年
4) 平均発生率：$\lambda_0=50$ ($N_T=500$)
5) 想定発生率　　：$\lambda_{asm}=5\sim150$
6) 標本データ個数：$N=50$ ($\nu=0.1$)
7) 繰り返し回数　：1000回

なお，解析を行った標本データは λ_{asm} のすべてのケースについて共通である。また，データに対するあてはめ関数としては母関数と同じ FT-I 型を使用した。

まず，発生率 λ の推定誤差による母数推定値の変化は図-14のようになる。λ を低く推定し，極値の総数 N_T を少なく見積ると，下位の極値データはプロッティング・ポジション $\hat{F}_i$ が低く計算される。この結果式(65)の基準化変量 $y_{(i)}$ が小さくなる。極値データの下限値は変らないので，$y_{(i)}$ が小さいままで同一の $x_{(i)}$ が得られるため

には $\hat{B}$ の値が大きくなければならない。この結果，位置母数の推定値 $\hat{B}$ は λ_{asm}/λ_0 が減少するにつれて増大することになる。尺度母数の推定値 $\hat{A}$ は比較的変化が少ないけれども，λ_{asm}/λ_0 の減少につれて減少する傾向を伴う。

次に，再現確率統計量の推定値 $\hat{x}_R$ については，真値 x_R との差を標本データの標準偏差 σ_x で除して無次元化した量 $(\hat{x}_R-x_R)/\sigma_x$ についてその平均値を求め，これを $\hat{x}_R$ の偏り量として定義した結果を図-**15** に示す。ここには再現期間が $R=10$，100，および 1000 年の結果を示しており，R の増加につれて $\hat{x}_R$ の偏り量が大きくなることは予想されるとおりである。また，極値の発生率 λ を小さく見積るにつれて偏り量が増大し，特に $\lambda_{asm}\leq 0.2\lambda_0$ でこの傾向が著しい。しかしながら，発生率 λ の推定値が真値の 0.5 倍以上であれば再現確率統計量の推定誤差はかなり小さくなるといえる。

以上の検討は FT-I 型分布のみを取り上げたものであり，必らずしも一般的な結論とはいえないかもしれないが，実用的観点からすれば部分極値資料に対して式(71)を用いてプロッティング・ポジションを計算する際には極値の総数 N_T はおおよその値を与えればよいと判断される。目安としては，真値の 0.5 倍以上あればよく，やや大き目の値を用いれば無難と思われる。したがって，本報告の次章以降の検討では式(71)の方法を用いてプロッティング・ポジションを計算することとした。

4. 分布関数が既知の場合の母数および再現確率統計量の推定値の信頼区間

4.1 極値統計における標本資料の標準偏差

先に図-**3**で例示したように，特定の分布関数を有する母集団から比較的少数のデータを標本として抽出した場合には，標本ごとにデータの値が大きく変動する。このため，一つの標本に対して分布関数をあてはめ，その結果に基づいて再現確率統計量を推定しても，その値は必らずしも母集団の値と一致しない。我々にできることは，標本ごとの推定値を多数集めたときにその平均値が母集団の値に近ずく（不偏倚性）ように解析手法をくふうすることである。標本ごとにあてはめた分布関数がばらつくことについては，既に1952年に Benson[32] が例示している。

極値統計資料の標本ごとの変動の度合は，標本資料の標準偏差値の変動状況から推測することができる。ここでは標準偏差として次式の不偏推定値を用いる。

$$\sigma_x=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2} \tag{72}$$

各分布関数の標準偏差は式(7)，(12)，および(16)で計算される。少数データの標本の標準偏差に関しては，FT-I 型分布についての Gumbel[1]（6.2.3節）および角屋[33]の計算結果がある。ただし，$\nu<1$ の部分極値資料の標準偏差については検討されていないようである。

本報告ではプロッティング公式の数値的検討において各条件ごとに10,000組の全数極値資料の標本を各母集団から抽出しているので，この資料について σ_x の平均値 $\bar{\sigma}_x$ およびその変異係数 C.V.$[\sigma_x]$ を計算した結果を表-**5**に掲載した。ここにはデータ採択率が $\nu=0.5$ および $\nu=0.25$ の部分極値資料の場合についてそれぞれ 10,000 組の標本について検討した結果も載せている。この場合，平均発生率は $\lambda=20$および40に設定した。なお，$\bar{\sigma}_x$ については使用した母数の絶対値の影響を避けるため，理論値 σ_0 に対する比率で示してある。また変異係数は 10,000個の標本から計算した σ_x の標準偏差 $\sigma[\sigma_x]$ を平均値 $\bar{\sigma}_x$ で除して求めている。すなわち，

$$\mathrm{C.V.}[\sigma_x]=\sigma[\sigma_x]/\bar{\sigma}_x \tag{73}$$

表-**5**によると，全数極値資料である $\nu=1.0$ の場合は $\bar{\sigma}_x/\sigma_0$ の比が 1 よりも僅かに小さく，標本の大きさNが小さくなるにつれて漸減する。ただし，FT-I 型分布における減少の度合は角屋による値よりも少な目である。また，Gumbel による計算値は今回の値や角屋による値よりもかなり小さい。分布関数の母数推定法として積率法を FT-I 型分布に適用する際に Gumbel による σ_x の計算値を使うと，再現確率統計量の推定結果に正の偏り量が生じることを **2.3** で述べたが，これは Gumbel による σ_x の値が実態よりも過小であることに起因すると思われる。ただし，Gumbel は σ_x の計算の詳細を説明していないので，なぜそのような差を生じたかは不明である。

部分極値資料の場合は，平均値 $\bar{\sigma}_x$ の挙動が分布関数の形状によって異なる。確率密度関数が裾を長く引く $k=0.75$ のワイブル分布では，上位のデータのみを取り出すことによって標準偏差 $\bar{\sigma}_x$ の値が増加する。$k=1$ のワイブル分布（すなわち指数分布）では，ν をどのように取っても $\bar{\sigma}_x$ の値はほとんど変化しない。これは簡単な確率計算によって確認することができる。これに対して $k>1$のワイブル分布，FT-I 型分布，および対数正規分布ではデータ採択率 ν を下げるにつれて $\bar{\sigma}_x$ の値が減少する。これは確率密度関数の裾の部分の伸びが小さいため，データの散らばり方が少なくなることに対応している。ν の低下による $\bar{\sigma}_x$ の減少は対数正規分布が最も著しく，極値の大きな値が出にくいことを示唆している。

一方，標準偏差の変異係数はデータ採択率 ν による変化の傾向が $\bar{\sigma}_x$ の場合と逆である。すなわち，$k=0.75$の

表-5　極値統計資料の標準偏差の平均値とその変異係数

(1) FT-I 型分布

N	平均値 $\bar{\sigma}_x/\sigma_0$			変異係数 C.V. $[\sigma_x]$		
	ν=1.0	0.50	0.25	ν=1.0	0.50	0.25
10	0.951	0.789	0.754	0.317	0.386	0.404
14	0.964	0.804	0.766	0.271	0.326	0.343
20	0.974	0.817	0.777	0.228	0.277	0.291
30	0.985	0.824	0.789	0.186	0.228	0.238
40	0.989	0.828	0.791	0.165	0.198	0.209
60	0.992	0.833	0.795	0.133	0.165	0.172
100	0.995	0.835	0.799	0.104	0.129	0.135

(2) 対数正規分布

N	平均値 $\bar{\sigma}_x/\sigma_0$			変異係数 C.V. $[\sigma_x]$		
	ν=1.0	0.50	0.25	ν=1.0	0.5	0.25
10	0.976	0.665	0.579	0.252	0.304	0.328
14	0.982	0.670	0.583	0.209	0.255	0.276
20	0.987	0.674	0.586	0.175	0.2 2	0.231
30	0.994	0.674	0.591	0.140	0.172	0.186
40	0.995	0.676	0.591	0.122	0.148	0.162
60	0.997	0.676	0.592	0.097	0.123	0.132
100	0.998	0.676	0.594	0.076	0.093	0.102

(3) ワイブル分布 (k=0.75)

N	平均値 $\bar{\sigma}_x/\sigma_0$			変異係数 C.V. $[\sigma_x]$		
	ν=1.0	0.50	0.25	ν=1.0	0.5	0.25
10	0.868	0.994	1.100	0.566	0.529	0.505
14	0.898	1.028	1.128	0.492	0.448	0.428
20	0.922	1.057	1.152	0.424	0.384	0.367
30	0.944	1.076	1.179	0.352	0.320	0.303
40	0.958	1.087	1.186	0.308	0.279	0.268
60	0.973	1.102	1.201	0.257	0.238	0.223
100	0.983	1.111	1.213	0.202	0.184	0.177

(4) ワイブル分布 (k=1.0)

N	平均値 $\bar{\sigma}_x/\sigma_0$			変異係数 C.V. $[\sigma_x]$		
	ν=1.0	0.50	0.25	ν=1.0	0.50	0.25
10	0.925	0.920	0.927	0.417	0.420	0.420
14	0.943	0.943	0.944	0.362	0.355	0.356
20	0.954	0.961	0.959	0.301	0.301	0.302
30	0.972	0.972	0.975	0.245	0.248	0.247
40	0.979	0.979	0.978	0.217	0.215	0.217
60	0.985	0.986	0.984	0.179	0.180	0.179
100	0.989	0.990	0.989	0.140	0.140	0.140

(5) ワイブル分布 (k=1.4)

N	平均値 $\bar{\sigma}_x/\sigma_0$			変異係数 C.V. $[\sigma_x]$		
	ν=1.0	0.50	0.25	ν=1.0	0.50	0.25
10	0.953	0.814	0.748	0.308	0.346	0.361
14	0.966	0.828	0.757	0.261	0.292	0.305
20	0.976	0.839	0.766	0.219	0.245	0.257
30	0.987	0.844	0.776	0.177	0.200	0.209
40	0.990	0.848	0.777	0.157	0.173	0.182
60	0.993	0.851	0.780	0.126	0.144	0.150
100	0.995	0.853	0.782	0.098	0.112	0.117

(6) ワイブル分布 (k=2.0)

N	平均値 $\bar{\sigma}_x/\sigma_0$			変異係数 C.V. $[\sigma_x]$		
	ν=1.0	0.50	0.25	ν=1.0	0.50	0.25
10	0.971	0.710	0.604	0.246	0.304	0.325
14	0.983	0.717	0.609	0.206	0.255	0.275
20	0.983	0.724	0.614	0.171	0.213	0.230
30	0.990	0.726	0.620	0.139	0.173	0.186
40	0.994	0.728	0.620	0.120	0.150	0.162
60	0.996	0.730	0.622	0.099	0.124	0.132
100	0.998	0.730	0.624	0.075	0.094	0.103

ワイブル分布では ν の低下につれて変異係数が減少し，指数分布ではほとんど変化しないのに対し，それ以外の分布関数では ν の低下につれて σ_x の変異係数が増加する。もっとも，変異係数の絶対値としては k=0.75 のワイブル分布が最大であり，N=10 の標本では C.V. $[\sigma_x]$ が0.57にも達する。変異係数の値がこのように大きい場合，σ_xの分布は正規分布から外れて強い正のひずみ度を示す。

分布関数の母数を積率法で推定する場合，尺度母数 A は標本の標準偏差 σ_x に比例する量として求められる（式7，12，16参照）。本報告のように最小2乗法によって母数推定を行っても事情は変らず，$\hat{A}$ は σ_x にほぼ比例する。したがって，σ_x の変動はそのまま $\hat{A}$ の変動となる。今回の数値シミュレーションのデータについて $\hat{A}/A$ の比の標準偏差を調べてみると，表-5 に記載した σ_x の変異係数よりも 0～6％小さいだけである。その意味で，σ_x の変異係数の大きさは母数推定値 $\hat{A}$ ひいては再現確率統計量 $\hat{x}_R$ の信頼区間の幅を規定する主要因といえる。

表-6 標本データの標準偏差の変異係数の比例定数 c
[C. V. $(\sigma_x)=cN^{-1/2}$]

分布関数	比例定数 c		
	$\nu=1.0$	0.50	0.50
FT-I 型分布	1.02	1.25	1.31
対数正規分布	0.77	0.95	1.03
ワイブル分布 ($k=0.75$)	1.92	1.75	1.66
ワイブル分布 ($k=1.0$)	1.36	1.36	1.36
ワイブル分布 ($k=1.4$)	0.98	1.10	1.15
ワイブル分布 ($k=2.0$)	0.76	0.95	1.03

表-5のC. V. $[\sigma_x]$ の数値を吟味すると，これらは標本の大きさNの $-1/2$ 乗にほぼ比例していることが分る。そこで，変異係数が $N^{-1/2}$ に比例するものと見なし，その比例定数の値を $N=10\sim100$ のデータに対する算術平均として求めた結果を表-6に示す。この結果は本質的には表-5の C. V. $[\sigma_x]$ のデータと同一であるが，標本の大きさNの効果を取り除いた形で表したことによって，分布関数ごとの特徴をより鮮明に表示している。すなわち，ワイブル分布では形状母数kが小さくなるにつれて σ_x の変異係数が増大する。$k=2.0$ のワイブル分布の変異係数は対数正規分布のものとほとんど同一であり，この2分布の特性が極めて類似していることをうかがわせる。また，FT-I 型分布は $k=1.4$のワイブル分布とかなり似た特性を示すことが分る。さらに，データ採択率 ν が低下するにつれ，分布関数による差異が減少してくる。これは，与えられた標本データから母関数の分布形状を推定しようとする場合，母関数の判別がむずかしくなることを示唆するものである。これについては **5.2** で詳述する。

4.2 分布関数の母数の推定値の信頼区間

与えられた標本データに対して特定の分布関数をあてはめて尺度母数Aおよび位置母数Bを推定する場合，得られた推定値 $\hat{A}$ および $\hat{B}$ が母集団における真値からどの程度離れている可能性があるかということは，統計学上でしばしば取り上げられる問題である。たとえば，製品の品質管理において母集団の分布関数として特定のものを設定しておいたとき，最初の標本資料から得られた母数の推定値が適切な信頼区間の中に入っていれば当初の分布関数の設定で良いと判断される。信頼区間から大幅に外れるのであれば，分布関数の設定をやり直す必要が生じる。

極値統計の問題では，母数推定よりも再現確率統計量の推定の方が関心事である。また，母数推定値の信頼区間については研究発表があまり多くない。著者の知り得た範囲では，Lawless[21] がFT-I 型分布および上方有界型のFT-III型分布について最尤法で尺度母数および位置母数を推定したときの信頼区間の近似式を導き，Challenor[22] がこの近似式に基づいて数表を作成しているのみである。しかし，今回の数値シミュレーションで作成された標本資料を解析すれば，母数推定値の信頼区間も容易に求めることができる。

FT-I 型分布，対数正規分布，ワイブル分布4種（$k=0.75$，1.0，1.4，および 2.0）の6種の分布について $N=10\sim100$の範囲で尺度母数の推定値 $\hat{A}$ および位置母数の推定値 $\hat{B}$ の信頼区間を解析した経果はまとめて**付表-B.1～B.6** に示してある。信頼区間としては非超過百分率 2.5%，5%，25%，75%，95%，および 97.5%の分位値を用い，参考として標準偏差の不偏推定量を示してある。信頼区間の無次元表示としては真値推定の便宜を考え，Petruaskas と Aagaard[15] や Lawless[21] などにならい，次の補助統計量（ancillary statistics）を用いている。

$$A/\hat{A}, \qquad (\hat{B}-B)/\hat{A}$$

データ採択率としては $\nu=1.0$，0.5，および0.25の3種類について計算した結果が提示されている。

付表-B.1～B.6 の結果について特徴的な点を以下にまとめて述べる。

(A) 尺度母数について

1) 補助統計量 $A/\hat{A}$ はすべて正のひずみ度を持つ分布形であり，平均値は 1.0 よりもやや大きい。これは推定値 $\hat{A}$ を分母に取っているためであり，$\hat{A}/A$ の形では平均値は 1.0 に極く近い値となる。

2) 分布関数ごとの傾向は表-5に示した標本データの標準偏差の傾向と同一であり，分布幅が最も広いのは $k=0.75$ のワイブル分布，最も狭いのは対数正規分布である。また $k=2.0$ のワイブル分布は対数正規分布と非常に良く似た分布を示す。

3) データ採択率が減少すると，$k=0.75$ のワイブル分布は推定値 $\hat{A}$ の分布幅が減少，$k=1.0$ のワイブル分布は無変化，それ以外の分布関数では推定値 $\hat{A}$ の分布幅が増大する。この傾向も **表-5** と同一である。

4) $A/\hat{A}$ の標準偏差は表-5 の C. V. $[\sigma_x]$ よりもやや大きく，特に標本の大きさNが小さいときに顕著である。これも推定値 $\hat{A}$ を分母に取っているためであり，$\hat{A}/A$ の形では表-5に関連して述べたように σ_x の変異係数よりもやや小さ目である。

(B) 位置母数について

1) 補助統計量 $(\hat{B}-B)/\hat{A}$ の分布形状も $A/\hat{A}$ と同様に正のひずみ度を持ち，平均値は0よりもやや正に偏った値をとる。ただし，対数正規分布で $\nu=1.0$ の場合は0を中心とした対称な分布形を示す。なお，補助統計量として $\hat{B}/B$ を用いた場合には平均値が1.0に極く近い値となる。

2) いずれの分布関数においても，データ採択率を下げると $(\hat{B}-B)/\hat{A}$ の分散が大幅に増大する。すなわち，位置母数の推定精度が低下する。これは，分布関数の裾の部分のみを使って全体像を推定しようとすることから必然的に発生する現象である。

3) 尺度母数に関しては対数正規分布と $k=2.0$ のワイブル分布がかなり似た性質を示すが，位置母数に関してはやや異なる傾向を示す。すなわち，データ採択率が $\nu=1.0$ の場合は $k=2.0$ のワイブル分布の方が位置母数の推定値の分散が小さいのに対し，データ採択率が $\nu=0.5$ 以下では対数正規分布の方が少ない分散を示している。これは，対数正規分布における分布の裾の引き方が少ないことに起因すると推測される。

付表-B.1～B.6 に示した母数推定値の信頼区間は，ここで用いた母数推定法，すなわち最適なプロッティング公式を用いた最小2乗法に基づく結果であって，母数推定法が異なれば信頼区間もまた変化する。最小2乗法を用いてもプロッティング公式を変えると異なる結果が得られる。たとえば，FT-I 型分布に対してバーネット公式でプロッティング・ポジションを計算すると，$\nu=1.0$ の場合には $A/\hat{A}$ の標準偏差が10%程度減少し，信頼区間もそれに応じて狭くなる。$(\hat{B}-B)/\hat{A}$ の標準偏差も1～3%減少する。また，最尤法を用いる場合の Lawless の近似解[21] に基づいて Challenor[22] が計算した FT-I 型分布の母数推定値の信頼区間の結果と比べると，最尤法で母数推定を行うと $\hat{A}/A$ の標準偏差がバーネット公式を用いた最小2乗法による値よりも15～20%程度小さくなる。ただし，$(\hat{B}-B)/\hat{A}$ についてはほとんど差がなく，むしろバーネット公式による推定値の方が僅かながら分散の度合が少ない。

一方，積率法で母数を推定した場合の信頼区間については研究報告を見かけないので比較ができない。ただし 4.1 で言及したように，今回の数値シミュレーションの結果では最小2乗法で求めた推定値 $\hat{A}/A$ の標準偏差が標本データの標準偏差 σ_x の変異係数よりも0～6%小さかった。すなわち，積率法のように σ_x に直接に比例する量として $\hat{A}$ を求める場合よりも若干ながら変動が小さく抑えられたことになる。

いずれにしても，母数推定法による推定値の信頼区間の差異は小さい。採択した推定法における推定値の信頼区間を把握してさえいれば，どの方法を用いても差し支えないものと思われる。

4.3 再現確率統計量の推定値の信頼区間

(1) 再現確率統計量の算定式

これまでに述べたように，1組の標本データに対して推定した分布関数の母数の値は，標本ごとにかなり大きく変動する。母数の絶対値が変れば再現期間 R 年に対する統計量の推定値 $\hat{x}_R$ も異ならざるを得ない。R 年確率統計量 x_R については先に式(26)，(27)としてその定義を与えたが，ここで具体的な算定式を示しておく。

1) FT-I 型分布

$$\left.\begin{aligned}\hat{x}_R&=\hat{A}\,y_R+\hat{B}\\ y_R&=-\ln[-\ln(1-1/\lambda R)]\end{aligned}\right\}\tag{74}$$

2) 対数正規分布

$$\left.\begin{aligned}\hat{x}_R&=\exp[\hat{A}\,y_R+B]\\ y_R&=\Phi^{-1}(1/\lambda R)\end{aligned}\right\}\tag{75}$$

ここに，$\Phi(t)$ は次式で定義される標準正規分布の累積分布関数であり，Φ^{-1} はその逆関数である。

$$\Phi(t)=\frac{1}{\sqrt{2\pi}}\int_t^{\infty}\exp[-u^2/2]\,du\tag{76}$$

3) ワイブル分布

$$\left.\begin{aligned}\hat{x}_R&=\hat{A}\,y_R+\hat{B}\\ y_R&=[-\ln(1/\lambda R)]^{1/k}\end{aligned}\right\}\tag{77}$$

なお，期間最大値資料の場合は式(74)，(75)，(77)において $\lambda=1$ と置けばよい。また，母集団における R 年確率統計量 x_R は，母集団の尺度母数 A および位置母数 B を使い，

$$x_R=A\,y_R+B\tag{78}$$

によって求める。基準化変量は式(74)～(77)と同じである。

(2) 全数極値資料における $\hat{x}_R$ の信頼区間

FT-I 型分布に関しては母数推定を最尤法で行ったときの理論値が Lawless[2),21)] によって与えられているが，具体的には数値積分が必要であるため一般的でない。また，最小2乗法で母数推定を行った結果に対しては直接には適用できない。一方，Gumbel[1)] (6.2.3節) は百分率等の分位値に基づく信頼区間は明示していないものの，再現確率統計量の分散を次のように与えている。

$$\frac{N\sigma^2(\hat{x}_R)}{\sigma^2}=1+1.1396(y_R-\gamma)\frac{\sqrt{6}}{\pi}+1.1(y_R-\gamma)^2\frac{6}{\pi^2}\tag{79}$$

ここに，σは母集団の標準偏差，Nは標本の大きさ，γはオイラーの定数（0.5772…）である。なお，この算定式は Resio[34] や Simiu と Filliben[35] が引用しており，また Lettenmaier と Burges[18] は数値実験結果に比べるとやや大き目の値を与えるとしている。

極値資料の標本データを解析する立場からは母集団の標準偏差σは未知であるので，標本データの値 σ_x をその代りに用いることにして式(79)を書き直したのが次式である。

$$\sigma(\hat{x}_R)=\frac{\sigma_x}{\sqrt{N}}[1+0.8885(y_R-\gamma)+0.6687(y_R-\gamma)^2]^{1/2} \qquad (80)$$

