土木学会論文集B Vol.66 No.1,66-75,2010.2

治水計画策定における統計的手法 —SLSC及び費用便益分析に関する考察—

葛葉　泰久[1]

[1]正会員　三重大学大学院教授　大学院生物資源学研究科（〒514-8507 三重県津市栗真町屋町1577）
E-mail: kuzuha@bio.mie-u.ac.jp

この論文では，治水計画策定におけるいくつかの問題のうち，宝・高棹のSLSC（標準最小二乗規準）と費用便益分析に関し，現在用いられている手法の問題点を指摘し，その解を示すことを目的としている．SLSCは非常によくできた客観的規準だが，それ自体，確率変数である．つまり，サンプル数に応じて異なるSLSCの閾値を用いるべきであり，確率分布の適合度の評価結果を報告する際には，使った閾値がSLSC分布のどの程度のパーセント点にあるかを示すのがよい．また，治水経済調査マニュアルの費用便益分析法に関し，最近のソフトウェアを用いた定積分を用いる方法で算定した被害額の期待値（真値に近い）と，マニュアルに従って求めた数値積分値がかなり異なる場合があることを示す．

Key Words : *standard least-square criterion, cost benefit analysis, flood control planning, Monte-Carlo simulation, probability density function*

1. 序論

本論文では，治水計画策定の種々のプロセスのうち，「標準最小二乗規準[1]（以下，慣例に従って“SLSC”と称す）を使用した適合度評価」と「費用便益比算定における便益計算」について考察を行う．いずれも，現在用いられている方法の問題点を指摘し，改善点を示すことを目的としている．本章では，それぞれの現在用いられている手法の概略と問題点について，簡単に解説する．

(1) SLSC法

SLSCは，宝と高棹（宝・高棹[1]）が1988年に発表した（ただし，それにさきがけ，まず，査読のない雑誌に1986年（高棹，宝ら[2]）に公表されている），（水文データへの）確率分布あてはめの適合度を評価する規準である．「30年分の年最大の1時間降水量データ」のような水文データがあったとして，この水文量の母集団がいかなる確率分布に従うか，それを評価する際に用いられる．非常に簡単に説明するならば，まず，何種類かの確率分布を候補に挙げ，水文データを，その何種類かの確率分布に対応した確率紙上にプロットする．水文データが，一番直線に近く並んだ確率紙，つまりそれに対応する確率分布が，高い評価を得る．SLSC法では，通常，SLSCという値を計算し，それが一番小さい統計分布を選ぶが，要は，目で見て，「どの確率紙上にデータをプロットした場合に直線に近いか」を判定する図式推定法の一種である．ただし，「目で見て」という，主観的な要素をできるだけ排し，SLSC値という，客観的に求められる規準（指標）を用いる．

さて，ここで何が問題かというと，それは以下の二つに集約される．ただし，この問題は，理論に内在する問題ではなく，正当な使い方をしない使用者の問題であろう．

a) SLSC < 0.04 （以降，この0.04などという値を“基準値”と称す）や，SLSC < 0.02 等のように，SLSCに，それがある基準値より小さかったら，適合度は充分であるというような記述がされている場合がある．この基準値として，0.04, 0.02などが多く用いられている．SLSCは，上述のように，「データの並び方がどの程度直線に近いか」を評価するものであるから，複数候補間の相対評価には適しているが，本来，「適合度が充分か否か」というような絶対評価には向かない．便宜上，後述するような理由により，絶対評価にも用いられているが，「なぜ0.04という値が用いられているか」という根拠は，常に認識すべきであろう．
b) そもそも，データの規模（データの個数）が異なれば，基準値は変わるはずである．後に示すが，水文データが30個の場合，SLSC < 0.04 を満たすのは困難でも，データが100個あれば，比較的容易である．

このような，データの個数を考慮していないSLSC法の使用例が多い．気象庁も確率降水量を求める際に，SLSCを用いているが，その際にSLSCを絶対評価的に使っており，サンプルの個数には注意を払っていないようである（参考：http://www.data.kishou.go.jp/climate/riskmap/cal_qt.html）．

(2) 便益計算法

公共事業の評価においては，その事業に要する費用（C: cost）と，事業による便益（B: benefit）を計算し，その比 B/C が1.0より大きいことが要求される．これは，特に治水事業だけに限ったことではない．地方公共団体などの事業主体は，治水事業に関して，治水経済調査マニュアル（案）（国土交通省河川局）[3]（以下，"マニュアル"と称す）を用いて便益の計算をすることが多いようである．

このマニュアルを用いた便益計算法には，いくつかの問題点や課題が指摘されている（劉[4]，湧川[5]，森・高木[6]など）．しかし，本稿では，便益を求めるに際し用いられる期待値計算法について，よりよい方法を提案することだけを目的としており，これらの問題点（またはそれに対する対策）については，特に言及しない．

