

波高の極値統計解析における母数推定法の比較

愛媛大学工学部 正 員 ○山口正隆
 (附) ハザマ 正 員 中村雄二
 愛媛大学大学院 学生員 大木泰憲

1. はじめに 波高の極値統計解析に用いられる各種理論確率分布(Weibull分布、Gumbel分布等)の母数推定法には、多くの方法が存在する。一般的には、積率法、PWM法、最尤法および最小2乗法の4種類が使用されているようだが、その特性についての詳しい検討や比較は行われていない。そこで本研究では、母集団分布を指定した場合のモンテカルロシミュレーションを実施し、4種類の母数推定法の特性を明らかにする。
2. 母数推定法 4種類の母数推定法について簡単に説明する。①積率法: 確率分布の積率量である平均値、標準偏差、ひずみ係数などを、標本データから求められるこれらの推定量と等値であるとして母数を求める。②PWM法: 積率法と同様の方法であるが、積率法で使用する確率変量の原点まわりの積率のかわりに、確率密度関数の未超過確率および超過確率の累乗で重み付けした積率を使用する。③最尤法: 各標本データに対する確率密度関数 $f(x_i)$ の積で定義される尤度 L が最大となるように母数を求める。④最小2乗法: 線形と非線形最小2乗法があり、前者は、2母数と未超過確率 F の関数で表される基準化変量ととの間の線形関係を利用して、最小2乗法により母数を求める。後者は、母数が3つ以上の分布に対して、未超過確率に対する基準化変量と順序統計量に対する基準化変量との差の平均2乗誤差が最小となるように母数を求める。
3. モンテカルロシミュレーション 確率分布が明らかな母集団（母数およびすべての統計量が既知）から任意数の標本を抜き出して、4つの母数推定法により母数を推定して確率波高を求め、真値との比較を行う。シミュレーションの手順は次のようである。①母集団の確率分布と母数を指定する。②確率波高の真値を計算する。③一様乱数により母集団から大きさ N の標本を抽出する。④得られた標本に対して4種類の母数推定法によりそれぞれ母数、適合度指標（Q-Qプロットにおける相関係数 ρ 、SLSC）および確率波高を求める。⑤③、④を M 回($M=5000$)繰返し、適合度指標の平均値、確率波高の平均値と偏倚および2乗平均平方根誤差RMSEを求める。ただし、母数を推定することができない場合があるので、必ずしもデータの個数は M 個とはなっていない。⑥標本個数の影響を調べるために N を変えて($N=10, 20, 30, 40, 50, 70, 100, 200, 500, 1000$)③～⑤を実施する。
4. シミュレーション結果 ここでは例として、3母数Weibull分布に従う母集団に対するシミュレーション結果を提示する。表-1にこの結果を示し、併せて表をグラフにして図-1に示す。

表-1

母集団 標本個数		3母数Weibull分布 母数 $k=1.8, b=-1.0, X_c=5.0$									
		10	20	30	40	50	70	100	200	500	1000
適合度 相関係数	積率法	0.975	0.985	0.988	0.991	0.992	0.994	0.996	0.997	0.999	0.999
	PWM法	0.979	0.986	0.989	0.991	0.992	0.994	0.995	0.997	0.999	0.999
	最尤法	0.955	0.981	0.986	0.988	0.990	0.992	0.994	0.997	0.998	0.999
	最小2乗法	0.974	0.983	0.987	0.990	0.991	0.993	0.995	0.997	0.998	0.999
		0.048	0.042	0.038	0.035	0.034	0.031	0.028	0.024	0.018	0.015
適合度 SLSC	積率法	0.048	0.042	0.038	0.035	0.034	0.031	0.028	0.024	0.018	0.015
	PWM法	0.044	0.040	0.037	0.034	0.033	0.031	0.028	0.023	0.017	0.014
	最尤法	0.174	0.073	0.058	0.051	0.046	0.039	0.033	0.024	0.016	0.012
	最小2乗法	0.065	0.056	0.050	0.045	0.042	0.037	0.032	0.026	0.020	0.017
		-4.797	-3.755	-3.413	-2.980	-2.733	-2.468	-2.113	-1.562	-1.020	-0.707
100年確率波高 bias % (真値10.34m)	積率法	-4.797	-3.755	-3.413	-2.980	-2.733	-2.468	-2.113	-1.562	-1.020	-0.707
	PWM法	2.150	-0.664	-1.066	-1.085	-1.129	-1.153	-1.058	-0.930	-0.601	-0.472
	最尤法	27.108	2.413	1.902	1.730	1.443	0.940	0.697	0.342	0.222	0.125
	最小2乗法	6.414	3.633	2.613	2.183	1.780	1.248	0.981	0.335	-0.592	-1.383
		21.433	15.161	12.877	11.032	9.982	8.569	7.194	5.159	3.304	2.376
100年確率波高 RMSE %	積率法	21.433	15.161	12.877	11.032	9.982	8.569	7.194	5.159	3.304	2.376
	PWM法	25.269	16.266	13.551	11.446	10.290	8.691	7.184	5.076	3.208	2.302
	最尤法	46.045	18.264	14.398	12.368	11.045	9.091	7.339	5.008	3.086	2.201
	最小2乗法	25.889	16.736	13.706	11.743	10.505	9.078	7.866	6.112	4.511	3.581