これまでに述べてきた数値実験のうち，データ採択率が $\nu=1.0$ のものについては平均発生率λを指定せず，結果的には $\lambda=1$ である期間最大値資料と同形式で取り扱った。そして，再現期間Rを標本中のデータ個数Nの0.2倍から100倍まで9通りに変えてR年確率統計量 $\hat{x}_R$ を推定した。推定結果を無次元化する補助統計量としては，先に図-15で例示した次の形のものを使用した。

$$z=(\hat{x}_R-x_R)/\sigma_x \qquad (81)$$

無次元化の基準量としては一般に尺度母数推定値$\hat{A}$が用いられるが，既に述べたように$\hat{A}$はσ_xにほぼ比例する量であり，標本データが与えられた場合には$\hat{A}$よりもσ_xの方が直接に計算される量であることによる。さらに，式(81)の補助統計量であればその標準偏差は式(80)のような表示式との比較が容易である。

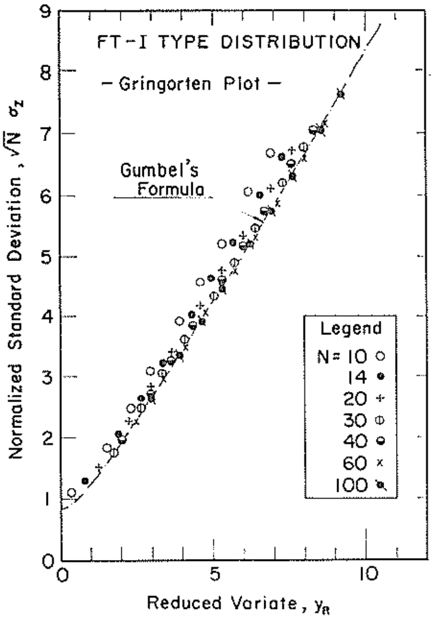


図-16　FT-I 型分布における再現確率統計量の標準偏差

数値実験結果のうち，まず FT-I 型分布についてまとめたのが図-16である。補助統計量zの標準偏差σ_zは厳

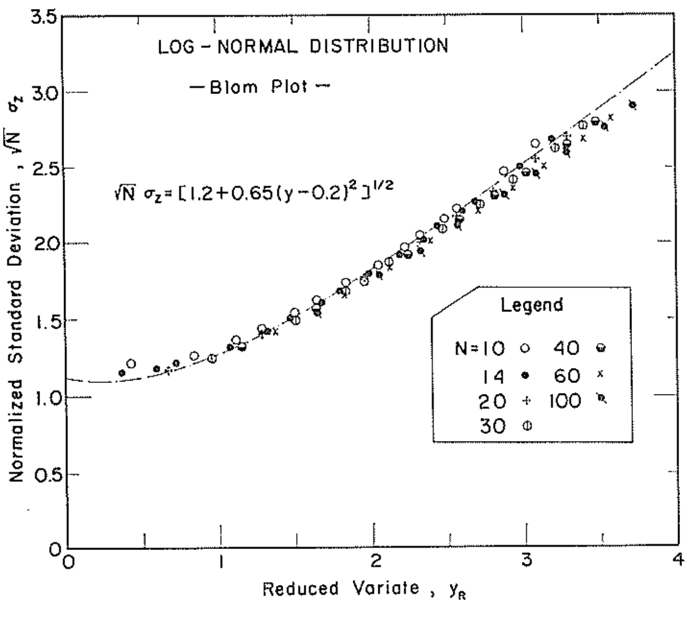


図-17　対数正規分布における再現確率統計量の標準偏差

密には $\sigma(\hat{x}_R)/\sigma_x$ に等しくないが，便宜的に等しいものとして式(80)による計算値を曲線で図示してある。図から明らかなように，式(80)は$N\geq30$に対して良い近似を与える。しかしながら，標本の大きさの影響も認められるので，Nを取り込んだ実験式を試行錯誤によって次のように作成した。

$$\sqrt{N}\sigma_z=[1.0+ay_R{}^2]^{1/2} \qquad (82)$$

ここに，

$$a=0.64\exp[9.0N^{-1.3}] \qquad (83)$$

なお，プロッティング公式としてバーネット公式を使うと $\hat{x}_R$ の標準偏差は1～5％程度減少する。

次に，対数正規分布については理論式を見つけることができなかったので，数値実験の結果のみを示したのが**図-17**である。図中の曲線はデータにあてはめた実験式であり，次のように与えたものである。

$$\sqrt{N}\sigma_z=[1.2+0.65(y_R-0.2)^2]^{1/2} \qquad (84)$$

対数正規分布の場合は標本の大きさの影響が少ない。しかし，正規分布の場合は FT-I 型分布と同程度に標本の大きさの影響を受ける。正規分布の標準偏差の実験式としては次のようなものを得ている。

$$\sqrt{N}\sigma_z=[1.2+0.50\exp(11.0N^{-1.3})y_R{}^2]^{1/2} \qquad (85)$$

ワイブル分布の場合には標本の大きさNの影響がFT-I 型分布よりも強く現われる。**図-18～20** は $k=0.75$, 1.0, および2.0の場合について$N=10$, 20, および100の結果のみを示したものである。こうしたデータに直線を仮にあてはめると，$k\geq1.0$の場合には回帰直線が横軸と交差する。このため，実験式としては次のように y_R の座標軸を若干ずらす方式を導入した。

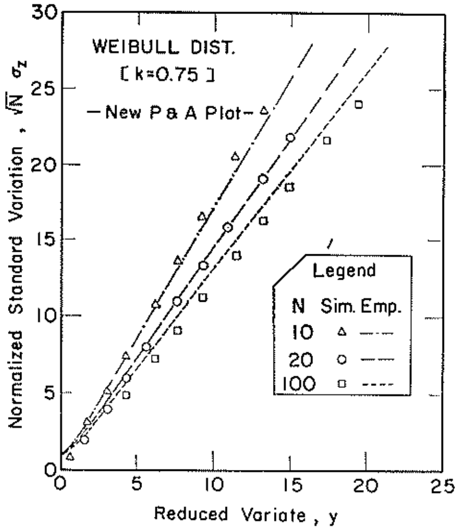


図-18 ワイブル分布（$k=0.75$）における再現確率統計量の標準偏差

$$\sqrt{N}\sigma_z=[1.0+a(y_R-c)^2]^{1/2} \qquad (86)$$

ここに，

$$a=a_1\exp[11.4N^{-1.3}] \qquad (87)$$

係数 a_1 および c はワイブル分布の形状母数ごとに次のように与えた。

$$\left.\begin{array}{ll} k=0.75: & a_1=1.65,\ c=0 \\ k=1.0\ : & a_1=1.92,\ c=0.3 \\ k=1.4\ : & a_1=2.05,\ c=0.4 \\ k=2.0\ : & a_1=2.22,\ c=0.5 \end{array}\right\} \qquad (88)$$

図-18～20 に記入した曲線は式(86)～(88) を用いた実験

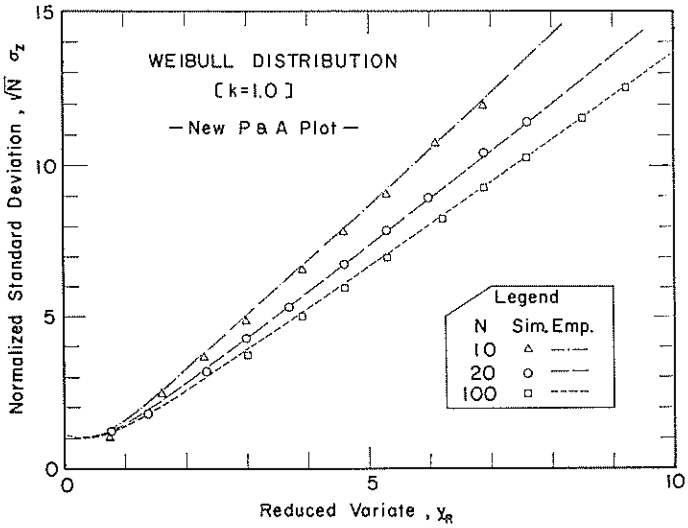


図-19 ワイブル分布（$k=1.0$）における再現確率統計量の標準偏差

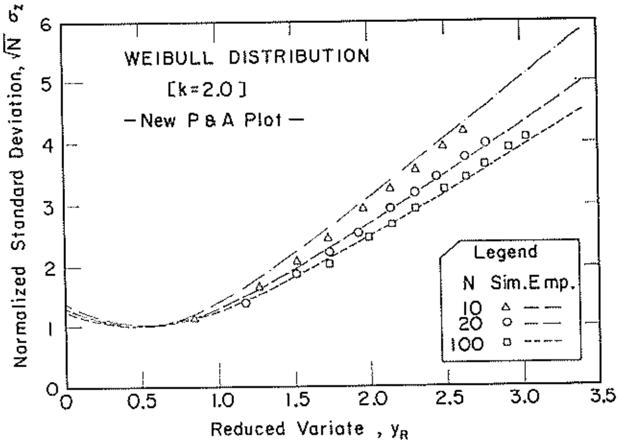


図-20 ワイブル分布（k=2.0）における再現確率統計量の標準偏差

式の値である。y_R の値の小さいところでやや過大な値を与える傾向が見られるものの，数値シミュレーションの結果にほぼ合致しているといえる。なお，式(86)によると $y_R<c$ の領域で標準偏差が増大することになる。これは FT-I 型分布に対する式(80)の理論式を参考にして実験式を設定したことに基づくもので，データによる検証は行っていない。

再現確率統計量の推定値の信頼区間は，本来は **付表-B.1～B.6** に示したような百分率の分位値として表す必要がある。推定値の分散状況がほぼ正規分布で近似できるときは標準偏差のみで表示できるけれども，補助統計量 z の分布は必らずしも正規分布で表示できない。そこで，z の累積度数分布で非超過確率が2.5%，5%，25%，75%，95%，および97.5%に相当する分位値 z_p を求め，zの標準偏差 σ_z で除して無次元化した結果を**図-21～23**に示す。それぞれ，FT-I 型分布（グリンゴルテン公式使用），対数正規分布（ブロム公式使用），および k=1.0 のワイブル分布（新 P&A 公式使用）に対するものであ

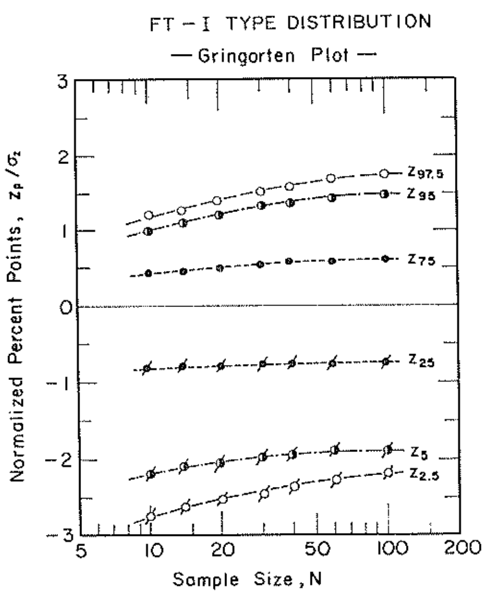


図-21 FT-I 型分布における再現確率統計量の信頼区間の算定図

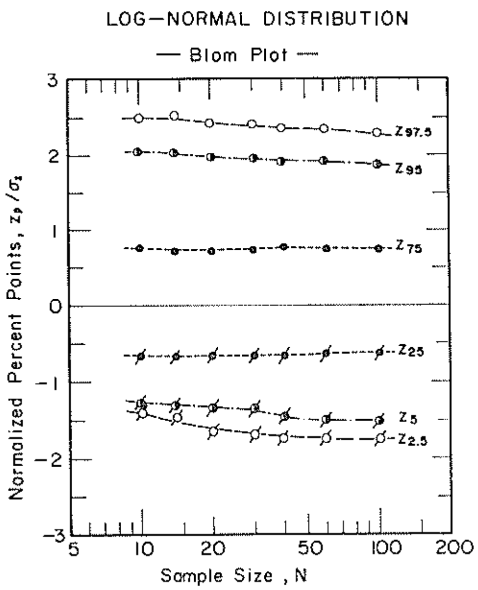


図-22 対数正規分布における再現確率統計量の信頼区間の算定図

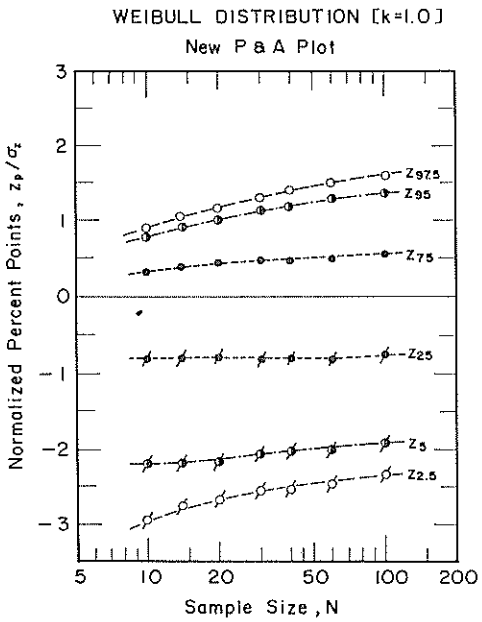


図-23　ワイブル分布（k=1.0）における再現確率統計量の信頼区間の算定図

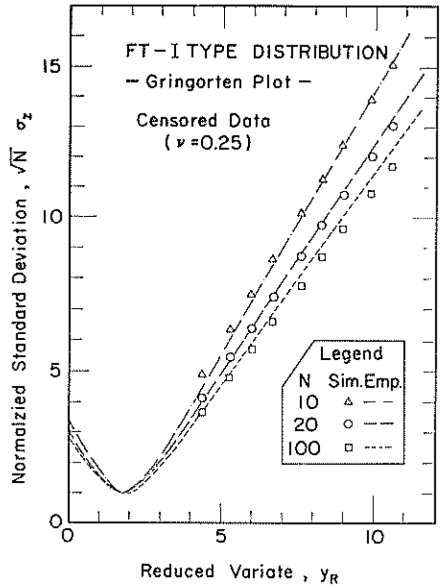


図-24　FT-I 型分布の部分極値資料における再現確率統計量の標準偏差

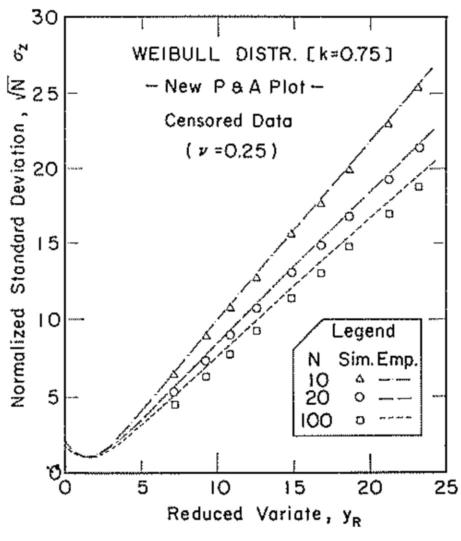


図-25　ワイブル分布（k=0.75）の部分極値資料における再現確率統計量の標準偏差

る。再現期間としては$R=10N$に対するものを示している。$R=10N$ 以外の結果も z_p/σ_z の形で表示すると再現期間による差があまり目立たなくなる。なお，補助統計量 z の分布が正規分布と見なされる場合には，図中に示した分位値が次のような値をとる。

$$z_{97.5}=1.96\sigma_z \qquad z_{2.5}=-1.96\sigma_z$$
$$z_{95}=1.64\sigma_z, \qquad z_5=-1.64\sigma_z$$
$$z_{75}=0.67\sigma_z, \qquad z_{25}=-0.67\sigma_z$$

標本が大きくなると，補助統計量 z の分布は正規分布に近付く。図で示した中では対数正規分布の場合が最も収束が速く，ワイブル分布は最も遅い。

(3)　部分極値資料における $\hat{x}_R$ の信頼区間

極値統計資料のうちで絶対値の小さいものを切り捨てた部分極値資料を使って再現確率統計量を推定すると，その推定結果は全数資料の場合よりも大きな分散を示すことになる。この問題に関する理論的検討は行われていないと思われるので，ここでは数値実験による結果に基づいて述べる。

部分極値資料に関する数値シミュレーションでは全数極値資料の場合とは異なり，平均発生率 λ をあらかじめ設定しておく必要がある。本報告ではデータ採択率として ν=0.5 および 0.25 の場合について検討し，それぞれ λ=20 および40とした。したがって，1年間当りのデータ採択個数はいずれも10個である。標本の大きさはN=10～100で変化させたので，形式的にはK=1～10年の期間のデータを対象としたことになる。再現確率統計量はR=2～1000 年の再現期間に対して推定作業を行った。もっとも，ここで用いた λ の設定値や再現期間Rの絶対値は数値実験の実施上の必要性から仮に用いたものであ

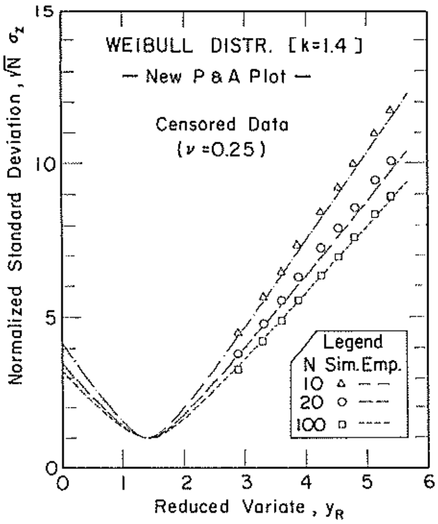


図-26 ワイブル分布（$k=1.4$）の部分極値資料における再現確率統計量の標準偏差

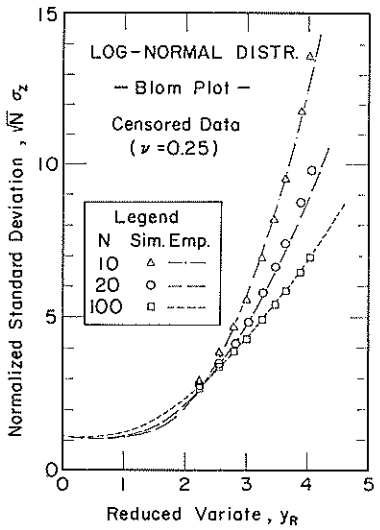


図-27 対数正規分布の部分極値資料における再現確率統計量の標準偏差

り，補助統計量 z と基準化変量 y_R の関係に直してしまえば λ や R のどのような値に対しても一般性を保つものである。

図-24～27 は全数極値資料の場合と同様に，補助統計量 z の標準偏差 σ_z と基準化変量 y_R の関係を図示したもので，それぞれ FT-I 型分布，ワイブル分布の $k=0.75$ と $k=1.4$，および対数正規分布の場合であり，データ採択率はいずれも $\nu=0.25$ である。対数正規分布の場合は下に凸な曲線状をなすけれども，FT-I 型分布とワイブル分布はほぼ直線関係で表される。また，FT-I 型分布とワイブル分布について全数極値資料の場合の $\sigma_z \sim y_R$ の関係と比べてみると，部分極値資料の場合はデータが全体として右側へ移動している。ただし，$\sigma_z \sim y_R$ の関係を表す線の勾配は変化している。

データが y_R 軸の右方へ移っているのは，ここで使ったデータが標本の大きさ N を固定したままデータ採択率 ν を変化させているため，平均発生率 λ も同時に変化していることによる。すなわち，式(74)～(77)の y_R の算定式の右辺に λ が含まれており，λ が大きくなると y_R の値も大きくなるためである。実際の極値資料のように λ が既に定まっている場合のデータ採択率 ν は，全極値データの中から上位どのくらいの範囲を採択するかを表すパラメータである。ν が小さくなるにつれて標本データの平均値 $\bar{x}$ は増大し，その標準偏差 σ_x は分布関数によって増大あるいは減少を示す。この ν の変化に伴う $\bar{x}$ の変化を $k=1.0$ のワイブル分布の場合について求めてみると

$$\bar{x}_\nu = B + A(1 - \ln \nu) \tag{89}$$

となる。これは y_R としては $-\ln \nu$ の変化に相当する。一方，式(77)で $k=1.0$ の場合について $\lambda \propto 1/\nu$ と考えると，y_R の変化量は同じく $-\ln \nu$ で表される。さらに，FT-I 型分布の再現確率統計量の分散を与える式(80)においても，y_R に対する補正量的に導入されているオイラー定数 γ は，偶然の一致かも知れないが式(6)で示したように平均値 $E[x]$ の係数に等しい。

このようなところから，部分極値資料の $\hat{x}_R$ の標準偏差については y_R の値を $\ln \nu$ のある倍数で補正する方法によって実験式を組み立てることが可能と考えられる。本報告で採用した実験式は次のようなものである。

$$\sqrt{N}\sigma_z = [1.0 + a(y_R - c + \alpha \ln \nu)^2]^{1/2} \tag{90}$$

これは式(86)の実験式に $+\alpha \ln \nu$ の補正を行ったものである。また係数 a については，$\sigma_z \sim y_R$ の勾配が分布関数によって異なる点を考慮し，式(87)を次のように修正した。

$$a = a_1 \exp[a_2 N^{-1.3} + \kappa(-\ln \nu)^{1/2}] \tag{91}$$

係数 a_1，a_2，c および κ は，式(83)および(88)式も含めて**表-7** のように与えた。**図-24～26** の中の曲線はこうした実験式による値であり，数値シミュレーションのデータをほぼ的確に表現している。**表-7** において係数 κ が $k=0.75$ のワイブル分布の場合に負となっているのは，この分布関数においては ν の減少につれて σ_x が減少することに対応している。また $k=1.0$ のワイブル分布において $\kappa=0$ であることも **4.1** で述べたことと照応してい

表-7　再現確率統計量の標準偏差の実験式の係数
—分布関数既知の場合—

$$\sigma[\hat{x}_R]=\sigma_x[1+a(y_R-c+\alpha\ln\nu)^2]^{1/2}/N^{1/2}$$
$$a=a_1\exp[a_2N^{-1.3}+\kappa(-\ln\nu)^{1/2}]$$

分布関数	a_1	a_2	κ	c	α
ワイブル (k=0.75)	1.65	11.4	−0.63	0	1.15
同上 (k=1.0)	1.92	11.4	0	0.3	0.90
同上 (k=1.4)	2.05	11.4	0.69	0.4	0.72
同上 (k=2.0)	2.24	11.4	1.34	0.5	0.54
FT-I 型	0.64	9.0	0.93	0	1.33