治水工事の便益を算定する場合，たとえばリターンピリオド 2, 5, 10, 30, 50, 80, 100年に相当する降水が起こす洪水・氾濫による被害額から，被害額の期待値を求め（治水工事前と治水工事後のそれを求める）その差をもって便益（工事の効果）とする．ただし，本来，確率密度関数を用いた積分で期待値を算定すべきところを，非常に簡易的な手法を用いることが推奨されている．そこで，3.では，定義に戻って積分計算をすることを提案するものである．

2. SLSC法

(1) 問題点

年最大降水量や年最大流量などの水文極値データに何らかの確率分布をあてはめる際に，その適合度を評価する必要がある．その場合，少なくとも我国では，宝・高棹ら[1),2)]が提案した SLSC が，標準的に用いられている．もともと，SLSC は，異なる確率分布の適合度を比較する際に「相対評価を行なう」ために提案されたと思われるが，しばしば，「SLSC ≤ 0.04 なら充分な適合度と判断される」というように，様々な場面で，絶対的な基準値があるかのように考えられることがある．最近は，0.04 という基準値が用いられることが多いように思われるが，おそらく，その根拠は，田中・宝[7]と，中小河川計画の手引（改訂案）[8]（著者が入手した版は平成 18 年版であるが，正式には公開されていない）ではないかと思われる．

まず，高棹・宝[1)]は，例えば SLSC = 0.02 の解釈を，「プロット点の平分線からの偏差が，平均して確率紙の縦軸の長さの 2%程度に収まっていることを意味する」としたうえで，SLSC ≤ 0.02 なら「充分な適合性を表している」，SLSC > 0.03 なら，「他の分布形への当てはめを試みた方が良い」としている．ここで，主観を排することを目的として作成された SLSC ではあるが，「SLSC がいくらなら適合性があると言えるか」を決めるためには，主観による判断が必要なのは言うまでもなく，それ自体はおかしなことではない．ただし，確率紙を見て主観で 0.02 という基準値を決めたのは，「今までの解析結果との継続性」であろうと思われる．つまり，今まで充分適合性があると判断してきたケースは，SLSC でいうと 0.02 程度であろうと判断したということと思われる．

次に，田中・宝[7)]が行ったことを要約すれば，「日本の水文データをいろいろ調べた結果，0.02 という基準値は少々厳しすぎるので，0.04 に緩めるのが適当」と提案したということである．また，ここで重要なのは，「例えば基準値を 0.03 にした場合に厳しすぎる 」ことの根拠に，河川流量の計測方法に起因する問題も挙げられており，田中・宝が記述しているように，「河川流量の適合度の基準値を SLSC=0.04 程度に設定することが適当」ではあっても，他の水文量，統計量に関しては，その結論をすぐに適用できない可能性があることである．

以上のように，0.02（高棹・宝らが 1986 年に示した基準値）や，0.04（田中・宝が 1999 年に示した基準値）という値は，それぞれの著者らが非常に慎重にデータ解析を行なった結果得られたものではあるが，「データのばらつきが 2%, 4%程度である」という以上の意味は無く，例えば「SLSC ≤ 0.04 を満たしている場合には適合度として充分である」というのは，著者らの意図するところではないと思われる．

使用者側の認識を見てみると，例えば，国土交通省の「流量確率の選定及び計画降雨の時間，地域分布について（那賀川水系）（資料 1-2）」という資料（第 32 回河川整備基本方針検討小委員会で配布された資料）には，「SLSC：全国の実流域データから求められた SLSC 最小値の分布は，ほとんど SLSC ≤ 0.04 であるため，SLSC ≤ 0.04 のものを採用」と書かれている（この資料は，http://www.mlit.go.jp/river/shinngikai_blog/shaseishin/kasenbunkakai/shouiinkai/kihonhoushin/060207/pdf/s1-2.pdf から入手できる）．これから，使用者の側でも，「SLSC ≤ 0.04」というのを，便宜的な基準として用いていることが分かる．

また，例えば静岡県の「b)太田川治水対策に関する公開質問状 II への回答」という文書には，「確率分布モデルの適合度を客観的に評価する方法の一つとして，SLSCがあります．SLSCの値は標本数に依存することがわかっており，標本数が 40 程度の場合には，SLSC ≤ 0.04　を一つの規準として考えるのが妥当と考えます」と書かれており（この文書は，http://doboku.pref.shizuoka.jp/desaki/ootagawa1/news/kaitou/situmonkaitou2.pdf から入手できる），使用者側も，サンプル数によって SLSC の基準値を変えるべきなのは認識しているようである．