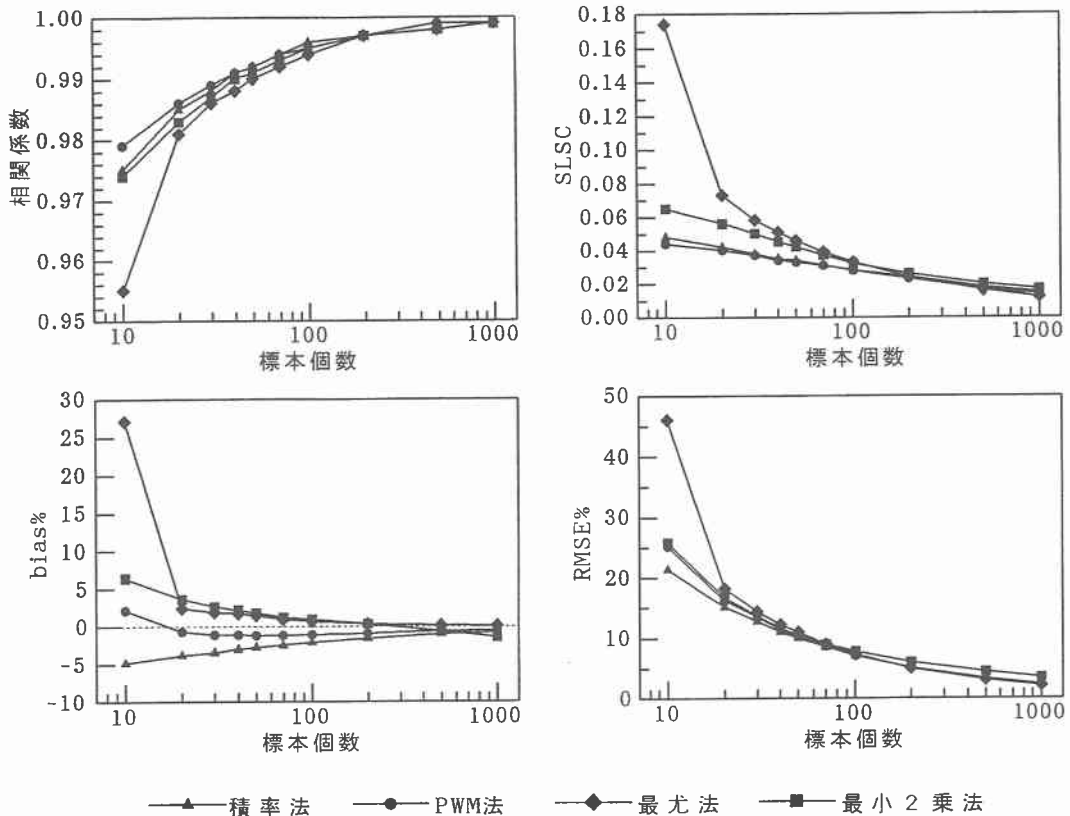


図-1

この表と図より次のことがわかる。①相関係数の基準によると、 $N \leq 30$ の場合PWM法、それ以上では積率法が最も優れている。標本個数に関係なく最尤法が最も劣っている。②SLSCの基準によると、 $N \leq 200$ の場合PWM法、それ以上では最尤法が最も優れている。適合度の絶対的な指標である $SLSC \leq 0.02$ となるのは、いずれの方法も $N > 200$ になってからである。③bias%はbiasを確率波高の真値で割ってパーセント表示したものである。biasの絶対値が最も小さいのは $N \leq 50$ の場合PWM法、それ以上の場合最尤法($N=200$ では最小2乗法)である。積率法およびPWM法($N=10$ を除く)は負のbiasを与える。最尤法および最小2乗法($N \leq 200$)は正のbiasを与える。④最尤法は標本個数が少ない(20以下)場合、解を与えにくく適合度も芳しくない。⑤PWM法は $N \leq 50$ の場合、最も優れた適合度を示し、しかもbiasが最も小さい。また、 $N \geq 50$ の場合でも適合度は他と比べて同じか優れているし、biasは最尤法に次いで小さい。⑥RMSE%が最も小さな値となるのは $N \leq 70$ で積率法、 $N=100$ でPWM法、 $N \geq 200$ で最尤法である。ただし、この3つの方法による値は $N=10$ の場合の最尤法を除いてかなり近い。 $N \geq 100$ で最小2乗法は他の3つに比べて大きな値となっている。

以上、シミュレーション結果を総合的にみると、3母数Weibull分布を極値統計解析の理論確率分布として用いる場合、その母数推定にはPWM法を使用するのが適当であると判断される。また、本研究の結果は、従来の研究において3母数Weibull分布の母数を積率法で推定している場合には、確率波高を過小評価し、最尤法や最小2乗法で推定している場合には過大評価している可能性を示唆する。

5. まとめ モンテカルロシミュレーションを行うことによって、母数推定法の特性を明らかにすれば、どの母数推定法を使用するのが最適であるかを客観的に判断することができるようになる。しかもbiasの特性を knowing ことで、従来の研究における確率波高がもつ偏倚を推測することも可能である。