る。なお，図-18〜20の場合と同様に，y_R の小さい範囲で σ_z が上昇する部分は実験式の構造に基づくものであり，データによる検証は未了である。

一方，対数正規分布については図-17と図-27を比べると明らかなように，ν が小さくなると σ_z が急増し，しかも y_R の増加に伴う σ_z の変化が著しく大きくなる。このため $\sigma_z\sim y_R$ の関係曲線の曲率自体が変化し，式(90)，(91) の実験式を適用することができない。ここでは，今回得られた $\nu=0.5$ および0.25のデータのみを対象とし，次のような実験式を作成した。

$$\sqrt{N}\sigma_z=[1.2+a(y_R-c)^q]^{1/2} \qquad (92)$$

ここに，$c=0.2$であり，a および q は以下による。

$$a=\begin{cases}0.65 & :\nu=1.0\\ 1.55\exp[-4.6\,N^{-0.6}] & :\nu=0.5\\ 1.18\exp[-8.8\,N^{-0.6}] & :\nu=0.25\end{cases} \qquad (93)$$

$$q=\begin{cases}2.0 & :\nu=1.0\\ 2.0\exp[1.96\,N^{-0.5}] & :\nu=0.5\\ 2.5\exp[2.34\,N^{-0.5}] & :\nu=0.25\end{cases} \qquad (94)$$

図-27の中の曲線はこれらの実験によるものである。図示はしていないが，$\nu=0.5$のデータについてもほぼ同程度の適合度を示している。対数正規分布における再現確率統計量の推定値の標準偏差が全数極値資料と部分極値資料とでこのように異なる特性を示す理由は不明である。しかし，主として部分極値資料を対象として解析する場合には，対数正規分布はやや使いにくい分布関数であるといえよう。

5. 分布関数の選択および再現確率統計量の推定

5.1 分布関数の選択基準

極値統計の問題では，対象とする現象がある特定の分布関数を有する母集団に属すると先験的に決めてしまう場合が少なくない。また，気象や水文量に関しては，多数の地点における標本資料にいろいろな分布関数をあてはめた経験に基づいて，特定の分布関数を推奨する場合も多い。米国の風速資料について検討したSimiuほか[36]の研究では，18地点の毎年最大値資料（37年間）に対してFT-I型分布およびFT-II型分布（形状母数を含み，フレッシェ分布ともいう）をあてはめて最適分布関数の出現頻度を数えた。その一方で，FT-I 型分布の母集団からモンテカルロ法でN=37の標本を300組抽出し，各標本に対する最適分布関数の出現頻度を調べた。そして，両方の出現頻度状況がほぼ一致するところから，毎年最大風速の分布はFT-I型とみなしてよいと結論している。ただし，ハリケーンの直接の影響を受ける地点については分布関数が異なる可能性があると述べている。

最近，中西[4]は各地の毎年最大雨量の分布関数を対象とし，各標本をそれぞれ2等分してその片方で分布関数の母数推定を行い，得られた推定値を用いて残りの半分の資料の中の最大値の再現期間を推定することを試みている。そして，こうして得られた各地の標本最大値の再現期間の分布を統計的に検定することによって，当初に仮定した分布関数の妥当性を確めようとしている。現在のところ，各地点ごとの最大雨量分布が異なるためか，全地点に適用可能な分布関数は見出されていないようである。

極値分布の汎関数を見出す目的ではなく，単純に一つの標本資料が得られたときにどのような分布関数をあてはめるべきかはむずかしい問題である。先験的に特定の分布関数を優先させることなしに種々の分布関数を適用し，最も適合すると判断されるものを採用するのが一番合理的であると考えられる。しかし，その場合であっても各分布関数に対する標本資料の適合性をどのように判定するかについての定説はない。今までに提案されている判定基準の幾つかを列挙すると以下のとおりである。

1) 図式推定法（最小2乗法によるあてはめを含む）においては目視で各分布関数に対するデータの適合度を判断する。
2) 上記の定量的評価の一つとして，順序統計量 $x_{(i)}$ と基準化変量 $y_{(i)}$の相関係数の大きさで判定する。
3) 上記の別法として，高棹ほか[37]による Standard Least-Square Criterion (SLSC) を使う。
4) 類似の方法として，Petruaskas と Aagaard[15]による Mean Square Deviation (MSD) を使う。
5) 最尤法で用いられる式(21)の尤度関数の値あるいはその対数値が大きいものほど適合が良いとする。
6) ジャックナイフ法を用いて再現確率統計量の変動を調べ，変動量の小さいものほど適合が良いとする。

以上のうち，第1は主観的判断であり，普遍化がむずかしいのでここでは除外する。第2の $x_{(i)}$ と $y_{(i)}$ の相関係数の大小で判断する方法は我が国で波浪解析に用いられているものである。これは，Simiu ほか[35),36)]が Maximum Probability Plot Coefficient Criterion と呼んでいるものと実質的に同じである。ただし，Simiu ほかの母数推定法は最小2乗法とやや異なり，また基準化変量を求めるときのプロッティング・ポジションとして期待値 $E[x_{(i)}]$ ではなく，中央値 $x_{(i)\mathrm{median}}$ を使う点で若干の差異はある。

高棹ほか[37)]の SLSC は次のように定義されている。

$$\mathrm{SLSC}=\frac{\left\{\frac{1}{N}\sum_{i=1}^{N}\left[x_{(i)}-(\hat{A}y_{(i)}+\hat{B})\right]^2\right\}^{1/2}}{|F^{-1}(0.99)-F^{-1}(0.01)|} \quad (95)$$

すなわち，最小2乗法で順序統計量 $x_{(i)}$ とその基準化変量 $y_{(i)}$ との間に直線をあてはめたときの残差の平均値を，非超過確率99%に対応する再現確率統計量と同じく1%に対する値との差で除して無次化したものである。高棹ほかは SLSC が0.02以下であれば分布関数への適合が良好と判定できるとしている。Petruaskas と Aagaard[15)] の MSD は式(95)の分母を尺度母数の推定値 $\hat{A}$ で置き換え，その結果を2乗したものである。MSD の絶対値は分布関数ごとに異なるので，あらかじめモンテカルロ法で MSD 値の分布を求めておき，その90%非超過値と比べて対象とする分布関数の採否を定める方法を用いている。この MSD に比べると SLSC の方が汎用性に富むと思われる。しかし，SLSC にしても $x_{(i)}$ と $y_{(i)}$ との関係が直線式からずれる程度を判定しようとするものであり，その目的では既に第2の相関係数が利用されており，SLSC が $x_{(i)}$ と $y_{(i)}$ の相関係数よりもさらに適切な指標であるとは考えにくい。

第5の最大尤度あるいは対数最大尤度の値で判断する方法[26)]は，分布関数の母数を最尤法で推定する場合には適切な方法ではないかと思われる。ただし，本報告のように最小2乗法を使って母数を推定した場合には，得られた分布関数が必らずしも最大尤度の条件を満していないので適用困難である。

第6のジャックナイフ法を使う方法は宝ほか[38)]が1987年に提案したもので，N個のデータからなる標本に対してデータを順に1個ずつ抜いた2次標本をN組作成し，各々に対して特定の分布関数のあてはめを行って適当な再現期間に対する再現確率統計量を計算する。これによってN個の推定値が得られるのでその変動範囲を計算する。この作業を想定される各種の分布関数について実行して変動範囲の最小のものを最適合分布関数とする。標本中のデータ個数Nの増加につれて計算時間が指数関数的に増大するのが難点であるが，新しい試みである。

以上のように，一つの標本資料に対して適合する分布関数を採択する方法もいろいろある。Petruaskas と Aagaard[15)]は最適合関数は定めず，90%信頼限界内の候補を複数個残し，最後は技術者の選択に委ねる方式を示している。しかし，それでは実務上いろいろ不具合である。本報告では，従来から実質的に用いられている相関係数が最大のものを最適合分布関数とする方法が最も分りやすく，かつ計算も単純であるところからこの方式を踏襲することにする。

5.2 母分布関数への適合度

(1) 9分布関数によるあてはめ

Petruaskas と Aagaard[15)]の研究では，分布関数として次の8種類のもののあてはめを試みることを提案している。

FT-I 型分布（Gumbel 分布）

Weibull 分布：k=0.75，0.85，1.0，1.1，1.25，1.5，および2.0

これらの分布関数を選んだ理由としては，FT-I 型分布が種々の極値統計に用いられていることと，分布関数の形状に幅を持たせることが挙げられているのみであり，形状母数 k の値にしても絶対的な根拠があったわけではないと考えられる。

本報告では当初，上記の8分布に加えて対数正規分布も入れた9種類を対象とし，一つの母関数から抽出された標本に対して最も適合する分布関数を求め，それが母関数にどの程度復帰するかを調べてみた。図-28はその結果の1例である。ここではデータ採択率が $\nu=1.0$ の場合を対象とし，母分布関数が FT-I 型分布，ワイブル分布($k=2.0$)，およびワイブル分布($k=1.0$)の場合について標本の大きさが $N=10$ および100の結果を示している。図の縦軸は最適合分布関数の出現率，横軸は各分布関数の指示記号であり，L-N は対数正規分布の略記号である。なお，シミュレーション回数はいずれも10,000回，母数はいずれも $A=1.0$，$B=5.0$である。

まず，母関数が FT-I 型分布の場合，$N=10$のように小さな標本では母関数が最適合分布として選定されるのは10%に満たず，対数正規分布を最適合と判定するケースが33%も生じる。また分布形状が大きく異なる$k=0.75$のワイブル分布を最適合とするケースが9%も存在する。標本が $N=100$ と大きなものになると FT-I 型の母関数に復帰する率が増えるけれども，その場合でも34%に過ぎない。これに対して，対数正規分布あるいは $k=2.0$ のワイブル分布のいずれかが適合度が良いと判定さ

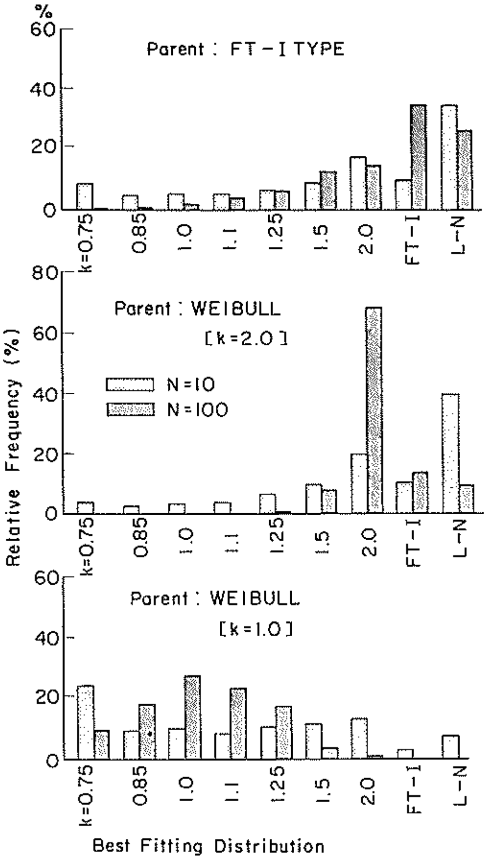


図-28　極値統計の標本資料に対する各種分布関数の適合状況

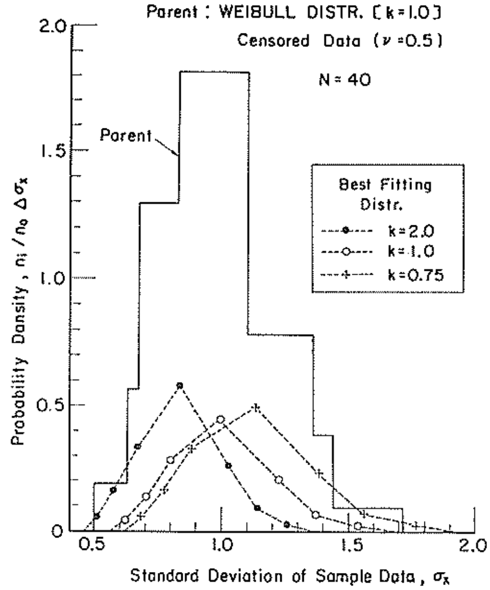


図-29　標本資料の標準偏差の分布状況

れるケースが合せて40%を超える。

$k=2.0$ のワイブル分布の場合には，標本が $N=100$ のように大きくなると他分布と見誤られる率が30%程度と低くなる。しかし，$N=10$ のような小標本では対数正規分布に取り違えられるケースが40%も発生する。

$k=1.0$ のワイブル分布の場合は母関数への的中率がさらに低く，$N=100$ の大標本でも $k=0.75$～1.25 のワイブル分布に分散してしまう。

このように，母分布関数からランダムに抽出された標本データに対して母分布よりも他分布の方が良く適合すると判定される原因は，標本データの変動性そのものにあると考えられる。図-29 は次項で述べる5分布関数あてはめ方式のデータであって図-28の9分布関数あてはめ方式の場合とやや異なるが，基本的性質は同じである。この図は，データ個数が $N=40$ の標本の標準偏差 σ_x の分布を調べたものであり，母関数は $k=1.0$ のワイブル分布，データ採択率は $\nu=0.5$（すなわち $N_T=80$ のデータのうち上位40個を採択）のケースである。階段状の実線は10,000組の標本の σ_x の分布を示し，σ_x が 0.5 から1.7付近にまで広がっていることを表している*。このデータに対して5種類の分布関数をあてはめると，$k=0.75$ のワイブル分布が28.5%，$k=1.0$ が21.7%，$k=1.4$ が17.3%，$k=2.0$ が22.9%，FT-I 型分布が 9.5%となる。これらの分布関数が最も適合すると判定された標本のそれぞれの組について σ_x の分布を調べた結果は3本の点線で示されている。$k=1.4$ のワイブル分布と FT-I 型分布の結果は省いてあるが，5分布関数の縦軸の値を足し合せると全体に対する σ_x の分布（階段状の実線）に一致するはずである。

図-29 に見られるように，$k=2.0$ のワイブル分布が最適合と判定された標本は σ_x が小さいものが多く，逆に $k=0.75$ のワイブル分布と判定されたものは σ_x が大きいグループに属する。σ_x が小さいということは標本自体において最大値と最小値の差が小さく，分布幅が狭いことを意味するので，分布関数として分布幅の狭い $k=2.0$ のワイブル分布が選択される率が高まるのは当然といえる。標本の大きさ N が増すにつれて母分布関数に復帰する割合が高まるのは，N の増加につれて $N^{1/2}$ に逆比例する形で σ_x の変動幅が減少し，図-29 の階段状の母集

* この分布は百分率分位値を使って推定したもので，下限値は 2.5%分位値から下方へ5%分位値との差の3倍だけ下った位置にあると仮定している。上限値についても同様である。

表-8 各母関数に対する最適合分布の出現率（0.1%単位）
—データ採択率 $\nu=1.0$ の場合—

あてはめ関数		母関数 / N	ワイブル分布 $k=0.75$	ワイブル分布 $k=1.0$	ワイブル分布 $k=1.4$	ワイブル分布 $k=2.0$	FT-I 型分布	計
ワイブル分布	$k=0.75$	10	**475**	287	140	51	113	1066
		20	**520**	267	68	9	65	929
		40	**588**	225	29	1	30	872
		100	**687**	173	3	—	5	868
		200	**792**	130	—	—	—	921
	$k=1.0$	10	227	**235**	174	97	128	861
		20	304	**323**	191	59	129	1005
		40	340	**443**	186	12	117	1099
		100	305	**615**	124	—	60	1105
		200	208	**751**	74	—	24	1056
	$k=1.4$	10	172	217	**198**	138	143	867
		20	146	279	**296**	136	163	1019
		40	70	293	**419**	94	162	1038
		100	8	210	**642**	41	133	1033
		200	1	120	**800**	8	86	1015
	$k=2.0$	10	112	223	400	**582**	484	1800
		20	27	112	346	**631**	438	1554
		40	2	34	267	**708**	380	1392
		100	—	2	151	**795**	272	1220
		200	—	—	71	**898**	196	1165
FT-I 型分布		10	14	38	89	132	**134**	406
		20	2	19	99	166	**206**	493
		40	—	5	99	185	**310**	599
		100	—	—	80	164	**530**	773
		200	—	—	55	94	**694**	842

団の σ_x の分布幅が狭まるためであり，このデータは表-5, 6 に示したところである。

極値統計資料として我々が入手するのは，あくまでも自然界に存在する母分布関数からランダムに抽出された標本に過ぎない。時と所が異なれば，標本データも大きく変らざるを得ない。既に図-3で例示したように，$N=10$の小標本ではデータ中の最大値が8倍以上変化することがあり得る。このようなランダム統計量としての標本データの変動性を考えるならば，標本データに対する分布関数のあてはめ手法を高度に追究することは徒労であるとさえいえよう。

(2) あてはめ分布関数の選定

前項で紹介したように，$N=10$～100 程度の大きさの標本にあっては分布関数の形状の微少な差異を判別することが不可能であり，あてはめの対象とする分布関数の種類を減らす必要がある。

まず，対数正規分布は $k=2.0$ のワイブル分布と特性が類似している。このことは表-5, 6 の標本資料の標準偏差の特性および付表-B.2, B.6 の母数推定値の信頼区間で確かめられる。また，図-28 の例のように対数正規分布は $k=2.0$ のワイブル分布や FT-I 型分布の標本データに対する適合度が高いため，本来の分布関数を見誤る危険度が大きい。$k=2.0$のワイブル分布と対数正規分布の間ではどちらを適合関数として使っても再現確率統計量の推定値の差は小さいけれども，本来は FT-I 型分布に属する標本データに対数正規分布関数をあてはめると，推定値が平均的に低く出る。さらに，対数正規分布を用いると基準化変数 y_R の算定に正規分布関数の逆

表-9 各母関数に対する最適合分布の出現率（0.1%単位）

—データ採択率 $\nu=0.5$ の場合—

あてはめ関数 ＼ 母関数		N	ワイブル分布				FT-I 型分布	計
			$k=0.75$	$k=1.0$	$k=1.4$	$k=2.0$		
ワイブル分布	$k=0.75$	10	**437**	329	229	166	268	1428
		20	**475**	302	175	90	235	1277
		40	**515**	285	118	41	200	1159
		100	**588**	253	51	3	130	1024
		200	**675**	184	15	—	87	961
	$k=1.0$	10	124	**121**	105	84	102	536
		20	174	**164**	129	85	139	690
		40	232	**217**	140	73	170	832
		100	289	**312**	132	25	223	980
		200	284	**432**	91	5	232	1043
	$k=1.4$	10	93	92	**87**	75	82	430
		20	118	139	**131**	109	124	620
		40	123	173	**182**	140	171	789
		100	75	217	**282**	188	244	1006
		200	26	181	**394**	201	295	1097
	$k=2.0$	10	309	410	522	**608**	492	2340
		20	188	328	478	**631**	421	2045
		40	85	229	437	**642**	341	1734
		100	12	95	354	**689**	216	1367
		200	—	33	287	**728**	114	1161
FT-I 型分布		10	38	48	57	67	**57**	267
		20	45	68	88	85	**82**	368
		40	46	95	123	105	**118**	486
		100	35	124	182	95	**187**	623
		200	15	171	215	67	**273**	739

関数の計算が必要になり，机上計算には不向きである。

以上の諸点を考慮し，本報告では対数正規分布の役割は $k=2.0$ のワイブル分布で代替可能であると判断し，これを標本資料に対するあてはめ分布の候補から除外することとした。

次に，ワイブル分布については図-28の $k=1.0$ のワイブル分布の例に見られるように，7種類の形状母数を導入しても相互間の差異を判別することがむずかしいところから，$k=0.75$，1.0，1.4，および2.0の4種類に限定することにした。なお，形状母数を0.75よりも小さくすると分布関数はその裾をさらに長く引くようになり，σ_x の変動性が一層増大し，再現確率統計量の推定値の信頼度が低下する。一方，形状母数を2.0よりも大きくすると分布幅が一層狭くなり，再現確率統計量を過小に推定するおそれがある。$k=0.75$や$k=2.0$の数値に特別の意味があるわけではないが，$k=0.75$ は確率密度関数が左上りの代表，$k=2.0$はワイブル分布系の中で対数正規分布を代替するものとしての性格を持ち，また Petruaskas と Aagaard の提案以来わが国の波浪統計で使われてきたこととの連続性を重んじて，このように選定したものである。

したがって，以下の検討においては次の **5** 分布関数をあてはめ関数として使用する。

FT-I 型分布

ワイブル分布：$k=0.75$，1.0，1.4，および2.0

(3) 5分布関数によるあてはめの適合率

表-8〜10は，上記の5分布関数をそれぞれ母関数として抽出された標本がどのような分布に最も適合するかを

表-10　各母関数に対する最適合分布の出現率（0.1%単位）
—データ採択率 $\nu=0.25$ の場合—

あてはめ関数 ＼ 母関数		N	ワイブル分布 $k=0.75$	ワイブル分布 $k=1.0$	ワイブル分布 $k=1.4$	ワイブル分布 $k=2.0$	FT-I 型分布	計
ワイブル分布	$k=0.75$	10	**424**	335	274	215	315	1563
		20	**446**	325	225	159	300	1456
		40	484	316	173	81	272	1327
		100	**566**	288	93	24	218	1189
	$k=1.0$	10	83	**77**	73	67	74	373
		20	117	**112**	94	78	96	497
		40	160	**144**	115	77	136	632
		100	212	**213**	123	50	190	788
	$k=1.4$	10	74	75	**70**	68	71	357
		20	98	108	**105**	94	107	513
		40	119	145	**150**	136	143	692
		100	109	190	**241**	189	202	931
	$k=2.0$	10	387	471	539	**606**	499	2502
		20	286	398	511	**605**	432	2233
		40	175	306	469	**616**	356	1921
		100	58	179	401	**643**	243	1523
FT-I 型分布		10	33	43	45	44	**42**	205
		20	52	56	65	64	**64**	302
		40	63	89	92	91	**94**	429
		100	55	130	142	94	**147**	569

示したものである。順に，データ採択率が $\nu=1.0$，0.5，および0.25の結果である。シミュレーション回数は $N=10$～40は10,000回，$N=100$は5,000回，$N=200$は2,000回である。出現率は 0.1%（パーミル）単位で表示してある。母関数が最適合分布となる場合は太字で示している。

まず，データ採択率が $\nu=1.0$ の場合を見ると，母関数への的中率が50%を超えるのは標本の大きさが60程度以上の場合と推測される。ただし，FT-I 型分布では $N\geq100$の場合である。データ個数が200を超えるような大きな標本であれば70%以上の割合で母分布関数に復帰する。逆に，$N=10$のような小標本であると母関数以外のものを最適合と判定する割合が高く，特に FT-I 型分布は母関数への的中率が低い。もっとも，$k=0.75$ および $k=2.0$ のワイブル分布の場合は母関数への的中率が高く，分布関数による違いが大きい。なお，表の最右欄は各行の数値を集計したものである。物理的な意味は薄いが，あえて意味付けをすると，次のようになる。すなわち，自然界にこの5分布関数を持つ5個の母集団が 1/5 ずつの割合で混在するとき，この全体集合からランダムに抽出された標本資料に対して最適合と判断される分布関数の度数が表の最右欄の数値である。たとえば $N=10$ の標本であれば $k=2.0$ のワイブル分布が最適合とされる回数は FT-I 型分布に対する回数の4倍以上となる。