そこで，以降では，このあたりの課題と解決策を，SLSC の確率密度関数を求めることによって，明確にしようと思う．

(2) モデルによる解析

前節で述べたような背景があり，“SLSC の正体”を，明らかにしておく必要があると考えたので，ここで，2 種類のモデルを用いて解析を行う．

a) 解析の前提

ある確率分布を仮定して生成されたデータを使い，その仮定した確率分布に関する SLSC を求めれば，非常に SLSC が小さくなるはずである．そこで，ある確率分布に従うデータ群を多数作成し，それの SLSC を求める．こうして求めた SLSC の分布の中で，0.02 や 0.04 といった値が，どの程度の位置にあるのかを知れば，基準値の意味が明らかになると考える．つまり，SLSC 自体を確率変数ととらえ，その確率密度関数を求める．そして，0.04 などの基準値の意味，つまり，「どの程度，条件として厳しいものか」等を探る．

b) モデル

主に Monte-Carlo シミュレーションを用いて解析を行う．このためのモデルを，以下，“数値モデル”と称す．あわせて，補助的に，“準解析モデル”を併用する．準解析モデルは，数値モデルによる解析の validation を行うために用いる．このモデルが，“準”解析モデルであるのは，現段階では，数式の複雑さゆえに数値積分を使っており，完全な解析モデルではないからである．

c) 数値モデルの validation

SLSC の定義式は以下の式 (1), (2) のように表せる．

$$SLSC = \frac{\sqrt{\sum_i \varepsilon(i)^2 / n}}{|\, y_{0.99} - y_{0.01} \,|} \tag{1}$$

$$\varepsilon(i)^2 = (y(i) - y(i)^*)^2 \tag{2}$$

ここで，y は，reduced variate であり，正規分布の場合は，次式(3)で表せる．なお，式(1)中の $y_{0.99}$ のように，y にサフィックスが付いている場合，つまり，y_T のように表記する場合，T は非超過確率（1 からリターンピリオドの逆数を引いたもの）を意味している．なお，$\varepsilon(i), y(i), y(i)^*, x(i)$ のような表記は，それらが順序統計量を構成していることを明示的に示すものである．

$$y(i) = \frac{x(i) - \mu}{\sigma} \tag{3}$$

多くの確率分布に対して，$y(i)$ は，次式(4)のように表せる（ただし，実際は，x と y が線形の関係にあれば，式(3), (4)によらなくてもよい）．

$$y(i) = \frac{x(i) - \xi}{\alpha} \quad \text{or} \quad y(i) = k\,\frac{x(i) - \xi}{\alpha} \tag{4}$$

ここで，ξ, α, k は，それぞれ location parameter, scale parameter, shape parameter と呼ばれるパラメータである (Hosking and Wallis[9])．なお，式 (2) の $y(i)$，$y(i)^*$ は，それぞれ，データから式 (3), (4) などを用いて求めた reduced variate と，i から，適当な plotting position 公式を介して求められた $x(F)$（非超過確率 F から x を求める式，つまり，$F(x)$の逆関数）に対応する reduced variate である．

いま，母集団の確率分布関数が $N[0,1^2]$ である N 個の昇順順序統計量 $x(i)$ ($i = 1 \sim N$) があるとする．Stuart and Ord[10] を参考にすると，昇順順序統計量の i 番目 $x(i)$ の確率密度関数は，次式(5)のように表現できる．

$$f(x(i)) = \binom{N}{i-1}\binom{N-i+1}{1}\{\Phi(x(i))\}^{i-1}\phi(x(i)) \cdot \{1 - \Phi(x(i))\}^{N-i} \tag{5}$$

ここで，ϕ と Φ は，x（順序統計量とする前の確率変数）が従う確率分布の密度関数，確率分布関数（累積密度関数）であるが，ここでは，標準正規分布 $N[0,1^2]$ を仮定している．

ここで，式(1)中の $\sum_i \varepsilon(i)^2$ が従う分布を求めてみよう．まず，

$$\begin{aligned} Y(i) &= (x(i) - y(i)^*)^2 \\ &= (y(i) - y(i)^*)^2 \\ &= \varepsilon(i)^2 \end{aligned} \tag{6}$$

なる新たな確率変数を考える．ここで，$y(i)^*$ は，i と plotting position 公式で値が決まるので，確率変数ではないことと，確率分布として標準正規分布 $N[0,1^2]$ を想定しているので，$y(i) = x(i)$ であることに注意されたい．N 個の昇順順序統計量のうち，i 番目の $Y(i)$ が従うべき確率密度関数を $g_i(Y_i)$ とおくと，それは， 確率変数の変換（例えば亀田・池淵・春名[11] を参照されたい） を行えば，$x(i) = y(i)^* \pm \sqrt{Y(i)}$ を介し，$|dY(i)/dx(i)|$ を用いることにより，式(5)からすぐに求めることができる．例えば，N=3 ならば，$\sum_i \varepsilon(i)^2$ の期待値は，以下の式(7)

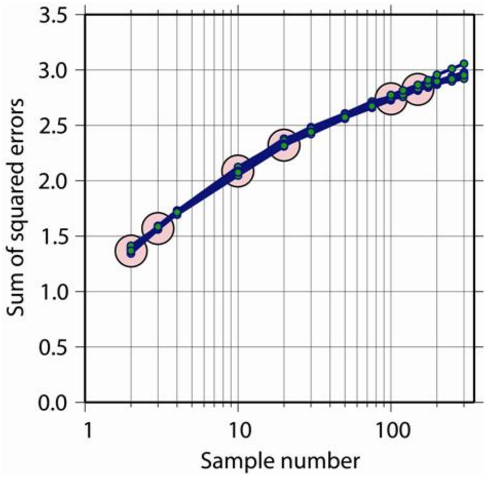

図-1 二乗誤差とサンプル数の関係．ピンクの円が準解析モデルによる解で緑の円がシミュレーション結果．緑の円は，ひとつに見える場合でも，複数（最大 10 個）の円が重なっていることに注意されたい．