次に，データ採択率 $\nu=0.5$ の場合は $\nu=1.0$ のときよりも母関数への的中率が全般的に低下し，また分布形状のかなり異なる関数を最適合とするケースが増える。すなわち，出現率が一様化に向う傾向にある。これは，先に示した図-1，2のような確率密度関数の右半分だけから全体を推定しようとするために分布関数の判別がむずかしくなることの影響である。

データ採択率が $\nu=0.25$ とさらに低くなると，この傾向が一層強くなる。分布形状が両極端である $k=0.75$ のワイブル分布および $k=2.0$ のワイブル分布は母関数に復帰する割合が高いけれども，分布形状がこれらの中間である $k=1.0$ および 1.4 のワイブル分布ならびに FT-I 型分布は両極端に引き寄せられてしまい，標本が小さいときにその傾向が顕著である。

表-8～10の結果は，標本データに対して最小2乗法による分布関数のあてはめを行い，相関係数の絶対値の大小で適合度を判定する方式の場合のものである。分布関数のあてはめを別の方法で行い，別の指標を用いて適合度を判定すると，ここに示したものとは若干異なる結果が得られるであろう。しかし，ランダム抽出標本におけるデータの変動性を考えるならば，母分布関数への的中率が大幅に向上するとは考えられない。これは4.2で述べた分布関数の母数の推定値の信頼区間について最尤法による理論値と比較検討した結果から推察されるところである。

5.3 再現確率統計量の推定値に対する偏倚補正

(1) 基本的な考え方

極値統計の一つの標本資料が与えられたとき，分布形状の異なる関数をあてはめてみてその中で最も適合する分布関数を選択するというのが最も合理的と考えられるけれども，実際問題としては表-8～10に示されたように標本中のデータ数Nが相当に大きくなければ正しい分布関数を選び出すことは困難である。たとえば，与えられた標本資料に最も適合する分布が $k=2.0$ のワイブル分布であると判定されたとしても，その標本が本当は $k=0.75$のワイブル母集団に所属するものであって，たまたま抽出されたデータが $k=2.0$ と見誤られやすいものであった可能性を否定することはできない。繰り返して述べるように，我々が知りたいのは母集団における再現確率統計量の値であって，標本データに適合する分布関数を外挿した値ではない。標本データに特定の分布関数をあてはめたとき，ぴたりと1直線上に乗ったからといって，その直線が母集団の分布関数に一致しているとは全くいえず，そうした可能性はむしろ低いと見るべきである。

上述の例のように，$k=0.75$ のワイブル母集団に属する一つの標本がたまたま $k=2.0$ のワイブル分布に適合したとすると，その分布関数を外挿して得られる再現確率統計量の推定値は母集団における真値よりもかなり低い値になる公算が大きい。すなわち，負の偏り量を伴うことになる。この偏り量の大きさは，母分布関数が特定されれば数値シミュレーション結果に基づいて推定可能である。しかし，極値統計の問題，特に波浪データに関しては母分布関数が未知であることから議論が出発している。母集団の分布関数を仮定できるのであれば，4.で述べたように適切なプロッティング公式を採用することによって偏り量を微小とすることができ，問題は生じない。

この矛盾を解決する一つの方法は，対象とする極値の母集団としては5種類の分布関数の取り得る可能性を確率の形で表し，確率計算によって偏り量の期待値を計算することである。いま，l 番目の分布関数$F_{(l)}$の存在確率を p_l とし，$F_{(l)}$ を母分布関数とする標本のうちでn番目の分布関数が最適合分布であると判定される割合を $Q_{l,n}$ とする。これは 表-8～10 の中の数値に比例する量と考えればよい。そして，$F_{(l)}$ の母分布関数が $F_{(n)}$ と見誤られることによって起きる再現確率統計量の偏り量の期待値（多数の繰り返し試行の平均で近似）を $\bar{Z}_{l,n}$ で表す。そうすると，最適合関数として判定されたn番目の分布関数を使ったときの偏り量 $\bar{Z}_n$ は次式で計算することができる。

$$\bar{Z}_n=\frac{\sum_{l=1}^{5} p_l Q_{l,n} \bar{Z}_{l,n}}{\sum_{l=1}^{5} p_l Q_{l,n}} \tag{96}$$

なお，$Q_{l,n}$ をnについて集計すればこれは一定値になる。表-8～10の例であれば，各欄の数値（Nが同一のもの）を縦に集計すればいずれも1000になる。（末尾の数字の四捨五入の関係で若干の誤差はある。）

式(96)を実際に適用するにあたっての問題は各分布関数の存在確率 p_l の設定である。対象とする極値データの種類によっては，既往の調査研究等によって最有力候補，最弱候補などのランク付けができる可能性もあろう。波浪データの場合にはそうした知見は現在までのところ得られていないので，最も単純な仮定である $p_1=p_2=\cdots=p_5=1/5$ を用いることにする。この結果，偏り量の期待値の計算式は次のように単純化される。

$$\bar{Z}_n=\frac{\sum_{l=1}^{5} Q_{l,n} \bar{Z}_{l,n}}{\sum_{l=1}^{5} Q_{l,n}} \tag{97}$$

表-8～10の最右欄の数値は上式の分母の値に比例する量である。

(2) 偏り量の期待値の計算

式(97)に基づいて計算を行うためには，$Q_{l,n}$ や $\bar{Z}_{l,n}$ の数値を求めておかなければならない。本報告では一連の数値シミュレーション作業によってこれを求めた。この数値計算は3.4で述べたプロッティング公式の数値比較のために行ったものと本質的には同じである。念のためにシミュレーション条件等を以下に記載しておく。

1) 母分布関数：FT-I 型分布およびワイブル分布4種
2) あてはめ関数：各母関数に対して上記の5種
3) 母　　　数：A＝1.0，B＝5.0
4) データ採択率および平均発生率の組合せ：

$\nu=1.0,\quad \lambda=10\ (\nu\lambda=10)$
$\nu=0.5,\quad \lambda=20$ （同上）
$\nu=0.25,\quad \lambda=40$ （同上）

5) 標本データ個数：$N=10,\ 20,\ 40,\ 100$, および 200

6) 繰り返し回数： $n_0=10,000$回（$N=10\sim40$）
$n_0=5,000$回（$N=100$）
$n_0=2,000$回（$N=200$）

7) 再現期間： $R=2\sim1000$（固定値）

表-8～10に記載した各分布関数の適合率の数値はこの数値シミュレーションによって得られた $Q_{l,n}$ の値である。なお，図-28 のデータはこれとは別のシリーズの計算によるものである。

再現確率統計量の偏り量 $\bar{Z}_{l,n}$ は4.3で導入した式(81)の補助統計量 z について求めた。すなわち，

$$\bar{Z}_{l,n}=\frac{1}{M_{l,n}}\sum_{i=1}^{M_{l,n}}[(\hat{x}_{R,n})_i-x_{R,l}]/(\sigma_x)_i \qquad (98)$$

ここに，$M_{l,n}$ は l 番目の母分布関数からの標本のうち n 番目の分布関数が最適合と判定された標本の数であり，$Q_{l,n}$ を整数化した数である。i はそうした標本の番号，$(\sigma_x)_i$ は i 番目の標本の標準偏差，$(\hat{x}_{R,n})_i$ は i 番目の標本に対して n 番目の関数をあてはめて求めた x_R の推定値，$x_{R,l}$ は母集団である l 番目の分布関数に基づく再現確率統計量すなわち真値である。当然のことながら，偏り量 $\bar{Z}_{l,n}$ は再現期間 R の増加につれて増大する。

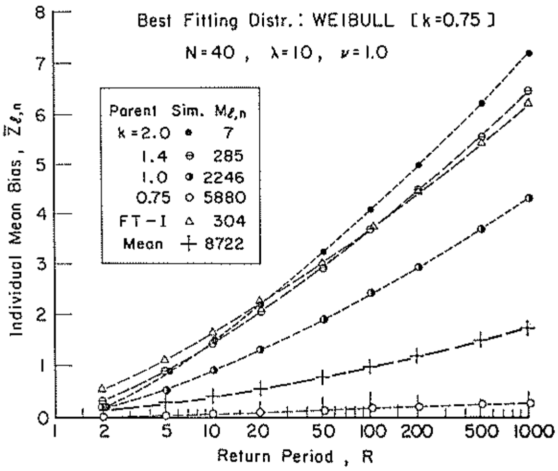


図-30 再現確率統計量の推定値の偏り量 $\bar{Z}_{l,n}$ の例

図-30は数値シミュレーションで得られた偏り量 $\bar{Z}_{l,n}$ の1例である。標本として $N=40$の大きさの全数極値資料を対象とし，5種類の母分布関数から抽出された標本のうちで $k=0.75$ のワイブル分布が最適合すると判定されたものについて式(98)で計算した結果である。図の横軸は再現期間 R である。各母関数から $k=0.75$ のワイブル分布に移った標本の個数 $M_{l,n}$ は図の凡例に記入してあり，各10,000回の試行（全体で50,000回）のうち $k=0.75$のワイブル分布は 8,722 回が最適合と判定されている。こうしたデータから式(98)によって偏り量の期待値 $\bar{Z}_n$ を計算した結果は，図中に＋印で記入してある。この場合は $k=0.75$ のワイブル分布を母関数とするデータが多いため，$\bar{Z}_n$ は比較的小さい値である。

こうした計算を $N=10\sim200$ の標本についても実施し

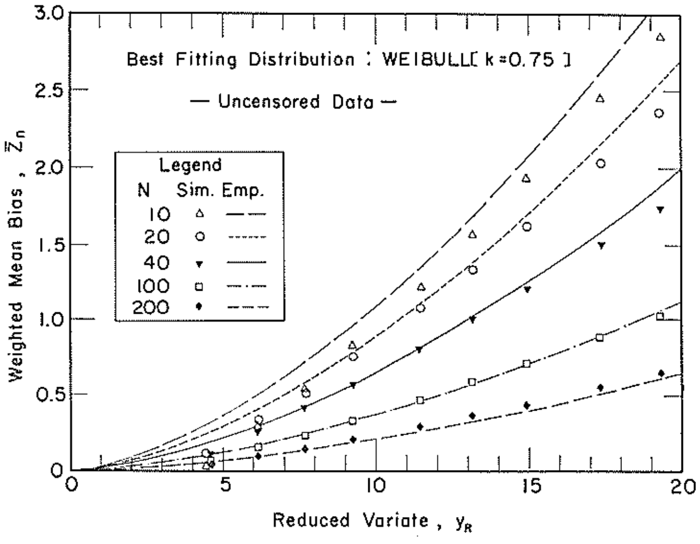


図-31 ワイブル分布（$k=0.75$）における偏り量の期待値 $\bar{Z}_n$ の算定結果

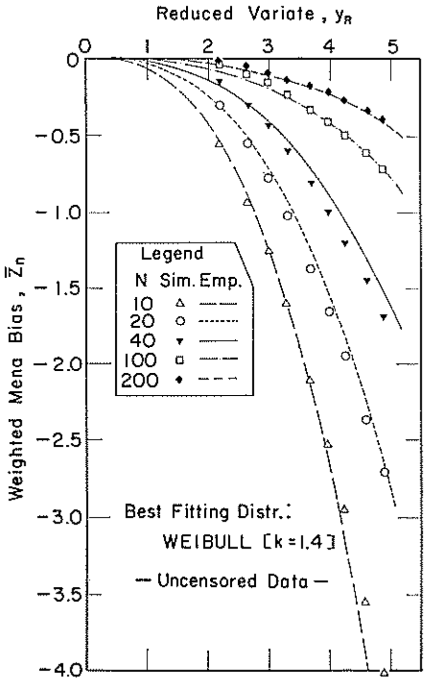


図-32 ワイブル分布（$k=1.4$）における偏り量の期待値 $\bar{Z}_n$ の算定結果

た結果を示した1例が図-31である。これも $k=0.75$ のワイブル分布が最適合と判定された標本についての結果であり，図-30 に記載した $\bar{Z}_n$ の計算結果もこの図の中に収録されている。ただし，横軸は基準化変量に戻してあり，また縦軸の縮尺も変えてある。図中の曲線は後述する実験式による推定値を表す。

$k=0.75$ のワイブル分布は他の4分布よりも分布の裾が長いため，この分布が最適合と判定されると再現確率統計量は母集団の真値よりも大きな値として推定される。したがって，偏り量の期待値は正の量となる。逆に，$k=1.4$ あるいは $k=2.0$ のワイブル分布が最適合と判定されると，これらの関数は分布幅が狭いため，再現確率統計量の推定値は真値よりも小さくなる確率が大きく，偏り量の期待値は負の量となる。図-32は $k=1.4$ のワイブル分布の場合の結果である。$k=1.0$ のワイブル分布が最適合と判定された場合も $\bar{Z}_n$ は負値をとるが，比較的小さな値である。FT-I 型分布の場合は，標本の大きさによって偏り量の期待値が正になったり，負になったりし，また再現期間の増大につれて負から正に変ったりするなど複雑な様相を示す。

データ採択率 ν の影響は，**4.3** の**図-24～26** と同様に，データを全体として y_R 軸の右方へずらす形で現われる。図-33はこれを示すもので，$k=0.75$のワイブル分布が最適合と判定された $N=10$ の標本について $\nu=1.0$, 0.5，および 0.25 のケースを比較したものである。図中の曲線は次項に述べる実験式の値を示すもので，この場合に関しては数値シミュレーションの結果と若干の差異が残っている。

(3) 偏り量の期待値に対する実験式のあてはめ

$\bar{Z}_n$ に対する実験式としては，分布関数が既知の場合の $\hat{x}_n$ の標準偏差に関する式(90)の形を参考にし，また次節に述べる推定値の信頼区間に関する実験式とも関連性を保つように考えて，次のような形を設定した。

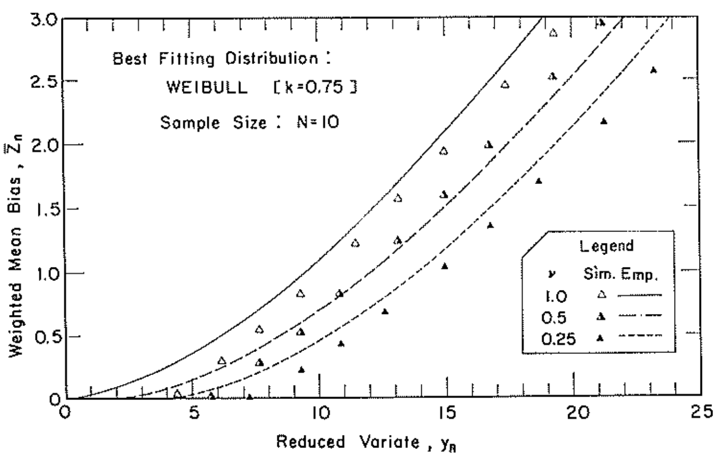


図-33 ワイブル分布（$k=0.75$）の部分極値資料における偏り量の期待値 $\bar{Z}_n$ の算定結果の例

$$\bar{Z}_n=\begin{cases}A_c(y_R+\alpha\ln\nu)^q & : y_R>-\alpha\ln\nu\\ 0 & : y_R\leq-\alpha\ln\nu\end{cases}\quad (99)$$

ここに，係数 A_c，α，および指数 q は分布関数ごとにそれぞれ次のように定めた。

1) ワイブル分布（$k=0.75$）

$$\left.\begin{array}{l}A_c=\begin{cases}0.030\exp[-0.6(\log_{10}N/4)^2] & : \nu=1.0\\ 0.025\exp[-0.7(\log_{10}N/15)^2] & : \nu=0.5,0.25\end{cases}\\ q=1.6,\quad \alpha=2.7\end{array}\right\}\quad (100)$$

2) ワイブル分布（$k=1.0$）

$$\left.\begin{array}{l}A_c=\begin{cases}-0.028\,N^{-0.25} & : \nu=1\\ -0.0022-0.006(\log_{10}N/50)^2 & : \nu=0.5,\ 0.25\end{cases}\\ q=2.1,\quad \alpha=1.0\end{array}\right\}\quad (101)$$

3) ワイブル分布（$k=1.4$）

$$\left.\begin{array}{l}A_c=\begin{cases}-0.40\,N^{-0.8} & : \nu=1\\ -0.10\,N^{-0.4} & : \nu=0.5,\ 0.25\end{cases}\\ q=2.7,\quad \alpha=0.5,\end{array}\right\}\quad (102)$$

4) ワイブル分布（$k=2.0$）

$$\left.\begin{array}{l}A_c=\begin{cases}-0.50\,N^{-0.7} & : \nu=1\\ -0.64\,N^{-0.6} & : \nu=0.5,\ 0.25\end{cases}\\ q=3.4,\quad \alpha=0.35\end{array}\right\}\quad (103)$$

5) FT-I 型分布

i) $\nu=1.0$

$$\left.\begin{array}{l}A_c=\begin{cases}0.046-0.40(\log_{10}60/N)^3 & : N<60\\ 0.046\exp[-2.5(\log_{10}N/60)^2] & : N\geq 60\end{cases}\\ q=1.0,\quad \alpha=0.9\end{array}\right\}\quad (104)$$

ii) $\nu=0.5,\ 0.25$

$$\left.\begin{array}{l}A_c=0.01-0.044(\log_{10}N/300)^4\\ q=1.0,\quad \alpha=0.9\end{array}\right\}\quad (105)$$

これらの係数を定めた方法は次のとおりである。まず y_R軸の移動量に関する α については，次節の信頼区間のデータも同時に考え，両者に対して分布関数ごとに同一の値となるようにデータの傾向から判断して定めた。次に指数 q については再現期間 $R=10$および $R=200$の 2 点に対して式(99)を適用してその初期値を求めた。q の初期値は同一分布関数内でも標本の大きさやデータ採択率によってある程度変化したけれども，統一性を保つため分布関数ごとの平均値を求め，それを丸めた数値とした。そして，A_c については以上で定めた α と q を使って $R=100$ の点に対する係数値を求め，これをNの関数として表示した。データ採択率 $\nu=0.5$ と $\nu=0.25$ とで A_c の値に若干の差異が残っていたけれども，ν による A_c の変化の定式化は困難であったので，式(100)～(105)のように $\nu=1$ の全数極値資料に対するものと，$\nu=0.5$ および0.25の部分極値資料に対するものの 2 本立てとしてとりまとめたものである。

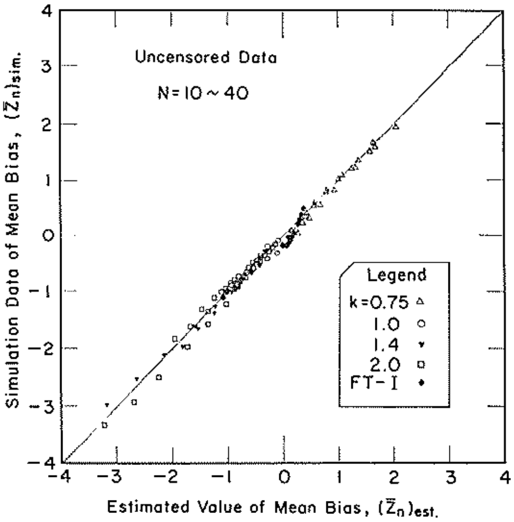


図-34　全数極値資料における偏り量の推定値と数値実験データとの比較

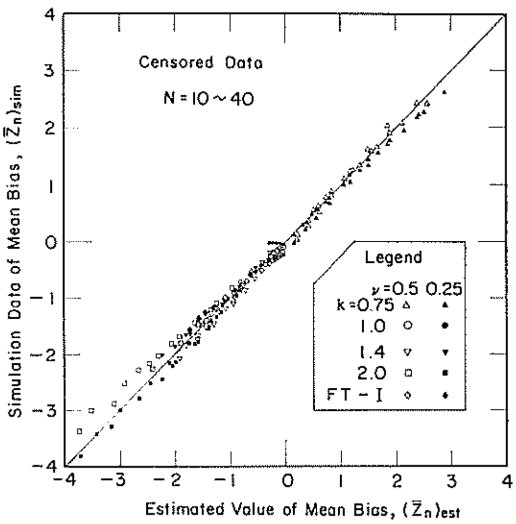


図-35　部分極値資料における偏り量の推定値と数値実験データとの比較

数値シミュレーション結果に対する実験式の適合度は図-34，35に示すとおりである。前者は $\nu=1.0$，後者は $\nu=0.5$ と $\nu=0.25$ に対するものである。いずれも$N=100$および200のデータは偏り量がもともと小さいため，図の混乱を避ける目的で省略してある。また，$R>20N_T$ のデータはそのように遠くまで外挿することは信頼度の

点で問題があるのでプロットしていない。

図-34，35を見ると，分布関数およびデータ採択率による固有の癖が少し残っているようでもあるが，シミュレーション結果と推定値の差は大半のものが0.2以下であり，偏り量の絶対誤差は標本の標準偏差の約0.2倍以下と見積られる。

(4) 再現確率統計量の推定値の補正方法

ここまで検討してきたのは，無次元量として式(98)で定義された補助統計量の平均値 $\bar{Z}_{l,n}$ を基礎データとしたものである。式(98)は，個々の標本について $(\hat{x}_R-x_R)/\sigma_x$ の比を求め，それを $M_{l,n}$ 個について計算したものである。そして，その $\bar{Z}_{l,n}$ を $l=1\sim5$ の母分布関数について重み付き平均をとった結果が偏り量の期待値 $\bar{Z}_n$ として計算され，これは式(99)の経験式によって推定可能となったものである。

実際の問題では，1組の標本に対して最適分布関数のあてはめが行われ，所定の再現期間 R に対する R 年確率統計量の推定値 $\hat{x}_R$ が求められる。この推定値は今までの検討で明らかなように，母分布関数が未知であることに起因する偏り量を伴っている。この補正方法として考えられる唯一の方法は，式(98)の級数和の各項の構造形を利用して次のようにすることである。

$$\left.\begin{aligned}(x_R)_{\mathrm{cor}}&=\hat{x}_{R,n}+\Delta x_R\\ \Delta x_R&=-\bar{Z}_r\,\sigma_x\end{aligned}\right\}\qquad(106)$$

しかしながら，この補正にしても完全なものではないことに注意する必要がある。それは，補正量 Δx_R の絶対値を定める偏差 σ_x が与えられた標本に特有な値をとり，$l=1\sim5$ の母集団から n 番目の分布関数を最適合分布とする標本の平均的な値にはなっていないためである。このため，$\bar{Z}_n$ が同一であっても標本データの σ_x が大きければ補正量も大きく，σ_x が小さければ補正量も小さくなる。すなわち，標本資料の変動性の影響は R 年確率統計量の補正においても現われる。このことは次節に述べる信頼区間の推定においても同様に発生する。ランダム変動現象を対象として統計解析を行う以上，こうした不確定性は受け入れざるを得ないといえる。なお，実験式の使用法の説明として以下に計算例を示しておく。

［計算例1］

ある港の9月の月最大有義波高のデータが $K=12$ 年分取得でき，1.90～5.68mの範囲に分布した。このデータの平均値および標準偏差は

$$\bar{x}=3.723\mathrm{m},\quad \sigma_x=1.250\mathrm{m}$$

であり，分布関数としては $k=2.0$ のワイブル分布が最も適合度が高かった。尺度母数および位置母数の推定値は次のとおりであった。

$$\hat{A}=2.787\mathrm{m},\quad \hat{B}=1.253\mathrm{m}$$

このデータについて30年確率波高の推定を行う。データは毎年1個の期間最大値資料であり $\lambda=1.0$ である。また全数極値資料に属し $\nu=1.0$ である。まず，偏り補正前の推定値を求めると，式(77)により

$$y_{30}=[-\ln(1/30)]^{1/2}=1.8442$$

$$\hat{x}_{30}=2.787\times1.8442+1.253=6.392\mathrm{m}$$

これに対して偏り補正のための補助統計量 $\bar{Z}_n$ を求めると，$N=12$ であるから式(101)によって係数が

$$A_c=-0.50\times12^{-0.7}=-0.0878$$

$$q=3.4$$

と求められる。この数値を式(99)に代入すると

$$\bar{Z}_n=-0.0878\times1.8442^{3.4}=-0.704$$

したがって，偏り補正量としては

$$\Delta x_{30}=-\bar{Z}_n\cdot\sigma_x=-(-0.704)\times1.250=0.88\mathrm{m}$$

となり，約14%の補正が必要と推定される。

以上により，30年確率波高の推定値は次のようになる。

$$x_{30}=\hat{x}_{30}+\Delta x_{30}=7.3\mathrm{m}$$

ここで数値を小数点以下1桁にとどめたのは，確率波高推定の不確定性を考慮したものである。

5.4 再現確率統計量の推定値の信頼区間

(1) 信頼区間の指標としての標準誤差の計算式

母分布関数が既知の場合には**4.3**で論じたように非超過百分率の分位値を使い，推定値の分布範囲を求めることが可能である。しかし，本章で扱っているような分布関数が未知の場合は分位値の計算が繁雑であり，得られた結果の取り扱いも面倒である。そこで，本報告では信頼区間の代表値として標準誤差のみを求めることとした。

再現確率量の推定値の標準誤差を求める際には，何らかの無次元化を行う必要があり，このため**4.3**と同様に式(81)の補助統計量 $z=(\hat{x}_R-x_R)/\sigma_x$ を用いる。そして偏り量 $\bar{Z}_{l,n}$ と同様に，母分布関数における真値 $x_{R,l}$ と最適合分布関数による推定値 $\hat{x}_{R,n}$ との2乗平均残差を次のように定義する。

$$S_{l,n}{}^2=\frac{1}{M_{l,n}}\sum_{i=1}^{M_{l,n}}[(\hat{x}_{R,n})_i-x_{R,l}]^2/(\sigma_x)_i{}^2\qquad(107)$$

他の記号の定義は式(98)と同じである。

最適合関数として判定された n 番目の分布関数を使って x_R を推定したときの標準誤差の期待値は，式(96)と同一の考え方により，一種の重み付き平均として計算する。その際は，式(97)と同様に各母関数の存在確率はすべて等しく $p_l=1/5$（ここに，$l=1\sim5$）であると仮定す

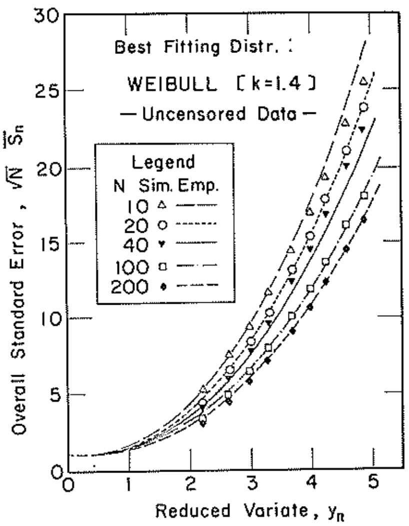


図-36 ワイブル分布（k=1.4）における再現確率統計量の標準誤差の算定結果

る。 n 番目の分布関数に対する標準誤差の期待値を S_n と表すと，この算定式は次のようになる。

$$S_n=\left(\frac{\sum_{l=1}^{5} Q_{l,n}\, S_{l,n}^2}{\sum_{l=1}^{5} Q_{l,n}}-\bar{Z}_n^2\right)^{1/2} \qquad (108)$$

実際の数値計算では $Q_{l,n}=M_{l,n}$ としている。式（108）は（97）式と同形式であるが，右辺第2項に平均値の偏り量 $\bar{Z}_n$ が補正項として入れてある。これは S_n として偏り補正を行った $(x_R)_{cor.}$ からの標準誤差を求めるためである。

(2) 標準誤差の期待値に対する実験式のあてはめ

5.3で述べた数値シミュレーションの結果から式(108)に基づいて $\bar{S}_n$ を計算した結果の1例を図-**36** に示す。標準誤差は標本の大きさ N の増加につれて減少するので，ここでは$\sqrt{N}\,S_n$の形で表示してある。この図は k=1.4 のワイブル分布が最適合と判定されたデータについてのもので，ν=1.0の全数極値資料の場合である。部分極値資料における ν の影響は 図-**37** にその1例を示すように，偏り量の期待値と同様にデータを全体として右方へずらす形で現われる。両図とも図中の曲線は以下の実験式を表している。

こうした再現確率統計量の推定値の標準誤差に対する実験式としては，ここでは次の形のものを設定した。

$$\sqrt{N}\,S_n=1.0+A_\sigma\,|\,y_R+\alpha\ln\nu\,|^{\,q} \qquad (109)$$

上式の絶対値記号は $y_R+\alpha\ln\nu$ が負の領域でも有限な大きさの S_n を与えるためのものである。

上式中の係数のうち α は **5.3** で述べたように $\bar{Z}_n$ と共通とし，係数 A_σ および指数 q は $\bar{Z}_n$ に対するものと同じ手法を使って定めたところ次のような結果を得た。

1) ワイブル分布（k=0.75）

$$\left.\begin{array}{l} A_\sigma=\begin{cases}0.57+0.18(\log_{10}N/20)^2 & :\nu=1.0\\ 0.41+0.22(\log_{10}N/20)^2 & :\nu=0.5,\ 0.25\end{cases}\\ q=1.2,\quad \alpha=2.7\end{array}\right\} \qquad (110)$$

2) ワイブル分布（k=1.0）

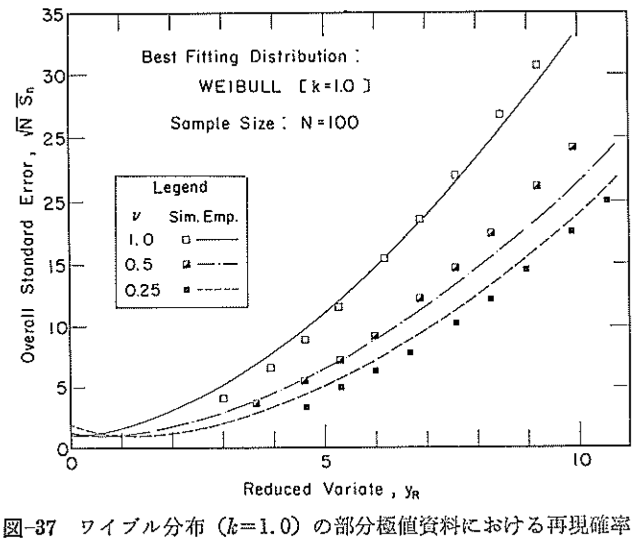


図-37 ワイブル分布（k=1.0）の部分極値資料における再現確率統計量の標準誤差の算定結果の例

$$A_\sigma=\begin{cases}0.55+0.15(\log_{10}N/15)^2 & :\nu=1.0\\ 0.38+0.17(\log_{10}N/20)^2 & :\nu=0.5,\ 0.25\end{cases}$$
$$q=1.7,\quad \alpha=1.0 \tag{111}$$

3) ワイブル分布 ($k=1.4$)

$$A_\sigma=\begin{cases}0.37+0.08(\log_{10}N/1000)^2 & :\nu=1.0\\ 0.46+0.09(\log_{10}N/20)^2 & :\nu=0.5,\ 0.25\end{cases}$$
$$q=2.3,\quad \alpha=0.5 \tag{112}$$

4) ワイブル分布 ($k=2.0$)

$$A_\sigma=\begin{cases}0.30+0.36(\log_{10}N/80)^2 & :\nu=1.0\\ 0.56+0.20(\log_{10}N/100)^2 & :\nu=0.5,\ 0.25\end{cases}$$
$$q=3.2,\quad \alpha=0.35 \tag{113}$$

5) FT-I 型分布

$$A_\sigma=\begin{cases}0.24+0.36(\log_{10}N/80)^2 & :\nu=1.0\\ 0.46+0.14(\log_{10}N/50)^2 & :\nu=0.5,\ 0.25\end{cases}$$
$$q=1.6,\quad \alpha=0.9 \tag{114}$$

これらの係数値を使い，式(109)によって標準誤差の期待値 S_n を推定した結果を数値シミュレーションにおける計算値と比較したのが図-38，39である。前者は $\nu=1.0$ の全数極値資料，後者は $\nu=0.5$ および0.25の部分極値資料に対応するものである。いずれも図の混乱を避けるため $N=10$ および100のデータのみをプロットしてある。$R>20N_T$ のデータを除いているのも図-34，35の場合と同じである。

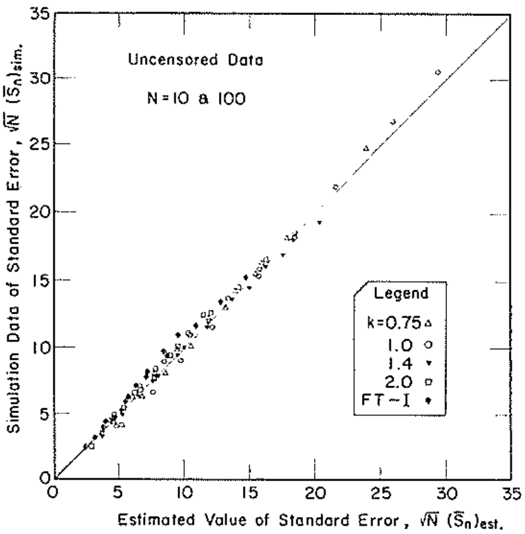


図-38 全数極値資料における再現確率統計量の標準誤差の推定値と数値実験データとの比較

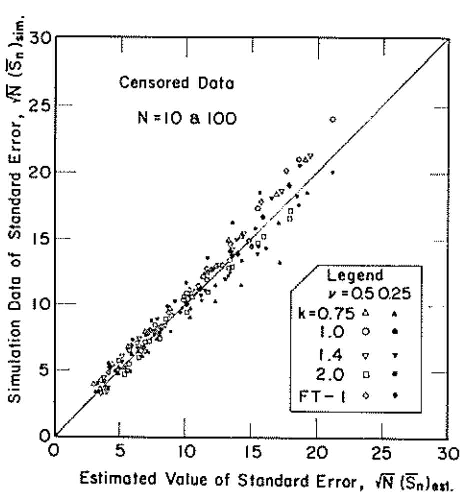


図-39 部分極値資料における再現確率統計量の標準誤差の推定値と数値実験データとの比較

まず，図-38の全数極値資料の場合は数値実験データが実験式によって的確に表現されており，実用上も十分な精度ではないかと思われる。これに対して図-39の部分極値資料は推定値と数値実験値との差がかなりあり，20%以上の差を示すものもある。しかし，今回の数値シミュレーションの範囲内で得られたデータの傾向が分布関数およびデータ採択率ごとにかなり異なっており，これ以上の推定精度を上げることは困難と判断し，とりあえず式(109)を式(110)～(114)の係数とともに提案するにとどめた。

なお，母分布関数が既知の場合には既に4.3で述べたように，再現確率統計量の推定値に対する標準偏差の実験式が式(90)のように統一的にまとめられ，全数極値資料と部分極値資料の両者を包含する形になっている。母分布関数が未知の場合の式(109)では，その係数を与える式が全数極値資料と部分極値資料とで別の形に組み立てられている。このため，今回は検討していない ν の範囲，すなわち，$0.5<\nu<1.0$ および $\nu<0.25$ の範囲の処理が今後の課題として残されている。当面の処置としては，格別の指針があるわけではないけれども，$0.8<\nu\leq 1$ は $\nu=1.0$ に対する実験式を適用し，$\nu\leq 0.8$ は $\nu=0.5$ および0.25に対する実験式を使うようなことで割り切らざるを得ない。ただし，ν が非常に小さいところまでそのまま適用することは疑問があるので，たとえば $\nu<0.15$ のものは式(109)の適用の際は $\nu=0.15$ と見なすというような便法をとる必要があろう。偏り量の補正に対する式(99)についても同様である。

(3) 信頼区間の推定方法

標準誤差が無次元量 S_n の形で推定されれば，その次元量は式 (105) の２乗平均残差の定義式を参照して次のように計算すればよい。

$$\sigma(\hat{x}_R)=S_n\sigma_x \qquad (115)$$

5.3 で示した計算例について 30年確率波高の標準誤差を推定すると次のようになる。

［計算例２］

まず，係数 A_σ を式(111)によって計算する。

$$A_\sigma=0.30+0.36(\log_{10}12/80)^2=0.544$$

基準化変量は $y_{30}=1.8442$ であるから，$q=3.2$の指数とともに式(107)に代入すると

$$\sqrt{N}\ S_n=1.0+0.544\times1.8442^{3.2}=4.86$$

したがって，

$$S_n=4.86/\sqrt{12}=1.40$$

結局，30年確率波高 $(x_{30})_{\mathrm{cor.}}=7.3\mathrm{m}$ に対する標準誤差の推定値は

$$\sigma(x_{30})=1.40\times1.25=1.75\mathrm{m}$$

信頼区間として１シグマ限界を用いると

$$\hat{x}_{30}=(7.27-1.75)\sim(7.27+1.75)=5.5\sim9.0\mathrm{m}$$

同じく２シグマ限界を用いると

$$\hat{x}_{30}=(7.27-2\times1.75)\sim(7.27+2\times1.75)$$
$$=3.8\sim10.8\mathrm{m}$$

一般に統計量が正規分布を従うとき，１個の標本が１シグマ限界内に入る確率は68%，２シグマ限界内に入る確率は 95% である。$\hat{x}_R$ の分布は正規分布からかなりずれた形状を取るものと考えられるけれども，現在のところ適当な推定方法がないので，とりあえずは正規分布による限界値を代用せざるを得ない。所定の再現期間に対する再現確率統計量を推定する場合には，中央値単独では情報として不十分であり，上記の例のように１シグマあるいは２シグマ限界の上・下限値として与えるべきであろう。

6. 極値統計解析におけるその他の諸問題

6.1 発生原因が異なる極値統計資料の取り扱い

(1) 基本的考え方

既に 2.1 で述べたように，ランダム変量を統計解析する際には対象とするデータがすべて同一の性質を有する母集団（分布関数は未知であってよい）に所属することについての心証を得ておく必要がある。すなわち，等質性(homogenuity)の確認である。波浪の場合であれば，台風による高波と低気圧による高波とは分離し，必要によっては冬型気圧配置に伴う高波も別立てとしそれぞれ別個のデータとして解析することが望ましい。このように，同一の地点において複数の極値統計資料が得られた場合，その結果をどのように取り扱うべきかの問題が生じる。

これに関して Carter と Challenor[39] の研究は一つの示唆を与える。この研究は，風速や波高のように年間で季節的に変化する現象については単純に毎年最大値を用いると季節的変化のために等質性の条件を満たさなくなることを指摘し，結果として再現確率統計量が低目に推定されることを例示したものである。この解決策として二人は毎月最大値の資料を使い，１～12月の月別に解析して12個の最適合分布関数 F_i を求め，この乗積として非超過確率を求めることを提案した。すなわち

$$F(x)=\mathrm{Prob.}[X\leqq x]=\prod_{j=1}^{12}F_j(x) \qquad (116)$$

この式は次のように解釈される。ある年にXがxの値を超えたとする。この事象は１月に起きたかも知れないし，８月に起きたかも知れない。１回だけかも知れないし，５回起きたかも知れない。したがって $X>x$の事象の確率を求めることは非常に複雑になる。しかし，問題を逆から見てXがxを超えないという非超過の事象について考えると，この非超過の事象が成立するためには１～12月のどの月においてもXはxを超えてはならない。すなわち，各月の非超過の事象が同時に成立しなければならない。したがって，１年間にXがxを超えない確率は，確率の定義によって毎月の非超過確率の乗積として与えられる。

この考え方は，台風，低気圧，冬型気圧配置等に伴う高波の極値分布についても同じように適用できる。波高値xを超える事象はいずれの気象原因によっても起り得ることであり，台風によって起きれば低気圧では起きないという排反事象ではない。したがって，n 個の気象原因別の分布関数 F_j が得られたとすると，これらを総合した分布関数としては次式で求めればよいことになる。

$$F(x)=\prod_{j=1}^{n}F_j(x) \qquad (117)$$

(2) 平均発生率 λ の調整

Carter と Challenor[39] が扱った例は毎月最大値という期間最大値資料であって月ごとに毎年１個ずつのデータであり，平均発生率はいずれも $\lambda=1$ である。しかし，波浪データでは極大値資料の形で極値解析を行う場合が多く，年間数個から数十個の発生事象が対象になる。しかも気象原因ごとに平均発生率が異なる。台風であれば年間数個，低気圧性であれば数十個である。4.3 で式(74)～(77)として提示したように，再現確率統計量の推

定においては超過確率を $1/\lambda R$ として算定するので，λ の値が異なる分布関数をそのまま式(116)に代入することはできない。

この解決策として，極大値資料に対する分布関数を毎年最大値資料に対するものに変換することを考える。具体的には，**2.4** で導いたポアソン分布の仮定に基づく式(31)を用いる。この結果，

$$F(x)=\prod_{j=1}^{n}\exp\{-\lambda_j[1-F_j(x)]\}$$
$$=\exp\left\{-\sum_{j=1}^{n}\lambda_j[1-F_j(x)]\right\} \qquad (118)$$

を得る。ただし，この方式では毎年最大値としての非超過確率であるので，再現期間1年またはそれ未満の値に対する再現確率統計量は定義上存在しない。また，式(33)に関連して述べたように，毎年最大値としての統計値は極大値分布に基づく値よりも再現期間に換算して約1/2年だけ短い期間に相当する値である。計算結果の利用に際してはこうした点に留意する必要がある。

(3) 乗積確率計算における偏り補正の方法

極値統計量が単一の分布関数で与えられたときは **4.3** に述べたようにして再現確率統計量が推定できる。分布関数が式(117)あるいは式(118)の乗積確率として求められたときは，x と $F(x)$ の関係をあらかじめ所定の x の刻みごとに計算しておき，そのあとで $F=1-1/R$ に対応する $\hat{x}_R$ を内挿法で推定する。

この際，分布関数として特定の関数が存在するわけではないので，**5.3**および**5.4**で提案した推定値に対する偏り補正および信頼区間推定の経験式が使用できない。これについては，次のような近似的方法を使うことが考えられる。

再現確率統計量の推定値に対する偏り補正量は式(106)のように $\Delta x_R=-\bar{Z}_n\,\sigma_x$ で与えられるが，この係数 $\bar{Z}_n$ は式(99)のように基準化変量 y の関数として組み立てられている。母数推定値が与えられていれば，y は式(20)によって直ちに計算される。したがって，乗積確率の計算においては x の設定値 x_u に対して基準化変量 y_u を算定し，これに基づいて偏り量補正量 Δx_u を計算する。そして，Δx_u だけずらした x の点，すなわち $x_u'=x_u-\Delta x_u$ における $F_j(x_u')$ の値を乗積確率計算用の $x=x_u$ における非超過確率と見なせばよい。

[計算例3]

5.3 の計算例1を使い，$x_u=7.3\mathrm{m}$ に対する非超過確率を偏り補正を行って推定する。まず，y_u の値は

$$y_u=(7.3-1.253)/2.787=2.170$$

偏り補正のための実験式の係数 A_c，q は計算例1で既に求められているので，これを使うと $\bar{Z}_n$ が次のように求められる。

$$\bar{Z}_n=-0.0878\times2.170^{3.4}=-1.223$$

偏り補正量 Δx_u および修正値 x_u' は次のようになる。

$$\Delta x_u=-(-1.223)\times1.250=1.529\mathrm{m}$$
$$x_u'=7.3-1.529=5.771\mathrm{m}$$

この点における非超過確率は

$$F(x_u')=1-\exp\{-[(5.771-1.253)/2.787]^2\}$$
$$=0.92777$$

この計算例の場合，得られた非超過確率 $F(x_u')$ に対する再現期間を求めると $R=13.9$年に相当する。計算例1と比べてみると分るように，$x=7.3\mathrm{m}$ は $R=30$ 年に対して偏り補正を行った推定値であるので，上記の結果は乗積計算のための偏り補正がやや行き過ぎであったことになる。この Δx_u による補正の適否は分布関数によって異なり，$k=2.0$のワイブル分布では過大補正の傾向があるのに対し，$k=0.75$ のワイブル分布では補正不足の傾向にある。分布関数の特性に応じてΔx_uにある倍率を乗じて偏り補正を行う方法等，今後さらに検討を進める必要がある。実際の問題に対しては，上記の方法によって偏り量を補正した計算と未補正の計算を同時に実施し，また各気象原因別の再現確率統計量の推定結果も併せて参照して判断する必要があるものと思われる。

(4) 乗積確率計算における信頼区間の推定

前項の偏り補正の場合と同じく，信頼区間についても理論的に根拠のある方法を案出することはむずかしい。本報告では信頼区間の指標として推定値からの標準誤差を用いている。標本が二つ以上ある場合の全体としての標準偏差は，それぞれの標本の分散和（標準偏差の2乗にデータ個数を乗じたもの）の合計を使って計算される。再現確率統計量の推定値に対する標準誤差は個々の標本データの標本偏差とは意味が異なるけれども，単純に同種のものと見なすと乗積確率計算における標準誤差として次のようなものを考えてもよいのではないかと思われる。

$$\sigma_{all}^2=\frac{\sum_{j=1}^{n}N_j\,\sigma_j(\hat{x}_R)^2}{\sum_{j=1}^{n}N_j} \qquad (119)$$