で求めることができる．

$$E\left[\sum_i \varepsilon(i)^2\right] = \int_0^\infty \int_{Y_1}^\infty \int_{Z_2}^\infty Z_3 g_1(Y(1)) g_2(Z(2)-Y(1)) \cdot g_3(Z(3)-Z(2)) dZ(3) dZ(2) dY(1) \tag{7}$$

ここで，$Z(i)$ は $\sum_{j=1}^{i} Y(j)$ である． しかし，実際には，$Y(1), Y(2), Y(3)$ の確率密度関数が独立であることゆえ，一次のモーメントの加法性から，この期待値は次式(8)のように，簡単に求めることができる．

$$E\left[\sum_i \varepsilon(i)^2\right] = \sum_i E\left[\varepsilon(i)^2\right] = \sum_i \int_{-\infty}^{\infty} Y(i) g_i(Y(i)) dY(i) \tag{8}$$

図-1 は，このようにして求めた $E\left[\sum_i \varepsilon(i)^2\right]$ と N（サンプル数）の関係を示したものである．図中，6 つのピンクの円が，この準解析モデルによる計算値である．また，緑の小さな円が，数値モデルによる解である．図には，全部で 10 セットのシミュレーション結果をプロットしてある．この 10 セットのシミュレーション結果は下記のようにして得た．

(i) 10 セットのシミュレーション，それぞれについて，まず 10,000 個の水文量を発生させた．各セットの“seed”は互いに異なる．

(ii) それぞれのセットについて，ある N について，SLSC を 10,000 個求め，その平均と分散を求めた．

(iii) 10,000 個の SLSC の求め方は，次のとおりである．つまり，「10,000 個の水文量から N 個をランダムに取って，SLSC を求める」ということを 10,000 回繰り返した．

ここで強調したいのは， 10 セットのシミュレーションによる数値解（緑）がほぼ一致し，また，それらが準解析モデルによる解（ピンク）とほぼ一致することである．これをもって，数値モデルの妥当性が確認できたと考え，以下，このモデルを用いた解析を行う．

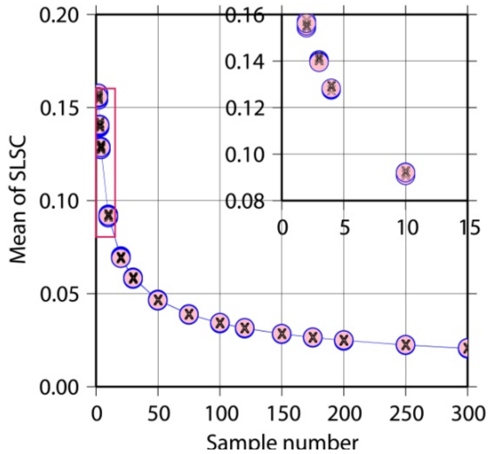

図-2 シミュレーションで求めた SLSC とサンプル数の関係．右上の図は，大きな図の中の赤い四角で囲った部分の拡大図である．

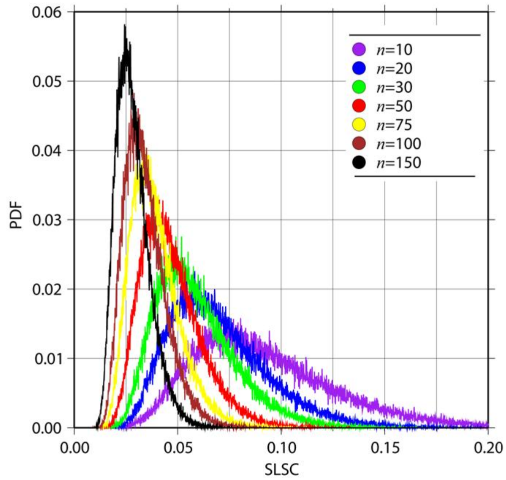

図-3 シミュレーションで求めた SLSC の確率密度関数．正規分布を仮定した．また，plotting position 式として，Cunnane の公式を用いた．この図と**表-1** は同じシミュレーション結果を用いている．

d) SLSC の平均値と分布

図-2 は，前項で説明した数値モデルで求めた SLSC とサンプル数 N の関係を示した図である．ここでも，**図-1** と同じく，10 個のセットを同時にプロットしている．ここで，水文データの母集団分布として，図中のピンクの○は $N[\mu,1]$ を，黒い×は $N[\mu,100^2]$ を仮定した．それらはほぼ重なっているので，SLSC の分布に（理論通り）分散は影響しないことを確認した（明らかに平均値も影響しない）．そこで，サンプル数 N ごとの SLSC の確率密度関数を示した**図-3** は，$N[\mu,\sigma^2]$ に等しく使える関数形（グラフ）となる（**表-1** は分位値をいくつか抜き出したものである）．例えば，サンプル数が 30 の場合，“SLSC = 0.04”は，16.2% のパーセント点に相当