ここに，N_j は j 番目の発生原因別資料におけるデータ個数である。

ただし，極値の発生原因によっては同じ x_R に対する再現期間 R の値が相当に異なることがある。たとえば，台風による波であれば $H_{1/3}=7\mathrm{m}$ が30年確率であるのに季節風による波とすると1000年確率以上となるというよ

うな場合である。一般に，信頼区間の幅は再現期間が長くなるほど広くなる。このため，高波の発生頻度が低い発生原因のデータと高いデータを混ぜて式(119)を適用すると，全体としての信頼区間幅は発生頻度の低いデータに強く影響を受けてしまい不合理である。これを避けるためには何らかの出現確率的なものを重みとして導入する必要がある。そこで，各 x_R ごとの超過確率 $1-F(x_R)$ を導入し，式(119)を次のように修正する。

$$\sigma_{all}^2=\frac{\sum_{j=1}^{n}N_j[1-F_j(\hat{x}_R)]\sigma_j(\hat{x}_R)^2}{\sum_{j=1}^{n}N_j[1-F_j(\hat{x}_R)]} \tag{120}$$

実際計算においては，乗積確率 $F(x)$ と x との数値計算結果から内挿によって所定の再現期間 R に対する $\hat{x}_R$ を求めるとともに，その基礎データである個々の $F_j(x)$ について同じく内挿によって $F_j(\hat{x}_R)$ を求めればよい。発生原因別の標本における標準誤差 $\sigma_j(\hat{x}_R)$ は式(109)，(115) によって基準化変量 y_R の関数として直接に求められる。

6.2 極値統計資料の統計期間とデータ個数について

極値統計解析を行う場合に常に問題となるのが対象とするデータの統計期間の長さおよびデータ個数である。再現期間 R 年に対する再現確率統計量を推定しようとする場合，最小限何年間のデータが必要か，また毎年何個以上のデータを抽出すべきかという問題である。毎年最大値，あるいは月ごとの毎月最大値を用いる場合には統計年数 K がそのままデータ個数 N となる（データ採択率 $\nu=1.0$ とする)。極大値資料を用いる場合には $N_T=\lambda K$ である。これから，期間最大値資料と極大値資料のどちらが有利であるかという問題も生じる。

データの統計期間および個数の問題は，R 年確率統計量の推定値の信頼度の観点から判断するのが適切である。信頼度の指標として $\hat{x}_R$ の標準誤差を与えると，母分布関数を既知とする場合には **4.3** の式(90)～(94)で推定され，母分布関数が未知の場合には **5.4** の式 (109)～(115) で求められるので，設定条件ごとに試算してみればよい。ただし，母分布関数が既知・未知いずれの場合にも標準誤差の大きさはデータ個数 N，対象期間の長さ K，平均発生率 λ，およびデータ採択率 ν のパラメータによって変るので，一般的な結論を述べることはむずかしい。ここでは次の仮想的なケースについて検討してみる。

［計算例４］

ある地点で低気圧による高波の平均発生率が $\lambda=30$ 個/年であり，その波高の極大値は $k=1.0$ のワイブル分布に従い，母数は $A_1=1.0$m，$B_1=2.5$m であるとする。この高波の毎年最大波高は FT-I 型分布に従い，その母数は $A_2=1.0$m，$B_2=5.9$m と想定される。高波の極値データの標準偏差は，前者の極大値資料が $\sigma_1=1.0$m，後者の毎年最大値資料が $\sigma_2=1.283$m であるとする。この条件の下で，毎年最大値資料および極大値資料による 100 年確率波高の標準誤差を比較する。ただし，データの取得期間は $K=30$年，極大値資料については採択データ個数を30および 100 とする。

i) 毎年最大値資料の場合

まず 100 年確率波高は式(74)により

$$y_{100}=-\ln[-\ln(1-1/100)]=4.600$$

$$\hat{x}_{100}=1.0\times4.600+5.9=10.5\text{m}$$

これに対する標準誤差は母分布関数が既知であるとして式(82)を適用すると，

$$\sqrt{N}\,\sigma[\hat{x}_{100}]/\sigma_2=[1.0+0.64\times\exp(9.0\times30^{-1.3})\times4.600^2]^{1/2}=4.01$$

$$\sigma[\hat{x}_{100}]=4.01\times1.283/\sqrt{30}=0.94\text{m}$$

ii) 極大値資料の場合

100年確率波高は式(77)で $k=1.0$ と置くことにより，

$$y_{100}=-\ln(1/30/100)=8.006$$

$$\hat{x}_{100}=1.0\times8.006+2.5=10.5\text{m}$$

と同一の値が得られる（母数をそのように与えてあることによる)。

これに対する標準誤差は式(90)で求めるものとすると，$N_T=30\times30=900$ であるから，データ採択率は $N=30$ で $\nu=1/30$，$N=100$ で $\nu=1/9$ である。計算を進めると

$$N=100：a=1.92\times\exp[11.4\times100^{-1.3}]=1.976$$

$$\sqrt{N}\sigma_z=[1.0+1.976\times(8.006-0.3-0.9\times\ln 9)^2]^{1/2}=8.11$$

$$\therefore\sigma[\hat{x}_{100}]=8.11\times1.0/\sqrt{100}=0.81\text{m}$$

$$N=30：a=1.92\times\exp[11.4\times30^{-1.3}]=2.202$$

$$\sqrt{N}\sigma_z=[1.0+2.202\times(8.006-0.3-0.9\times\ln 30)^2]^{1/2}=6.96$$

$$\therefore\sigma[\hat{x}_{100}]=6.96\times1.0/\sqrt{30}=1.27\text{m}$$

結局，毎年最大値資料に基づく 100 年確率波高の標準誤差は，30年で30個採った極大値資料よりは小さく，30年で 100 個採択した場合よりは大きいことになる。この例でいえば，年平均で２～３個選んだ極大値資料と同等であり，むしろ常識的な答である。ただし，極大値資料の場合にはデータ採択率をさらに高めて多数のデータを統計処理することによって，確率波高推定値の信頼度を高める余地がある。この意味では，実測および推算の両者とも極値データの個数の増大に努め，確率波高の精度

を向上させることが望まれる。

6.3 N年確率統計量とN年最大統計量

(1) 概　説

本報告では **2.4** で述べた再現期間の概念を用い，その期間内に平均して1回はその値を超えるような統計量を再現確率統計量と呼び，その統計的性質についていろいろ検討してきた。構造物の設計において自然外力の再現期間などの程度に設定すべきは判断に迷う事項の一つであるが，一般には構造物の供用期間内に設計外力に遭遇する確率の計算を一つの判断資料としている。再現期間と遭遇確率の関係については Borgman[40] が解説しており，成書等にもいろいろ紹介[41]されているので，ここでは触れない。

一方，構造物の設計法の分野では近年，信頼性設計法が脚光を浴びており，白石・上田[42]は港湾・海洋構造物の安全性について検討を加えている。この信頼性設計法では，R年確率統計量（本節では再現期間としてR年の代りにN年の表記を使用する）ではなくて，N年中に発生すると予想される最大値すなわちN年最大値を取り上げ，永年にわたるその平均値（厳密には確率計算に基づく期待値）を設計外力の指標として用いている。N年最大値の性質については白石・上田[42]も述べているが，ここではやや異なる観点から解説しておく。

2.4 で扱ったと同じように，ある事象の毎年最大値xを対象とし，分布関数を$F(x)$とする。これは分布関数の定義によって非超過確率を表す。分布関数が極大値資料を対象にして求められているときは，ポアソン分布を仮定した式(31)によって毎年最大値の分布関数に変換することができる。

N年間中の最大値を x_N と表し，その分布関数を$\Phi_N(x_N)$とすると，$\Phi_N(x_N)$は次のように書き表される。

$$\Phi_N(x_N)=[F(x_N)]^N \qquad (121)$$

これは **6.1** において毎月最大値の分布関数から毎年最大値の分布関数を非超過確率の乗積として求めた方法と同一の考え方に基づくものである。また，Gumbel[1]が「極値統計学」の中で指数分布（$k=1$ のワイブル分布）からFT-I 型の極値分布を導いた考え方と同じである。N年最大値の確率密度関数は式 (119) の微分によって次のように求められる。

$$\varphi_N(x_N)=N[F(x_N)]^{N-1}f(x_N) \qquad (122)$$

ここに $f(x)$ はxの確率密度関数である。

なお，式(121)，(122)はxの分布関数$F(x)$がN年間のどの年にも共通であること，すなわち長期気候変動その他による経年変化がないことを前提にしている。

(2) FT-I 型分布のN年最大値

この分布関数は式 (121) に式(4)を代入することにより，次のように求められる。ただし，演算の便宜上，式(20)の基準化変量を途中で使用する。

$$\begin{aligned}\Phi_N(y_N)&=\exp[-\exp(-y_N)]^N\\&=\exp[-N\exp(-y_N)]\\&=\exp\{-\exp[-(y_N-\ln N)]\}\end{aligned}$$

$$\therefore\quad \Phi_N(x_N)=\exp\{-\exp[-[x-(B+A\ln N)]/A]\} \qquad (123)$$

白石・上田[42]も述べているように，式 (123) の関数形は原式である式(4)と同一であり，位置母数が $A\ln N$ だけ大きくなったに過ぎない。このようにN年中の最大値に対しても分布関数の形が変化しないというのが，FT-I 型分布の特徴である。本報告では扱っていないが，この性質は Fisher-Tippett のII型，III型分布でも成立する。ワイブル分布は Fisher-Tippett のIII型分布の特殊な場合であって，この性質は受け継いでいない。

関数形が同一であることを利用すると，x_N の平均値その他は式(6)，(7)，(17)などを使って次のように書き表すことができる。

$$\left.\begin{aligned}&\text{平　均}:\bar{x}_N=B+A(\ln N+\gamma)\\&\text{中央値}:(x_N)_{\mathrm{median}}=B+A[\ln N-\ln(\ln 2)]\\&\text{最頻値}:(x_N)_{\mathrm{mode}}=B+A\ln N\\&\text{標準偏差}:\sigma(x_N)=\sigma(x)=\pi A/\sqrt{6}\end{aligned}\right\} \qquad (124)$$

このN年最大値の平均値 $\bar{x}_N$ の再現期間は次のように計算される。

$$\begin{aligned}R(\bar{x}_N)&=1/[1-F(\bar{x}_N)]\\&=1/\left\{1-\exp\left[-\frac{1}{N}e^{-\gamma}\right]\right\}\\&\doteqdot1/\left\{1-\left[1-\frac{1}{N}e^{-\gamma}+\frac{1}{2}\left(\frac{1}{N}e^{-\gamma}\right)^2+\cdots\right]\right\}\\&\doteqdot Ne^{\gamma}\left[1+\frac{1}{2N}e^{-\gamma}+\cdots\right]\\&\doteqdot1.781N+\frac{1}{2}\end{aligned} \qquad (125)$$

すなわち，N年最大値の平均値の再現期間は約 1.8N年となる。平均値そのものとN年確率統計値との比率は尺度・位置母数等の値によって変り，一概にいうことはできない。

(3) ワイブル分布のN年最大値

この場合は FT-I 型分布のような式の変形による関数形の誘導ができない。このため白石・上田[42]は式 (121) の乗積の形のままで数値積分する方法と級数和展開による方法を示し，それによって各地の風圧力や波高等のN年最大値の平均値や標準偏差を計算している。

こうした値を直接に求めることはできないが，ワイブ

ル分布のN年最大値の概略の性質は次のような近似式によってほぼ明らかにすることができる。

$$\Phi_N(y_N)=[1-\exp(-y_N{}^k)]^N \fallingdotseq \exp[-N\exp(-y_N{}^k)] \quad (126)$$

この近似精度は y_N あるいはNの値が大きいほど向上する。まず，中央値については

$$(y_N)_{\mathrm{median}}=[\ln N-\ln(\ln 2)]^{1/k} \quad (127)$$

この中央値の再現期間は

$$\begin{aligned} R&=1/\exp[-(y_N)_{\mathrm{median}}{}^k] \\ &\fallingdotseq \exp[\ln N-\ln(\ln 2)]=N/\ln 2 \\ &=1.44N \end{aligned} \quad (128)$$

この式 (128) の関係は，実はNが十分大きければどのような分布関数についても成立する (Gumbel[1] 3.1.3節)。

最頻値 (mode) を求めるため確率密度関数を求めると，

$$\varphi_N(y_N)=Nky^{k-1}\exp[-y^k-\exp(-y^k)] \quad (129)$$

最頻値は $\varphi_N'=0$ を与える y_N の値であり，次式の数値解として求められる。

$$\left.\begin{aligned} kz[Ne^{-z}-1]+k-1=0 \\ z=(y_{\mathrm{mode}})^k \end{aligned}\right\} \quad (130)$$

式(130)を解くために

$$z-\ln N=\varepsilon<<1$$

と仮定すると，z の近似解が次のように導かれる。

$$z=\frac{1}{2}\left\{\ln N+\sqrt{(\ln N)^2+4\left(1-\frac{1}{k}\right)}\right\} \fallingdotseq \ln N\left\{1+\frac{1}{(\ln N)^2}\left(1-\frac{1}{k}\right)\right\}$$

したがって，$(y_N)_{\mathrm{mode}}$ の近似解として次の結果が得られる。

$$(y_N)_{\mathrm{mode}}\fallingdotseq\left\{\ln N\left[1+\frac{1-1/k}{(\ln N)^2}\right]\right\}^{1/k} \quad (131)$$

N年確率統計量は，基準化変量の形では式 (77) により $(\ln N)^{1/k}$ で与えられるので，式 (131) の結果はkの値によってN年最大値の最頻値の相対的位置が変化することを示している。すなわち，$k=1$ならば最頻値とN年確率統計量は等しく，これは FT-I 型分布でも同じである($k=1$ のN年最大値分布はFT-I 型分布へ漸近する)。$k<1$ のときは最頻値が N 年確率統計量よりも小さくなる。N年最大値の中央値は式 (127) により分布形に無関係 ($1/k$ 乗の項を除外して考える) であるから，このことは y_N の分布の裾が長引くことを意味し，したがって y_N の平均値の再現期間は FT-I型分布の $1.78N$ よりも大きくなる。$k>1$ のときはその逆であり，y_N の平均値の再現期間は $1.78N$ よりも短くなると推定される。

6.4 信頼性設計法における外力の変動係数について

信頼性設計法では上述のN年最大値の平均値を外力の基準値として使い，N年ごとの期間の最大値が期間によって異なることに対しては，N年最大値の標準偏差と平均値の比すなわち変異係数（信頼性設計では変動係数という）を導入して処理している。しかし，そこで用いられるN年最大値の標準偏差は対象とする事象の分布関数が確固不動の場合のものである。すなわち，使用する分布関数およびその母数の値が母集団の値を正しく代表していることを前提としている。本報告で繰り返し述べてきたように，そした前提が成立する可能性はほとんどない。対象とする事象の母集団の分布関数は，標本に対してあてはめられた分布関数の上下にある幅で取られた信頼区間の中のどこかにあるけれども，その絶対位置は不確定である。これは母集団の分布型が既知であっても同じである。こうした母集団の分布関数の不確定性を考慮するためには，N年最大値の変動係数の算定式を次のように修正する必要があるものと思われる。

$$\mathrm{C.V.}[x_N]=[\delta_N{}^2+\sigma_N{}^2]^{1/2}/\bar{x}_N \quad (132)$$

ここに，δ_N：最適合分布関数を確定関数と見なして算定されるN年最大値の標準偏差

σ_N：母集団の分布関数が不確定であることに起因するN年最大値の標準誤差

［計算例 5］

ある港で有効統計年数9.88年の観測によって低気圧性の高波 ($H_{1/3}\geq 4.0$) を50個取得した。データ中の最大値は 7.13m であり，平均値は $\bar{x}=4.688$m，標準偏差は $\sigma_x=0.723$m であった。平均発生率を $\lambda=50$ と想定して分布関数のあてはめを行い，最適合関数として $\hat{A}=0.7363$m, $\hat{B}=2.264$mの $k=1.0$ のワイブル分布を得た。この条件で50年最大値の変動係数を算定する。ここで得られた分布関数は $\lambda=50$ の極大値資料に対するものであるので，これを式(31)によって毎年最大値資料の分布関数に変換する。

$$\begin{aligned} F(x)&=\exp\{-50\exp[-(x-2.264)/0.7363]\} \\ &=\exp\{-\exp[-(x-5.144)/0.7363]\} \end{aligned}$$

これによって分布関数は FT-I 型で近似されたので，50年最大値の平均および標準偏差を求めると，

$$\bar{x}_{50}=5.144+0.7363\times(\ln 50+0.5772)=8.45\text{m}$$

$$\delta_{50}=\pi\times 0.7363/\sqrt{6}=0.94\text{m}$$

ただし，説明を簡単にするために $\bar{x}_{50}$ に対する偏り補正は無視している。

一方，σ_{50} を求めるため原関数における $\bar{x}_{50}$ の基準化変量 $\bar{y}_{50}$ を算定すると

$$\bar{y}_{50}=(8.45-2.264)/0.7363=8.401$$

σ_{50} は元の分布関数に対する係数 A_σ が式(111)によって 0.407 と求められるので，次のように計算される。ただし，ν は0.15に制限しておく。

$$\sigma_{50}=0.723\times\{1.0+0.407\times(8.401+1.0\times\ln 0.15)^{1.7}\}/\sqrt{50}=0.723\times 10.82/\sqrt{50}=1.11\mathrm{m}$$

以上により，50年最大値の変動係数は次のような値となる。

$$\mathrm{C.V.}[x_{50}]=(0.94^2+1.11^2)^{1/2}/8.45=0.172$$

母分布関数の不確定性を考慮しなければ変動係数は 0.111であるから，50%以上増加することになる。

信頼性設計法における自然外力の取り扱いは，外力のランダム性を考慮に入れている点で合理的であり，確率論的な考え方が定着するにつれて普及するものと思われる。本報告で検討した母分布関数の不確定性による誤差も取り込むことによって，外力のランダム性の取り扱いをさらに一段と実際の現象に近ずくことができるのではないかと考えられる。

7. むすび

本報告はもっぱら数値シミュレーション法によって多数の極値統計資料の標本を作成し，その統計解析に基づいて今まで未解明であった諸問題に対する解答を与えるとともに，実用的な観点から幾つかの新しい提案を行ったものである。主要事項を列挙すると次のようになる。

1) 極値統計資料は，期間最大値資料と極大値資料とに分類され，その各々が全数極値資料と部分極値資料に細分されるが，これらはすべて平均発生率およびデータ採択率の二つのパラメータによって解釈することが可能である。

2) 極値データに対して分布関数をあてはめる際には，各分布関数に対して最適なプロット公式を用いるべきであり，FT-I 型分布に対してはグリンゴルテン公式，正規分布と対数正規分布に対してはブロム公式，ワイブル分布に対しては修正 P&A 公式が推奨される。ワイブル公式を極値統計データに適用するのは不適切である。

3) 部分極値資料のプロッティング確率の計算においては，対象期間中の全データ数に基づく式(71)の方式を用いるのが適当である。

4) FT-I 型分布，対数正規分布，および4種類のワイブル分布について，標本資料の標準偏差の平均値ならびに変異係数が $N=10\sim100$，$\nu=0.25$，0.5，および 1.0 に対して数表として取りまとめられた。

5) 同じく最小2乗法による尺度母数および位置母数の信頼区間が，非超過確率 2.5～97.5%の分位値の数表として取りまとめられた。

6) 再現確率統計量の推定値の標準偏差に関する実験式が，上述の分布関数があらかじめ指定された場合について取りまとめられた。

7) 母分布関数が未知の場合，データ個数の少ない標本に対して分布関数のあてはめを行っても，母分布関数が最適合と判定される比率は低く，特にデータ採択率の小さい標本においては母分布関数を適切に判別することは困難である。

8) 母分布関数が未知な場合には，あてはめ対象の分布関数の種類を制限する必要がある。本報告では，FT-I 型分布，ならびに $k=0.75$，1.0，1.4，および 2.0 のワイブル分布の5種類を競合させる方式を提案する。

9) 母分布関数が未知であることにより，再現確率統計量の推定値は母集団の真値から偏った値となる可能性がある。この補正方法について実験式を提示した。

10) 母分布関数が未知な場合の再現確率統計量の標準誤差について，その推定のための実験式を提示した。この実験式を利用することにより，再現確率統計量の信頼区間や再現期間の変動範囲その他の推定が可能となる。

11) 異なる気象原因による極値データはそれぞれ別個に解析すべきであり，得られた結果を組み合せて全体としての再現確率統計量を推定する方法を明示した。

12) 極値統計資料に基づく再現確率統計量の推定値は，使用した標本に固有な統計的誤差を必然的に伴っているものであり，信頼性設計法において外力として極値データの解析結果を利用する際には，推定値の不確定性に基づく統計的誤差を導入する必要がある。

本報告で検討した事項の中では，次の3項目が解明困難あるいは今後の課題として残されている。

i) 極値統計資料の母分布関数は何が正しいのか。

ii) 母分布関数が未知であるとして，再現確率統計量の推定値の偏り補正および標準誤差の推定方法をどのように改良すべきか。

iii) 標本資料中の異常値の取り扱い。

第1点は対象とするそれぞれの極値事象ごとに検討すべき課題である。波浪についていえば，統計期間が10年以上のものの実測データについてデータ採択率を高め，データ個数を増やした解析を各地点について実施するこ

とによって何らかの展望が開けるかも知れない。

第２点については，まず本報告で採用した p_l=const.（式.97）の仮定の妥当性が問題点として挙げられる。しかし，これに代るべき資料のない現状では当面解決不能であろう。また，今回は ν=0.25 と 0.5 の場合しか検討していないので，さらに ν の範囲を広げ ν=1.0 の場合を統合するような実験式の樹立が望まれる。さらに，乗積確率の計算における偏り補正の方法についても改良の余地がある。

第３点については，本報告で提示した再現確率統計量の推定値の信頼区間の一つの応用として解決できると思われるが，本報告では割愛する。

まえがきでも述べたように，本報告の内容は完結したものではなく，現状における幾つかの実務的提案をまとめたものである。今後の研究の進展によって，さらに理論的・実務的に優れた方法が生み出されることを期待するものである。

なお，本報告の取りまとめにあたっては当所水工部小舟浩治海象観測研究室長からいろいろ貴重な御意見を頂き，表現の不備な点等を改めることができた。ここに記して謝意を表する次第である。

（1987年11月30日受付）

参　考　文　献

1) Gumbel, E. J.: Statistics of Extremes, Columbia Univ. Press, New York, 1958.（河田・岩井・加瀬監訳；「極値統計学」，生産技術センター新社）
2) Lawless, J. F.: Statistical Models and Methods for Lifetime Data, John Wiley & Sons, New York, 1982, 580 p.; a) p. 92, b) p. 464.
3) 江藤剛治他：大雨の頻度，土木学会論文集，第 369 号/Ⅱ-5, 1986, pp. 165-174.
4) 中西祐啓：治水計画における外力の確率評価の可能性について，土木計画学研究・論文集，No. 4, 1986, pp. 165-172.
5) Wang, S. and LeMehaute, B.: Duration of measurements and long-term wave statistics *J. Wat., Port, Coast. & Ocn. Eng.*, Vol. 109, No. 2, 1983, ASCE, pp. 236-249.
6) 合田良実：波浪統計に関する二，三の考察，港湾技研資料，No. 39, 1967, pp. 239-255.
7) Muir, L. R. and El-Shaarawi, A. H.: On the calculation of extreme wave heights: a review, *Ocean Engng.*, Vol. 13, No. 1, 1986, pp. 93-118.
8) 広瀬宗一・高橋智晴：観測結果に基づく沿岸波浪の出現特性，昭和 57 年度 港湾技術研究所講演会講演集，1982, pp. 1-55.
9) Lawson, N. V. and Abernethy, C. L.: Long term wave statistics off Botany Bay, *Proc. 2nd Australian Conf. on Coastal and Ocean Engng.*, 1975, pp. 167-176.
10) Resio, D. T.: Some aspects of extreme wave prediction related to climatic variations, *Prepr. 10th Ann. Offshore Tech. Conf.*, 1978, OTC 3278.
11) 神田　徹・藤田睦博：「水文学一確率論的手法とその応用一」，技報堂出版，1982, 275p.; a) p. 8, b) p. 34, c) p. 16, d) p. 51, e) pp. 33-42, f) p. 20.
12) Isaacson, M. de St. Q. and MacKenzie, N. G.: Long-term distributions of ocean waves, *J. Wat.. Port, Coast. & Ocn. Div., Proc.* ASCE, Vol. 107, No. WW 2, 1981, pp. 93-109.
13) 森口繁一他：「数学公式Ⅰ」，岩波書店，1956, p.231.
14) Weibull, W.: A statistical theory of strength of materials, Ing. Vet. Ak. Handl., Stockholm, 1939, 151p.
15) Petruaskas, C. and Aagaard, P. M.: Extrapolation of historical storm data for estimating design wave heights, *Prepr. 2nd Ann. Offshore Tech. Conf.*, 1970, OTC 1190.
16) 合田良実：波浪の統計的性質とその応用，1975年度水工学に関する夏期研修会講義集，1975, pp. B-6-1〜18.
17) 江藤剛治・室田　明：一雨降雨の１確率模型，土木学会論文集，第345号/Ⅱ-1, 1984, pp. 101-109.
18) Lettenmaier, D. P. and Burges, S. J.: Gumble's extreme value I distribution: a new look, *J. Hydr. Div.*, Proc. ASCE, Vol. 108, No. HY4, 1982, pp. 502-514.
19) Carter, D. J. T. and Challenor, P. G.: Methods of fitting the Fisher-Tippet type I extreme value distribution, *Ocean Engng.*, Vol. 10, No. 3, 1983, pp. 191-199.
20) Cunnane, C.: Unbiased plotting positions-a review, *J. Hydrology*, Vol. 37, 1978, pp. 205-222.
21) Lawless, J. F.: Approximation to confidence intervals for parameters in the extreme value and Weibull distributions, *Biometrika*, Vol. 61, No. 1, 1974, pp. 123-129.
22) Challenor, P. G.: Confidence limits for extreme value statistics, *Inst. Oceanogr. Sciences*, Rept. No. 82, 1979, 27p.
23) Borgman, L. E.: Extreme statistics in ocean engineering, *Proc. Civil Engng. in the Ocean*/Ⅲ, 1975, pp. 117-133.
24) Langbein, W. B.: Annual floods and the partial-duration flood series, *Trans. Amer. Geophy. Union*, Vol. 30, No. 6, 1949, pp. 879-881.
25) Chow, Ven Te: Discussion of "Annual floods and the partial-duration flood series," by W. B. Langbein, *Trans. A. G. U.*, Vol. 31, No. 6, 1950, pp. 939-941.
26) Rao, C. R.（奥野忠一他訳）：「統計的推測とその応用」，東京図書，1977, 568p.
27) Kimball, F.: On the choice of plotting positions

on probability paper, *J. Amer. Statist. Assoc.*, Vol. 55, 1960, pp. 546-560.

28) Blom, G.: Statistical Estimates and Transformed Beta-Variables, John Wiley & Sons, New York, 1958, Chapt. 12.

29) Gringorten, I. I.: A plotting rule for extreme probability paper, *J. Geophys. Res.*, Vol. 68, No. 3, 1963, pp. 813-814.

30) Barnett, V.: Probability plotting methods and order statistics, *Applied Statistics*, Vol. 24, No. 1, 1975, pp. 95-108.

31) Earle, M. D. and Baer, L.: Effects of uncertainties on extreme wave heights, *J. Wat., Port, Coast. & Ocn. Div.*, Proc. ASCE, Vol. 108, No. WW4, 1982, pp. 456-478.

32) Benson, M. A.: Characteristics of frequency curves based on a theoretical 1000 years record, 1952. (Reproduced in "Flood-frequency analysis," U. S. Geol. Survey Water Supply Paper 1543-A, 1960, pp. 51-77).

33) 角屋 睦：極値分布とその一解法，農業土木研究，第23巻第6号，1956.

34) Resio, D. T.: Extreme wave heights for Cleveland Harbor, *Proc. Civil Engng. in the Ocean*/III, 1975, pp. 62-78.

35) Simiu, E. and Filliben, J. J.: Probability distributions of extreme wind speeds, *J. Struct. Div.*, Proc. ASCE, Vol. 102, No. ST9, 1976, pp. 1861-1877.

36) Simiu, E., Bietry, J., and Filliben, J. J.: Sampling errors in estimation on extreme winds, *J. Struct. Div.*, Proc. ASCE, Vol. 104, No. ST3, 1978, pp. 491-501.

37) 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.

38) 宝　馨・高棹琢馬・清水　章：確率水文量の変動性を基準とした確率分布モデルの評価手順，第42回土木学会年次学術講演会講演集，II-5, 1987, pp. 40-41.

39) Carter, D. J. T. and Challenor, P. G.: Estimating return values of environmental parameters, *Quart. J. R. Met. Soc.*, Vol. 107, 1981, pp. 259-266.

40) Borgman, L. E.: Risk criteria, *J. Wat. & Harb. Div.*, Proc. ASCE, Vol. 89, No. WW3, 1963, pp. 1-35.

41) たとえば運輸省港湾局監修：「港湾の施設の技術上の基準・同解説，I 超大型石油タンカー用施設，II 海上貯油基地施設」，日本港湾協会，1980, pp. 107-108.

42) 白石　悟・上田　茂：港湾構造物及び海洋構造物の安全性照査に関する検討―作用荷重の変動係数と荷重係数の算定，港湾技術研究所報告，第26巻第2号，1987, pp. 493-576.

主要記号表

A, $\hat{A}$	：分布関数の尺度母数およびその推定値
A_c	：再現確率統計量の推定値に対する偏り補正の係数（式. 99～105）
A_σ	：再現確率統計量の推定値の標準誤差の係数（式. 111～114）
B, $\hat{B}$	：分布関数の位置母数およびその推定値
c	：再現確率統計量の標準偏差の実験式中の係数（式. 90, 92）あるいは標本資料の標準偏差の変異係数の比例定数（表-6）
C. V. [ ]	：括弧内の変数の変異係数
E[ ]	：括弧内の変数の期待値
$f(x)$	：極値統計量 x の確率密度関数
$F(x)$	：極値統計量 x の分布関数
$F_*(x)$	：分布関数が極大値資料に基づくことを明示する場合の表記
$\hat{F}_i$	：昇順の順序統計量 $x_{(i)}$ に割り付けた非超過確率
$F_j(x)$	：j 番目の気象原因による極値統計量の分布関数
$\hat{F}_m$	：降順の順序統計量 $x_{(m)}$ に割り付けた非超過確率
i	：昇順の順位序数（$=1, 2, \cdots, N$）
j	：序数（$=1, 2, \cdots, n$）
k	：分布関数の形状母数
K	：極値資料の対象期間の年数
$L(\theta)$	：尤度関数（式. 21）
m	：降順の順位序数（$=1, 2, \cdots, N$）
$M_{l,n}$	：l 番目の母分布関数からの標本のうちで n 番目の分布関数が最適合と判定された数
n	：序数，あるいは標本数
N	：標本中のデータ個数
N_j	：j 番目の発生原因別標本資料の個数
N_T	：対象期間中に発生した極値の総数
p_l	：l 番目の分布関数の存在確率
$P_{(i)}$	：順序統計量 $x_{(i)}$ の非超過確率
P_n	：事象 x が n 年目に生起する確率（式. 23）
P_r	：ポアソン分布の確率（式. 28）
q	：再現確率統計量の推定値の偏り補正あるいはその標準誤差の実験式の指数（式. 99 および 109）
$Q_{l,n}$	：l 番目の母分布関数からの標本のうちで n 番目の分布関数が最適合と判定される割合
r	：序数（1, 2, …）

R, $\hat{R}$ ：再現期間およびその推定値
$S_{l,n}^2$ ：l番目の母分布関数の標本がn番目の分布関数に最適合と判定された場合の再現確率統計量の推定値の真値からの2乗平均残差（式.107）
$\bar{S}_n$ ：n番目の分布関数に対する再現確率統計量の標準誤差の期待値（式.108）
Var[ ]：括弧内の変数の分散
x ：極値統計量
$x_{(i)}, x_{(m)}$：極値の順序統計量
x_R, $\hat{x}_R$ ：R年確率統計量およびその推定値
$\bar{x}$ ：標本データの平均値
y ：極値統計量の規準化変量（式.20）
$y_{(i)}, y_{(m)}$：規準化変量の順序統計量
y_R ：R年確率統計量の規準化変量
z ：R年確率統計量に対する補助統計量（式.81）
z_p ：補助統計量zのp%分位値
$\bar{Z}_{l,n}$ ：l番目の母分布関数の標本がn番目の分布関数に最適合と判定された場合の再現確率統計量の偏り量の期待値
$\bar{Z}_n$ ：n番目の分布関数に対する再現確率統計量の推定値の偏り量の期待値

α ：プロッティング公式中の定数，あるいは再現確率統計量の推定値に関する実験式中の係数（式.99および109）
α' ：プロッティング公式中の定数（式.54）
β ：プロッティング公式中の定数
γ ：オイラーの定数（0.5772…）
ζ ：[0，1] の一様乱数
$\boldsymbol{\theta}$ ：分布関数の母数ベクトル
κ ：分布関数が既知の場合の再現確率統計量の標準偏差の実験式中の係数（式.91）
λ ：極値の平均発生率（式.1）
λ_j ：j番目の発生原因別極値の平均発生率
μ ：母集団における確率変量の期待値
ν ：データ採択率（式.2）
σ ：母数団における確率変量の標準偏差
σ[ ] ：括弧内の変数の標準偏差
σ_x ：標本資料の標準偏差（式.72）
σ_z ：補助統計量の標準偏差
$\varphi_{(i)}$ ：i番目の順序統計量の確率密度関数(式.39)
φ_N ：N年最大値の確率密度関数（式.122）
Φ_N ：N年最大値の分布関数（式.121）

付録A：波浪データの極値統計解析の手順について

1. 作業手順

本報告では極値統計の諸問題について種々の検討を行ったが，実際に高波の極値資料を解析するための手順は必らずしも明示されていない。そこで本付録ではこの手順をまとめて示し，実務計算の参考に供する次第である。

計算作業を項目順に挙げると次のようになる。

1) 波浪観測・推算資料を解析し，ある限界値以上の高波の極大波高を抽出する。

2) 前項のデータから高波の毎年最大値資料または極大値資料を作成する。この際，極値資料は台風，低気圧，冬型気圧配置等の気象原因別に別個の資料として取りまとめることを原則とする。極大値資料の場合は，対象とする統計期間中に発生したと考えられる高波の概数 N_T を気象原因別に推定し，これに基づいて平均発生率 λ およびデータ採択率 ν を算定する。

3) 極値資料を大きさの順に並べ替え，最大のものを $m=1$ とする。

4) あてはめ分布関数として FT-I 型分布（式.4）およびワイブル分布（式.13，$k=0.75$，1.0，1.4，および 2.0）を対象とし，それぞれグリンゴルデン公式（式.50）および修正 P&A 公式（式.53 と 式.69 の組み合せ）を用いて出現確率を算定する。

5) 部分極値資料の場合には式(71)を用いて出現確率を算定する。参考のために前項と併せて計算式を再録すると次のとおりである。

FT-I 型分布：

$$\hat{F}_m=1-\frac{m-0.44}{N_T+0.12} \quad : m=1, 2, \cdots, N \qquad (付 1)$$

ワイブル分布：

$$\hat{F}_m=1-\frac{m-(0.20+0.27/\sqrt{k})}{N_T+0.20+0.23/\sqrt{k}} \quad : m=1, 2, \cdots, N \qquad (付 2)$$

6) 基準化変量 $y_{(m)}$ を次式で計算する。

FT-I 型分布：

$$y_{(m)}=-\ln[-\ln \hat{F}_m] \qquad (付 3)$$

ワイブル分布：

$$y_{(m)}=[-\ln(1-\hat{F}_m)]^{1/k} \qquad (付 4)$$

7) 極値波高データ $H_{(m)}$ と基準化変量 $y_{(m)}$ との間に次の関係を仮定し，最小2乗法を適用して母数推定値 $\hat{A}$ および $\hat{B}$ を求める。

$$H_{(m)}=\hat{A}\, y_{(m)}+B \qquad (付 5)$$

8) 最小2乗法の適用の際に $H_{(m)}$ と $y_{(m)}$ の相関係数も同時に求め，FT-I 型分布および4種類のワイブル分布をあてはめた中で相関係数が最大のものを最適合分布関数と判定する。

9) 所定の再現期間 R に対する確率波高の推定値 H_R を次式によって計算する。

$$H_R=\hat{A}\, y_R+\hat{B} \qquad (付 6)$$

ここに，

$$y_R=\begin{cases}-\ln[-\ln(1-1/\lambda R)] & : \text{FT-I 型分布}\\ [-\ln(1/\lambda R)]^{1/k} & : \text{ワイブル分布}\end{cases} \qquad (付 7)$$

10) 単独の気象原因ごとに，確率波高の推定値に対する偏り補正を式(106)によって行うとともに，その標準誤差を式(115)によって求める。このため，高波の極値データの標準偏差 σ_H をあらかじめ求めておく。

11) 複数の気象原因による高波が重合している場合には，気象原因別の極値データに対して最適合とされた分布関数をそれぞれ用い，式(118)によって全体としての高波の非超過確率を計算する。この際は，偏り補正を行わないケースと，**6.1**(3)の方法で偏り補正を行うケースとを併せて実施し，単独の気象原因別の確率波高と比較参照して，補り補正量の値を決定する。また，R年確率波高の推定値に対する標準誤差 σ_H を式(120)で求める。

12) 設計波高の選定に関しては，R年確率波高の推定値そのままではなく，これにその標準誤差のある倍数を乗じたものを加算あるいは削減した値を基準として参照する。標準誤差に乗じる倍数の大きさは，波浪資料の信頼度，設計対象の重要度その他を勘案し，適宜選定するものとする。

2. 計算例

ある港で台風による高波の観測値が，有効統計年数 $K=10.74$ 年にわたって取得された。この期間にこの港に来襲したと考えられる台風による波浪は，天気図等から判断して $N_T=53$ 個と考えられる。このうち，有義波高の極大値が 4.0m を超えたのは 21 個であった。したがって，データ採択率は $\nu=21/53=0.396$ であり，平均発生率は $\lambda=53/10.74=4.93$個／年である。

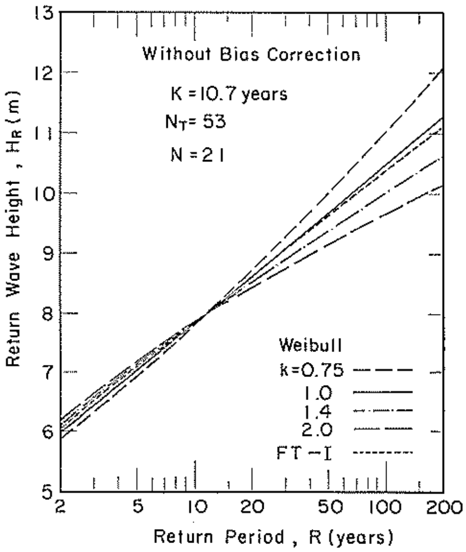
この21個の有義波高データを大きさ（降順）の順に並べて出現確率を算定し，基準化変量 $y_{(m)}$ を求めた結果を**付表-A.1** に示す。この表には最小2乗法によって求めた尺度母数および位置母数の推定値 $\hat{A}$ および $\hat{B}$，ならびに相関係数 r も記載してある。ただし，実際の計算では $y_{(m)}$ として有効桁数8桁のものを用いている。

付表-A.1 波高極値に対するプロッティング確率と基準化変量の計算例

順位 m	極値 $H_{(m)}$	FT-I 型		ワイブル(k=0.75)		ワイブル(k=1.0)		ワイブル(k=1.4)		ワイブル(k=2.0)	
		$\hat{F}_m$	$y_{(m)}$	$\hat{F}_m$	$y_{(m)}$	$\hat{F}_m$	$y_{(m)}$	$\hat{F}_m$	$y_{(m)}$	$\hat{F}_m$	$y_{(m)}$
1	8.36	0.9895	4.55	0.9909	7.86	0.9901	4.61	0.9893	2.95	0.9886	2.12
2	7.02	0.9706	3.51	0.9722	5.48	0.9714	3.55	0.9706	2.46	0.9699	1.87
3	6.94	0.9518	3.01	0.9535	4.46	0.9527	3.05	0.9518	2.21	0.9511	1.74
4	6.85	0.9330	2.67	0.9348	3.82	0.9339	2.72	0.9331	2.04	0.9324	1.64
5	6.74	0.9142	2.41	0.9161	3.35	0.9152	2.47	0.9144	1.90	0.9136	1.57
6	6.20	0.8953	2.20	0.8974	3.00	0.8965	2.27	0.8957	1.79	0.8945	1.50
7	5.92	0.8765	2.03	0.8787	2.71	0.8778	2.10	0.8769	1.70	0.8762	1.45
8	5.68	0.8577	1.87	0.8599	2.46	0.8591	1.96	0.8582	1.61	0.8574	1.40
9	5.57	0.8389	1.74	0.8412	2.26	0.8404	1.84	0.8395	1.54	0.8387	1.35
10	5.42	0.8200	1.62	0.8225	2.08	0.8216	1.72	0.8207	1.47	0.8199	1.31
11	5.34	0.8012	1.51	0.8038	1.92	0.8029	1.62	0.8020	1.41	0.8012	1.27
12	5.10	0.7824	1.41	0.7851	1.78	0.7842	1.53	0.7833	1.35	0.7825	1.24
13	5.09	0.7636	1.31	0.7664	1.65	0.7655	1.45	0.7646	1.30	0.7637	1.20
14	4.95	0.7447	1.22	0.7477	1.53	0.7468	1.37	0.7458	1.25	0.7450	1.17
15	4.81	0.7259	1.14	0.7290	1.43	0.7281	1.30	0.7271	1.21	0.7262	1.14
16	4.77	0.7071	1.06	0.7103	1.33	0.7093	1.24	0.7084	1.16	0.7075	1.11
17	4.63	0.6883	0.99	0.6916	1.24	0.6906	1.17	0.6896	1.12	0.6888	1.08
18	4.61	0.6694	0.91	0.6729	1.16	0.6719	1.11	0.6709	1.08	0.6700	1.05
19	4.41	0.6506	0.84	0.6542	1.08	0.6532	1.06	0.6522	1.04	0.6513	1.03
20	4.34	0.6318	0.78	0.6355	1.01	0.6345	1.01	0.6335	1.00	0.6325	1.00
21	4.11	0.6130	0.71	0.6168	0.95	0.6158	0.96	0.6147	0.97	0.6138	0.98
$\bar{H}=5.565$ $\sigma(H)=1.101$		$\hat{A}=1.091$ $\hat{B}=3.617$ $r=0.9842$		$\hat{A}=0.614$ $\hat{B}=4.029$ $r=0.9621$		$\hat{A}=1.147$ $B=3.374$ $r=0.9790$		$\hat{A}=2.084$ $\hat{B}=2.334$ $r=0.9878$		$\hat{A}=3.560$ $\hat{B}=0.786$ $r=0.9910$	

この計算の結果，分布関数としては $k=2.0$ のワイブル分布が最もデータに適合すると判定される。参考のために，ここで求めた母数推定値を使って5種類の分布関数による確率波高の推定値を比較した結果を **付図-A.1** に示す。再現期間が統計年数に等しい $R\fallingdotseq 11$年ではどの分布関数もほとんど同一の確率波高を与えるが，再現期間が長くなるにつれて差が次第に大きくなり，$k=0.75$ のワイブル分布が最大，$k=2.0$のワイブル分布が最小の確率波高の推定値を与える。

ここに示したのは偏り補正を行う前のものであり，補正を行った結果は**付表-A.2** に示すとおりである。補正量はそれぞれ括弧内に記入してあり，確率波高 H_R は補正後の値である。このように偏り補正を行うと，分布関数ごとの差異が減少し，大小の順位が逆転する結果となっている。**付表-A.2** には各推定値に対する標準誤差も示してある。

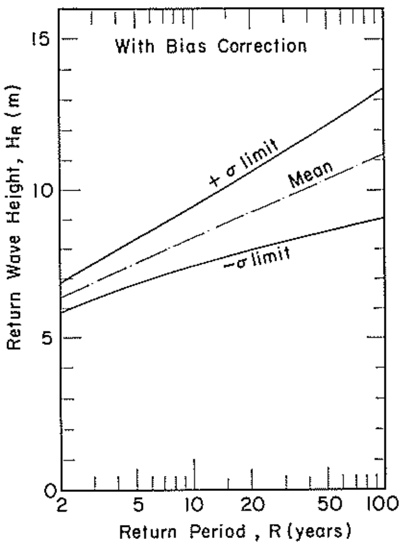


付図-A.1 各種分布関数における確率波高の推定結果

付表-A.2　確率波高とその標準誤差の計算結果

再現期間(年)	FT-I 型分布		ワイブル(k=0.75)		ワイブル(k=1.0)		ワイブル(k=1.4)		ワイブル(k=2.0)	
	H_R	$\sigma(H_R)$	H_R	$\sigma(H_R)$	H_R	$\sigma(H_R)$	H_R	$\sigma(H_R)$	H_R	$\sigma(H_R)$
2.0	6.16(−0.11)	0.44	5.87(+0.01)	0.28	6.01(−0.01)	0.39	6.17(−0.07)	0.46	6.38(−0.20)	0.51
5.0	7.27(−0.18)	0.69	6.83(+0.10)	0.50	7.09(−0.04)	0.61	7.29(−0.17)	0.69	7.58(−0.42)	0.77
10.0	8.09(−0.23)	0.93	7.58(+0.21)	0.70	7.92(−0.07)	0.82	8.11(−0.27)	0.90	8.44(−0.62)	1.02
20.0	8.90(−0.28)	1.20	8.34(+0.37)	0.94	8.75(−0.11)	1.07	8.92(−0.39)	1.16	9.28(−0.87)	1.30
50.0	9.98(−0.35)	1.60	9.36(+0.64)	1.30	9.87(−0.18)	1.45	9.97(−0.59)	1.54	10.38(−1.24)	1.73
100.0	10.79(−0.40)	1.94	10.13(+0.89)	1.60	10.73(−0.24)	1.78	10.77(−0.76)	1.87	11.22(−1.57)	2.10