表-1 シミュレーションで求めた SLSC．正規分布，Cunnane 式 [2)]を用いてシミュレーションを行った．

サンプル数	平均値	標準偏差	2.50%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
2	0.1561	0.0863	0.0289	0.0413	0.0587	0.0724	0.0839	0.0945	0.1039	0.1129	0.1216	0.1299	0.1394
3	0.1403	0.0680	0.0410	0.0514	0.0650	0.0748	0.0827	0.0903	0.0976	0.1051	0.1128	0.1203	0.1280
4	0.1282	0.0576	0.0445	0.0531	0.0640	0.0720	0.0792	0.0860	0.0923	0.0986	0.1050	0.1114	0.1183
10	0.0919	0.0347	0.0410	0.0461	0.0528	0.0579	0.0624	0.0663	0.0703	0.0742	0.0781	0.0820	0.0862
20	0.0695	0.0239	0.0340	0.0375	0.0423	0.0460	0.0491	0.0519	0.0547	0.0574	0.0601	0.0629	0.0657
30	0.0584	0.0192	0.0296	0.0326	0.0365	0.0394	0.0419	0.0443	0.0465	0.0487	0.0508	0.0530	0.0553
50	0.0466	0.0146	0.0246	0.0270	0.0299	0.0322	0.0342	0.0359	0.0377	0.0394	0.0410	0.0427	0.0444
75	0.0389	0.0118	0.0210	0.0229	0.0254	0.0273	0.0289	0.0303	0.0317	0.0331	0.0344	0.0358	0.0371
100	0.0342	0.0102	0.0187	0.0204	0.0225	0.0241	0.0255	0.0268	0.0280	0.0291	0.0303	0.0315	0.0327
120	0.0315	0.0092	0.0174	0.0189	0.0208	0.0223	0.0236	0.0247	0.0258	0.0269	0.0280	0.0290	0.0301
150	0.0284	0.0082	0.0158	0.0172	0.0190	0.0203	0.0214	0.0224	0.0234	0.0243	0.0253	0.0262	0.0272
175	0.0265	0.0076	0.0149	0.0162	0.0177	0.0190	0.0200	0.0209	0.0218	0.0227	0.0236	0.0245	0.0254
200	0.0249	0.0071	0.0141	0.0152	0.0168	0.0179	0.0189	0.0198	0.0206	0.0214	0.0222	0.0230	0.0238
250	0.0224	0.0063	0.0128	0.0138	0.0152	0.0162	0.0171	0.0179	0.0186	0.0193	0.0200	0.0208	0.0215
300	0.0206	0.0057	0.0119	0.0128	0.0140	0.0149	0.0157	0.0165	0.0171	0.0178	0.0184	0.0191	0.0198

サンプル数	55%	60%	65%	70%	75%	80%	85%	90%	95%	97.50%
2	0.1500	0.1615	0.1740	0.1882	0.2044	0.2231	0.2453	0.2754	0.3208	0.3634
3	0.1363	0.1451	0.1548	0.1659	0.1782	0.1928	0.2103	0.2338	0.2707	0.3046
4	0.1253	0.1327	0.1407	0.1496	0.1600	0.1724	0.1875	0.2075	0.2388	0.2663
10	0.0905	0.0950	0.0999	0.1053	0.1116	0.1188	0.1275	0.1392	0.1573	0.1743
20	0.0686	0.0717	0.0751	0.0788	0.0831	0.0880	0.0938	0.1018	0.1143	0.1260
30	0.0577	0.0602	0.0629	0.0659	0.0693	0.0733	0.0780	0.0844	0.0945	0.1037
50	0.0462	0.0482	0.0502	0.0524	0.0550	0.0579	0.0615	0.0663	0.0740	0.0810
75	0.0386	0.0402	0.0418	0.0436	0.0456	0.0480	0.0509	0.0548	0.0610	0.0669
100	0.0339	0.0353	0.0367	0.0382	0.0400	0.0420	0.0444	0.0478	0.0531	0.0581
120	0.0313	0.0324	0.0337	0.0351	0.0367	0.0385	0.0408	0.0439	0.0486	0.0531
150	0.0282	0.0293	0.0304	0.0317	0.0331	0.0347	0.0368	0.0395	0.0437	0.0476
175	0.0263	0.0273	0.0283	0.0295	0.0308	0.0323	0.0342	0.0366	0.0406	0.0442
200	0.0247	0.0257	0.0266	0.0277	0.0289	0.0303	0.0321	0.0343	0.0379	0.0413
250	0.0223	0.0231	0.0240	0.0249	0.0260	0.0272	0.0288	0.0308	0.0341	0.0372

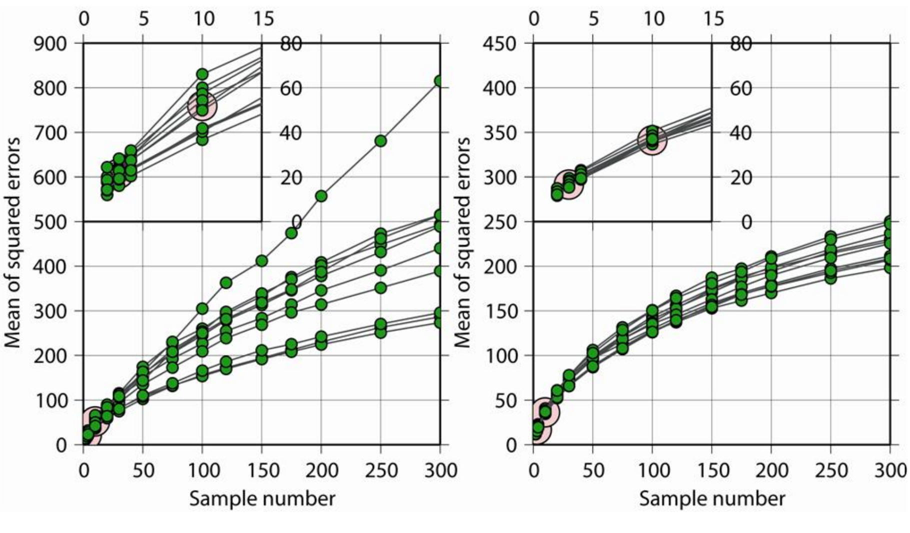

図-4 図-2 と同様の図．ただし GEV を用いた．“location, scale, shape parameters” は，それぞれ 72.84, 27.17, -0.35 とした．

する．しかし “SLSC＝0.02” は，サンプル数 100 個にして，ようやく，5%パーセント点となる程度の厳しい値である．

e) SLSC の分位値

著者は，以上の検討結果を踏まえ，以下のような SLSC の使用法を提案する．

(I)まず，どの程度のレベル（どの程度の分位値）の SLSC を基準値として用いるかを決める．**表-1** を使うならば，15%か 20%点に相当する SLSC が良い（これは，前述の「継続性」が理由である）．例えば 20%点を使うと，サンプル数 30 の場合は基準値が 0.042 で，サンプル数 30 の場合は基準値が 0.042 で，サンプル数が 100 個なら，0.026 が適当である．これを，サンプル数にかかわらず同じ値を用いるのは統計的におかしい．

(II)著者のこの**表-1** を用いて SLSC の基準値を算定しても良いが，もちろん，計算者自らが，自分の仮定する確率

分布を用いたシミュレーションを行ってもよい.

(III)計算者が 30 個のサンプルで，確率分布への当てはめを行い，SLSC＝0.04 を得たとする．その場合，計算者が,「SLSC = 0.04（サンプル数 30，正規分布，Cunnane 公式を仮定した場合の 16.2%の分位値）」と表示すれば，適合度がより客観的に評価できよう．

f) 裾の長い分布の SLSC

最後に，正規分布以外の確率分布を用いた場合について検討を行う．ここでは，計算者が，もし自分の用いた確率分布でシミュレーションを行い，その分布形の（確率密度関数の）裾が長い場合，シミュレーションの再現性が極めて悪いことを示す.

図-4 は，**図-1** と同様の図で，ただし，母集団の分布として GEV を仮定したものである．パラメータは，図中に示したものを用いている．左の図のケースは，何も操作することなくシミュレーションを行ったもので，右の図のケースは，確率密度関数の右の裾をカット（リターンピリオド 10,000 年以上をカット）してシミュレーションを行ったものである．PDF の裾の長い分布を用いてシミュレーションを行った場合，（この場合，リターンピリオド 10,000 年以上の）極値が現れるか否かによって結果が変わってくるので，シミュレーションの際には注意が必要である．そういう意味から，著者は，シミュレーションは（結果の再現性の良い）正規分布を仮定して行い，「SLSC=0.04（サンプル数 30，正規分布，Cunnane 公式を仮定した場合の 16.2%の分位値）」のように表示するのが良いと考える.

3. 便益算定法

(1) よく用いられている方法

マニュアルに記載されている便益計算法は，端的に言えば，本来，(9)を用いて計算すべき被害額総計の期待値 W を，離散的に数値積分する方法である.

$$W = \int f(x)h(x)dx \tag{9}$$

ここで，$x, f(x), h(x)$ は，それぞれ，降水量（または洪水流量など．ただし，以下，降水量と考える），その確率密度関数，降水量が x の時の被害額を表す関数（被害関数と称す）である．本論文の趣旨は，「$f(x)$ と $h(x)$ が分かった段階で，数値積分ではなく，解析的に定積分を実施して W を求めた方が正確で早い．最近は，計算機ソフトウェアでそれが容易にできる」ということである．

普通に行われている方法は以下のとおりである.