注：括弧内は偏り補正量であり，確率波高は補正後の値である。



付図-A.2　確率波高の推定値の信頼区間の例

この計算例では k=2.0 のワイブル分布が最適合と判定されているので，確率波高の推定値の信頼区間として仮に 1 シグマ限界を用いた場合の確率波高の範囲を示したのが **付図-A.2** である。たとえば，50年確率波高の存在範囲は 8.7～12.1m と推定され，この範囲のどの値を設計波高として用いるかは設計者の判断に委ねられることになる。この判断の基準については今後さらに検討していかなければならない課題である。

付表-B.1　FT-I 型分布の母数推定値の信頼区間（グリンゴルテン公式による）

データ採択率	N	尺度母数 $A/\hat{A}$							位置母数 $(\hat{B}-B)/\hat{A}$						
		2.5%	5%	25%	75%	95%	97.5%	S.D.	2.5%	5%	25%	75%	95%	97.5%	S.D.
$\nu=1.0$	10	0.58	0.64	0.86	1.30	1.81	2.04	0.37	−0.76	−0.61	−0.21	0.25	0.66	0.83	0.40
	14	0.62	0.68	0.87	1.24	1.62	1.78	0.30	−0.61	−0.50	−0.19	0.22	0.54	0.67	0.32
	20	0.67	0.72	0.88	1.19	1.50	1.61	0.24	−0.48	−0.40	−0.15	0.18	0.45	0.54	0.26
	30	0.71	0.75	0.90	1.14	1.37	1.47	0.19	−0.39	−0.32	−0.12	0.14	0.35	0.43	0.20
	40	0.74	0.78	0.91	1.13	1.32	1.39	0.17	−0.33	−0.28	−0.11	0.12	0.30	0.36	0.18
	60	0.78	0.81	0.92	1.10	1.25	1.30	0.13	−0.27	−0.22	−0.09	0.10	0.24	0.29	0.14
	100	0.82	0.85	0.94	1.08	1.19	1.23	0.10	−0.21	−0.17	−0.07	0.07	0.18	0.22	0.11
$\nu=0.5$	10	0.52	0.58	0.83	1.38	2.04	2.35	0.48	−0.67	−0.58	−0.25	0.44	1.25	1.59	0.59
	14	0.57	0.63	0.84	1.30	1.82	2.03	0.38	−0.59	−0.52	−0.23	0.36	1.02	1.28	0.49
	20	0.61	0.66	0.86	1.23	1.63	1.78	0.30	−0.53	−0.46	−0.20	0.29	0.78	0.99	0.39
	30	0.66	0.71	0.88	1.19	1.48	1.61	0.24	−0.46	−0.40	−0.17	0.23	0.61	0.76	0.31
	40	0.70	0.74	0.89	1.16	1.40	1.49	0.20	−0.40	−0.35	−0.15	0.20	0.52	0.64	0.27
	60	0.74	0.78	0.90	1.12	1.32	1.40	0.17	−0.35	−0.30	−0.13	0.16	0.41	0.51	0.22
	100	0.79	0.82	0.92	1.10	1.24	1.30	0.13	−0.29	−0.25	−0.11	0.13	0.32	0.38	0.17
$\nu=0.25$	10	0.51	0.57	0.82	1.40	2.11	2.48	0.52	−0.95	−0.82	−0.35	0.71	2.02	2.64	0.93
	14	0.55	0.61	0.84	1.32	1.87	2.12	0.40	−0.87	−0.75	−0.31	0.61	1.64	2.06	0.76
	20	0.60	0.65	0.85	1.25	1.68	1.86	0.32	−0.79	−0.68	−0.28	0.47	1.25	1.57	0.61
	30	0.65	0.70	0.87	1.19	1.52	1.63	0.25	−0.69	−0.60	−0.26	0.37	0.97	1.17	0.48
	40	0.69	0.73	0.89	1.16	1.44	1.54	0.22	−0.62	−0.54	−0.23	0.32	0.83	1.03	0.42
	60	0.73	0.77	0.90	1.13	1.34	1.41	0.18	−0.53	−0.46	−0.19	0.26	0.65	0.79	0.34
	100	0.78	0.81	0.92	1.10	1.25	1.31	0.14	−0.44	−0.38	−0.16	0.20	0.50	0.60	0.27

付表-B.2 対数正規分布の母数推定値の信頼区間（ブロム公式による）

データ採択率	N	尺度母数 $A/\hat{A}$							位置母数 $(\hat{B}-B)/\hat{A}$						
		2.5%	5%	25%	75%	95%	97.5%	S.D.	2.5%	5%	25%	75%	95%	97.5%	S.D.
$\nu=1.0$	10	0.66	0.71	0.87	1.21	1.61	1.78	0.29	−0.69	−0.57	−0.22	0.21	0.57	0.70	0.35
	14	0.70	0.74	0.89	1.16	1.46	1.58	0.22	−0.57	−0.47	−0.19	0.18	0.47	0.58	0.29
	20	0.75	0.78	0.90	1.13	1.37	1.46	0.18	−0.45	−0.38	−0.15	0.16	0.40	0.48	0.23
	30	0.79	0.82	0.92	1.10	1.27	1.33	0.14	−0.36	−0.30	−0.12	0.13	0.31	0.37	0.19
	40	0.81	0.83	0.93	1.08	1.22	1.27	0.12	−0.31	−0.26	−0.11	0.11	0.27	0.32	0.16
	60	0.84	0.86	0.94	1.07	1.17	1.21	0.09	−0.26	−0.21	−0.09	0.09	0.22	0.26	0.13
	100	0.87	0.89	0.95	1.05	1.13	1.16	0.07	−0.20	−0.16	−0.07	0.07	0.16	0.19	0.10
$\nu=0.5$	10	0.62	0.66	0.85	1.26	1.77	1.98	0.36	−0.50	−0.43	−0.19	0.24	0.70	0.91	0.36
	14	0.66	0.70	0.86	1.20	1.59	1.76	0.28	−0.43	−0.36	−0.16	0.19	0.56	0.72	0.29
	20	0.70	0.74	0.88	1.16	1.45	1.57	0.22	−0.36	−0.31	−0.13	0.16	0.43	0.54	0.23
	30	0.75	0.78	0.90	1.12	1.34	1.44	0.17	−0.30	−0.26	−0.11	0.13	0.34	0.43	0.18
	40	0.77	0.80	0.91	1.10	1.28	1.35	0.15	−0.27	−0.23	−0.10	0.11	0.28	0.35	0.16
	60	0.80	0.83	0.93	1.08	1.23	1.28	0.12	−0.23	−0.19	−0.08	0.09	0.23	0.28	0.13
	100	0.85	0.87	0.94	1.06	1.17	1.20	0.09	−0.17	−0.15	−0.07	0.07	0.17	0.21	0.10
$\nu=0.25$	10	0.59	0.64	0.83	1.28	1.84	2.10	0.39	−0.53	−0.46	−0.22	0.31	0.97	1.26	0.47
	14	0.64	0.68	0.86	1.22	1.65	1.82	0.31	−0.47	−0.41	−0.18	0.28	0.78	1.00	0.38
	20	0.68	0.72	0.87	1.18	1.50	1.63	0.25	−0.42	−0.36	−0.16	0.22	0.59	0.74	0.30
	30	0.72	0.76	0.89	1.13	1.38	1.47	0.19	−0.36	−0.31	−0.14	0.16	0.46	0.56	0.23
	40	0.76	0.79	0.90	1.11	1.32	1.40	0.16	−0.32	−0.27	−0.13	0.14	0.39	0.47	0.20
	60	0.79	0.82	0.92	1.09	1.24	1.29	0.13	−0.27	−0.23	−0.10	0.12	0.30	0.37	0.16
	100	0.83	0.85	0.94	1.07	1.18	1.22	0.10	−0.22	−0.19	−0.08	0.08	0.22	0.27	0.12

付表-B.3 ワイブル分布（$k=0.75$）の母数推定値の信頼区間（新 P&A 公式による）

データ採択率	N	尺度母数 $A/\hat{A}$							位置母数 $(\hat{B}-B)/\hat{A}$						
		2.5%	5%	25%	75%	95%	97.5%	S.D.	2.5%	5%	25%	75%	95%	97.5%	S.D.
ν=1.0	10	0.42	0.50	0.82	1.71	3.00	3.58	0.86	−0.41	−0.35	−0.12	0.37	0.92	1.13	0.40
	14	0.47	0.53	0.83	1.56	2.47	2.87	0.65	−0.39	−0.34	−0.12	0.31	0.75	0.92	0.34
	20	0.50	0.57	0.84	1.44	2.16	2.48	0.51	−0.36	−0.31	−0.11	0.27	0.62	0.76	0.29
	30	0.55	0.62	0.85	1.33	1.87	2.09	0.40	−0.33	−0.28	−0.10	0.22	0.51	0.62	0.24
	40	0.59	0.65	0.87	1.28	1.72	1.89	0.33	−0.31	−0.26	−0.09	0.19	0.44	0.53	0.21
	60	0.64	0.68	0.88	1.22	1.54	1.67	0.27	−0.28	−0.24	−0.08	0.16	0.35	0.42	0.18
	100	0.70	0.74	0.90	1.17	1.41	1.50	0.20	−0.24	−0.21	−0.07	0.12	0.28	0.33	0.15
ν=0.5	10	0.44	0.52	0.83	1.63	2.70	3.22	0.74	−0.85	−0.74	−0.26	0.82	2.05	2.66	0.92
	14	0.48	0.56	0.83	1.48	2.30	2.65	0.57	−0.80	−0.68	−0.25	0.67	1.68	2.15	0.76
	20	0.52	0.59	0.85	1.38	1.99	2.25	0.45	−0.75	−0.64	−0.23	0.55	1.34	1.66	0.62
	30	0.58	0.64	0.86	1.30	1.76	1.94	0.35	−0.68	−0.57	−0.20	0.44	1.05	1.29	0.50
	40	0.62	0.67	0.87	1.25	1.63	1.76	0.30	−0.62	−0.52	−0.20	0.38	0.89	1.09	0.44
	60	0.66	0.71	0.89	1.20	1.49	1.61	0.24	−0.56	−0.48	−0.18	0.32	0.73	0.88	0.37
	100	0.72	0.77	0.90	1.15	1.37	1.45	0.19	−0.46	−0.39	−0.15	0.24	0.55	0.67	0.29
ν=0.25	10	0.46	0.53	0.83	1.57	2.57	3.08	0.70	−1.39	−1.18	−0.42	1.33	3.11	>3.2	1.52
	14	0.49	0.57	0.84	1.47	2.22	2.55	0.54	−1.29	−1.11	−0.39	1.13	2.63	>2.7	1.24
	20	0.55	0.61	0.85	1.37	1.94	2.20	0.43	−1.20	−1.01	−0.36	0.89	2.17	>2.3	1.01
	30	0.60	0.65	0.87	1.28	1.71	1.88	0.33	−1.08	−0.92	−0.34	0.69	1.70	1.80	0.80
	40	0.63	0.68	0.88	1.24	1.61	1.75	0.28	−0.98	−0.84	−0.31	0.62	1.46	1.56	0.71
	60	0.67	0.72	0.89	1.19	1.46	1.55	0.23	−0.86	−0.72	−0.27	0.49	1.13	1.27	0.57
	100	0.73	0.77	0.91	1.14	1.35	1.41	0.18	−0.73	−0.62	−0.23	0.37	0.86	0.98	0.45

付表-B.4 ワイブル分布（k=1.0）の母数推定値の信頼区間（新 P&A 公式による）

データ採択率	N	尺度母数 $A/\hat{A}$							位置母数 $(\hat{B}-B)/\hat{A}$						
		2.5%	5%	25%	75%	95%	97.5%	S.D.	2.5%	5%	25%	75%	95%	97.5%	S.D.
ν=1.0	10	0.50	0.57	0.83	1.44	2.22	2.58	0.55	−0.34	−0.30	−0.12	0.24	0.66	0.83	0.30
	14	0.55	0.61	0.84	1.37	1.96	2.24	0.44	−0.32	−0.27	−0.11	0.21	0.53	0.65	0.25
	20	0.60	0.66	0.86	1.29	1.75	1.95	0.35	−0.28	−0.24	−0.09	0.17	0.43	0.53	0.21
	30	0.65	0.70	0.88	1.22	1.55	1.68	0.26	−0.25	−0.21	−0.08	0.13	0.33	0.41	0.16
	40	0.69	0.73	0.89	1.18	1.47	1.58	0.23	−0.22	−0.19	−0.08	0.12	0.28	0.35	0.15
	60	0.72	0.76	0.90	1.15	1.36	1.44	0.18	−0.20	−0.17	−0.07	0.10	0.23	0.28	0.12
	100	0.78	0.81	0.92	1.11	1.28	1.33	0.14	−0.16	−0.14	−0.05	0.07	0.18	0.21	0.09
ν=0.5	10	0.51	0.57	0.84	1.46	2.24	2.62	0.56	−0.66	−0.58	−0.23	0.55	1.44	1.89	0.67
	14	0.55	0.62	0.84	1.35	1.96	2.21	0.43	−0.60	−0.52	−0.21	0.44	1.16	1.52	0.54
	20	0.59	0.65	0.86	1.27	1.73	1.92	0.34	−0.55	−0.48	−0.19	0.35	0.91	1.14	0.43
	30	0.65	0.70	0.88	1.22	1.55	1.70	0.27	−0.49	−0.42	−0.16	0.28	0.70	0.87	0.35
	40	0.68	0.72	0.89	1.18	1.46	1.56	0.23	−0.44	−0.37	−0.16	0.24	0.59	0.73	0.30
	60	0.72	0.77	0.90	1.14	1.37	1.45	0.19	−0.38	−0.32	−0.13	0.19	0.47	0.58	0.24
	100	0.77	0.81	0.92	1.11	1.27	1.33	0.14	−0.31	−0.27	−0.11	0.15	0.36	0.44	0.19
ν=0.25	10	0.51	0.57	0.83	1.44	2.22	2.64	0.56	−0.98	−0.84	−0.34	0.82	2.26	2.91	1.01
	14	0.55	0.61	0.84	1.36	1.95	2.22	0.43	−0.90	−0.78	−0.30	0.69	1.82	2.26	0.82
	20	0.60	0.65	0.86	1.28	1.74	1.95	0.35	−0.81	−0.70	−0.28	0.54	1.38	1.72	0.66
	30	0.65	0.69	0.87	1.21	1.56	1.69	0.26	−0.71	−0.62	−0.25	0.41	1.06	1.29	0.52
	40	0.69	0.73	0.89	1.18	1.48	1.59	0.23	−0.65	−0.56	−0.23	0.36	0.91	1.13	0.45
	60	0.73	0.77	0.90	1.15	1.36	1.45	0.18	−0.56	−0.47	−0.19	0.29	0.71	0.87	0.36
	100	0.77	0.81	0.92	1.11	1.27	1.33	0.14	−0.46	−0.39	−0.16	0.22	0.54	0.65	0.28

付表-B.5 ワイブル分布（$k=1.4$）の母数推定値の信頼区間（新 P&A 公式による）

データ採択率	N	尺度母数 $A/\hat{A}$							位置母数 $(\hat{B}-B)/A$						
		2.5%	5%	25%	75%	95%	97.5%	S.D.	2.5%	5%	25%	75%	95%	97.5%	S.D.
$\nu=1.0$	10	0.60	0.65	0.86	1.30	1.84	2.07	0.38	−0.30	−0.27	−0.12	0.19	0.54	0.70	0.26
	14	0.64	0.69	0.87	1.24	1.63	1.80	0.30	−0.27	−0.24	−0.10	0.15	0.42	0.52	0.20
	20	0.69	0.73	0.88	1.19	1.50	1.62	0.24	−0.24	−0.20	−0.09	0.13	0.34	0.42	0.17
	30	0.72	0.76	0.90	1.14	1.37	1.47	0.19	−0.20	−0.17	−0.07	0.10	0.25	0.31	0.13
	40	0.75	0.79	0.91	1.12	1.32	1.39	0.16	−0.18	−0.16	−0.07	0.08	0.22	0.26	0.11
	60	0.79	0.82	0.92	1.10	1.24	1.29	0.13	−0.15	−0.13	−0.06	0.07	0.17	0.21	0.09
	100	0.83	0.86	0.94	1.07	1.19	1.22	0.10	−0.12	−0.10	−0.05	0.05	0.13	0.16	0.07
$\nu=0.5$	10	0.57	0.62	0.84	1.35	1.96	2.25	0.44	−0.54	−0.48	−0.20	0.40	1.10	1.47	0.52
	14	0.60	0.66	0.85	1.26	1.75	1.95	0.35	−0.48	−0.42	−0.19	0.32	0.88	1.15	0.42
	20	0.65	0.69	0.87	1.21	1.57	1.71	0.27	−0.44	−0.38	−0.16	0.25	0.68	0.85	0.33
	30	0.70	0.74	0.89	1.17	1.43	1.55	0.21	−0.38	−0.33	−0.14	0.20	0.52	0.65	0.26
	40	0.73	0.77	0.90	1.14	1.36	1.44	0.18	−0.33	−0.29	−0.13	0.17	0.44	0.53	0.22
	60	0.77	0.80	0.91	1.11	1.29	1.36	0.15	−0.29	−0.25	−0.11	0.13	0.34	0.43	0.18
	100	0.81	0.84	0.93	1.08	1.21	1.26	0.11	−0.24	−0.20	−0.09	0.10	0.26	0.31	0.14
$\nu=0.25$	10	0.55	0.61	0.83	1.36	2.01	2.34	0.47	−0.74	−0.64	−0.27	0.55	1.59	2.07	0.73
	14	0.59	0.65	0.85	1.28	1.78	2.00	0.36	−0.67	−0.58	−0.24	0.46	1.26	1.58	0.58
	20	0.64	0.68	0.86	1.22	1.60	1.76	0.29	−0.60	−0.52	−0.22	0.36	0.96	1.19	0.46
	30	0.69	0.73	0.88	1.17	1.46	1.56	0.22	−0.51	−0.45	−0.20	0.27	0.74	0.89	0.36
	40	0.72	0.76	0.90	1.14	1.39	1.48	0.19	−0.46	−0.40	−0.17	0.24	0.62	0.77	0.31
	60	0.76	0.80	0.91	1.12	1.30	1.37	0.15	−0.40	−0.34	−0.14	0.19	0.48	0.59	0.25
	100	0.81	0.84	0.93	1.09	1.22	1.27	0.12	−0.32	−0.27	−0.12	0.14	0.36	0.44	0.19

付表-B.6 ワイブル分布（$k=2.0$）の母数推定値の信頼区間（新 P&A 公式による）

データ採択率	N	尺度母数 $A/\hat{A}$ 2.5%	5%	25%	75%	95%	97.5%	S.D.	位置母数 $(\hat{B}-B)/\hat{A}$ 2.5%	5%	25%	75%	95%	97.5%	S.D.
$\nu=1.0$	10	0.65	0.70	0.87	1.22	1.62	1.81	0.29	−0.30	−0.26	−0.12	0.17	0.51	0.66	0.24
	14	0.69	0.74	0.88	1.17	1.48	1.60	0.23	−0.26	−0.23	−0.10	0.13	0.39	0.49	0.19
	20	0.74	0.77	0.90	1.14	1.37	1.46	0.19	−0.22	−0.20	−0.09	0.11	0.31	0.39	0.15
	30	0.78	0.81	0.92	1.11	1.28	1.36	0.15	−0.19	−0.17	−0.07	0.08	0.24	0.29	0.12
	40	0.80	0.83	0.92	1.09	1.24	1.29	0.13	−0.17	−0.15	−0.06	0.07	0.20	0.24	0.10
	60	0.83	0.85	0.94	1.07	1.18	1.23	0.10	−0.14	−0.12	−0.05	0.06	0.15	0.19	0.08
	100	0.87	0.89	0.95	1.05	1.14	1.17	0.08	−0.11	−0.10	−0.04	0.04	0.11	0.14	0.06
$\nu=0.5$	10	0.60	0.65	0.84	1.29	1.83	2.06	0.39	−0.48	−0.42	−0.19	0.32	0.93	1.22	0.44
	14	0.64	0.69	0.86	1.22	1.64	1.82	0.30	−0.42	−0.36	−0.17	0.26	0.73	0.96	0.35
	20	0.68	0.72	0.88	1.17	1.48	1.61	0.24	−0.37	−0.33	−0.14	0.20	0.56	0.71	0.28
	30	0.73	0.77	0.89	1.14	1.37	1.47	0.19	−0.32	−0.28	−0.12	0.16	0.43	0.54	0.22
	40	0.76	0.79	0.91	1.11	1.30	1.37	0.16	−0.28	−0.24	−0.11	0.13	0.36	0.44	0.19
	60	0.79	0.82	0.92	1.09	1.24	1.30	0.13	−0.25	−0.21	−0.09	0.11	0.28	0.35	0.15
	100	0.84	0.86	0.94	1.07	1.18	1.22	0.10	−0.19	−0.16	−0.07	0.08	0.21	0.25	0.11
$\nu=0.25$	10	0.58	0.63	0.83	1.30	1.89	2.17	0.42	−0.60	−0.52	−0.24	0.41	1.23	1.62	0.57
	14	0.63	0.67	0.85	1.24	1.68	1.88	0.32	−0.53	−0.47	−0.21	0.34	0.97	1.22	0.45
	20	0.67	0.71	0.87	1.19	1.53	1.67	0.26	−0.48	−0.42	−0.18	0.26	0.74	0.92	0.36
	30	0.71	0.75	0.89	1.14	1.40	1.49	0.20	−0.40	−0.35	−0.16	0.20	0.56	0.68	0.28
	40	0.75	0.78	0.90	1.12	1.34	1.42	0.17	−0.37	−0.31	−0.14	0.18	0.48	0.59	0.24
	60	0.78	0.81	0.92	1.10	1.26	1.31	0.14	−0.31	−0.27	−0.12	0.14	0.36	0.43	0.19
	100	0.82	0.85	0.93	1.07	1.19	1.23	0.10	−0.25	−0.22	−0.09	0.10	0.26	0.33	0.15