(I)いくつかのリターンピリオド T に対し，$g(T)$（上記の $h(x)$ と同じだが，ただしリターンピリオドを独立変数とする）を求める．つまり T に対する流出・氾濫計算を行い，被害額を求める．T として，例えば 2, 5, 10, 30, 50, 80, 100 年などをとる（これを以下，T_i; i = 1～n とする）.

(II)W を，次の式(10)で求める.

$$W = \sum_{i=1}^{n-1}\left[\{Q(T_i) - Q(T_{i+1})\} \times \{h(T_i) + h(T_{i+1})\}/2\right] \qquad Q(T_i) = 1/T_i \tag{10}$$

式(10)のΣ[　]の[　]の中の式に関し，乗算記号の前は $P\{x(T = T_i) \le x \le x(T = T_{i+1})\}$ を，乗算記号の後はその区間（$x(T = T_i) \le x \le x(T = T_{i+1})$）の被害関数の平均値の推定値を表している.

(III)実際は，流域をいくつかのサブ流域（ブロック）に分けて，各ブロックの被害額を合算し，流域全体の被害額を算出する.

(2) 解析的な手法（定積分による方法）

本節では，式(9)の W を（計算機ソフトウェアを用いて）定積分によって求めるが，以下，“実際はソフトウェアが数値積分を行っている場合”であっても，“ユーザーが被積分関数と積分範囲を指定したら解が示される”場合ならば，それを“定積分”と称す.

確率密度関数は，データをいくつかの確率分布に当てはめ，適合度評価を行ったものが存在すると仮定する．$f(x)$ は，連続型確率変数に対して定義されるものであり，本論文では，$f(x)$ は，任意の x に対して真値を与えるものと仮定する．つまり，$f(x)$ の不確実性については議論しない．それに対し，被害関数 $h(x)$ または $g(T)$ の方は，通常，いくつか選んだリターンピリオド T について，離散的に $g(T)$ が与えられる（被害額を計算する）にすぎないので，そもそもは連続関数ではない．しかし，以下では，被害関数 $h(x)$ を連続的に定義されるものとし，また，任意の x に対して真値を与えるものとする．つまり，何らかの補間を行って，適当な $h(x)$ がすでに求められているとする.

ここでは，実在の河川流域におけるデータを用いる（一級河川雲出川に関わるデータ）．降水量は Gumbel 分布に従い，確率分布関数は次式(11)で与えられるとする（**図-5** に確率密度関数を示す）.

$$F(x) = \exp\{-\exp(2.42 - 0.0196x)\} \tag{11}$$

また，$h(x)$ については，**図-6** 中の赤い丸が，氾濫計算等により算定された流域全体の被害額であるが，ここで，それらを結んだ，スプライン関数を用いた曲線（**図-6** 中の青い線）が，任意の正数 x に対して被害額の真値を与える被害関数 $h(x)$ であると仮定する.

著者は，数学ソフトウェア Mathematica 7.0 を用いて(9)

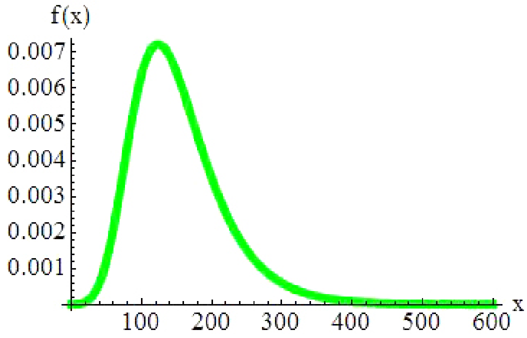

図-5 降水量の確率密度関数（x 軸は日降水量で単位は mm）

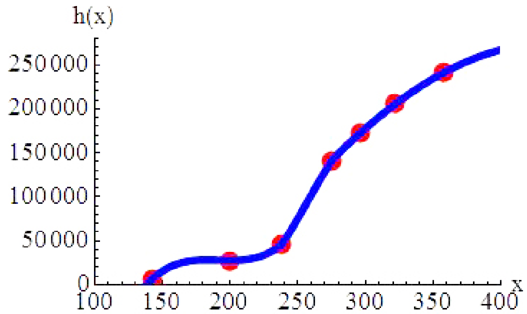

図-6 被害関数とその補間曲線

の定積分を行った．得られた期待被害額は $W = 24,331$（百万円）である．$x = 238$を境に2つのスプライン曲線を結合した場合（図は省略）は，W=23, 044（百万円）となる．これに対し，マニュアルによる方法では，W=22,109（百万円）が得られる．マニュアルによる方法については，**表-2** を参照されたい．表の上部に，わかりやすいように説明を書いた行を加えたが，その他は，マニュアルとほぼ同様の表記にしてある．また，(II)の「事業を実施した場合」の被害額をすべて 0 としたので，(III)の「年平均被害軽減額」は，そのまま，事業を実施しない場合の被害額の期待値を表している．

上述の場合は，被害額の期待値に大きな差異はないが，以下，計算方法によって結果が変わる場合の例を示そう．降水量の確率分布関数として，上記の場合と同じように，式(11)を用いる．被害関数として，次の 3 つの場合を仮定する．

Case A

$$h(x) = \begin{cases} 0 & x < 150 \\ 500(x-150) & x \geq 150 \end{cases} \tag{12}$$

Case B

$$h(x) = \begin{cases} 0 & x < 250 \\ 500(x-250) & x \geq 250 \end{cases} \tag{13}$$

Case C

$$h(x) = \begin{cases} 0 & x < 250 \\ 500(x-250) & 350 \geq x \geq 250 \\ 500(350-250) & x > 350 \end{cases} \tag{14}$$

結果として，Case A は，式(9)による W とマニュアルによる W がそれぞれ，13,147，12,987 であり，大きな差がなかったが，Case B は同様に，2,080 と 1,447，Case C は 1,783 と 1,427 と，いずれも大きな差が出た．

表-3 は，**表-2** と同様の表であるが，この Case A に関するものである．右端の 2 列は，ここで提案する定積分による方法で求めた $\int f(x)h(x)dx$ を表している．2 ヶ所の水色のセルが，マニュアル，積分による方法で求めた期待被害額を表しているが，これが大きく異ならないのは，対応する黄色のセルの値が大きく異ならないこと，また，マニュアルによる方法で無視される，非超過確率が非常に大きい範囲の $\int f(x)h(x)dx$（ピンクのセル）が，相対的に小さいためである．それと比して，Case B に関する **表-4** では，非超過確率の大きい範囲の $\int f(x)h(x)dx$ が相対的に大きい．

以上の原因を示したのが，**図-7** である．**図-7(1)**には，確率密度関数 $f(x)$ に加え，ここで用いた 8 つのリターンピリオド（1.1 年～100 年）に対応する降水量 x を y 軸に平行な直線で表示してある．つまり，図中の B～I が，**表-2, 3, 4** の右半分で規定する範囲に対応する（A は無視される）．また，**図-7(2), (3), (4)** は，それぞれ，Cases A, B, C の $\int f(x)h(x)dx$ を示している．

この図より，I の範囲（ここでのケースでは，リターンピリオド 100 年以上の範囲）を無視して数値積分を行うことにより，$\int f(x)dx$ については大きな誤差を生じないが，$\int f(x)h(x)dx$ については，大きな誤差を生じる様子が分かる．Case B はこれを示すために，「x の増加に従って増加し続ける $h(x)$ を用いたケース」であるが，上限を設けた場合も，Case C では，数値積分による誤差を生じている．

この問題は，計算者の，極値の確率分布の信頼性に対する考え方の問題である．「リターンピリオド 100 年以上の極端事象は実際には生じないので無視する．その部分の確率密度関数は信用できない」と考えるなら，マニュアルによる数値積分でも問題はないが，あくまで確率密度関数を厳格に用いるなら，本論文で提案する，定積分による方法が好ましい．

なお，本手法を用いる際に，以下の点に注意する必要がある．

(1) 被害関数は，降水量または洪水流量 x を独立変数とする関数 $h(x)$ として表現する必要がある．

表-2 マニュアルに記載されている年被害軽減期待額算出表

					被害軽減額の区間平均	超過確率の差を取って求めた確率	左の２つの欄の積	左の欄の累計
流量規模	年平均超過確率(N)	被害額(百万円) (I) 事業を実施しない場合	(II) 事業を実施した場合	(III) 被害軽減額 ((I)－(II))	区間平均被害額（百万円）	区間確率	年平均被害額（百万円）	年平均被害軽減期待額（百万円）
1.1年	0.909	0.0	0.0	0.0				
					3,127.5	0.409	1,279.4	1,279.4
2年	0.500	6,255.0	0.0	6,255.0				
					17,021.5	0.300	5,106.5	6,385.9
5年	0.200	27,788.0	0.0	27,788.0				
					36,906.5	0.100	3,690.7	10,076.5
10年	0.100	46,025.0	0.0	46,025.0				
					93,412.0	0.050	4,670.6	14,747.1
20年	0.050	140,799.0	0.0	140,799.0				
					156,518.5	0.017	2,608.6	17,355.8
30年	0.033	172,238.0	0.0	172,238.0				
					188,917.0	0.013	2,518.9	19,874.7
50年	0.020	205,596.0	0.0	205,596.0				
					223,476.5	0.010	2,234.8	22,109.4
100年	0.010	241,357.0	0.0	241,357.0				

表-3 Case A に関する年被害軽減期待額算出表

流量規模	年平均超過確率(N)	被害額(百万円) (I) 事業を実施しない場合	(II) 事業を実施した場合	(III) 被害軽減額 ((I)－(II))	区間平均被害額（百万円）	区間確率	年平均被害額（百万円）	年平均被害軽減期待額（百万円）	積分値	積分値の合計
1.1年	0.909	0.0	0.0	0.0						
					0.0	0.409	0.0	0.0	0.0	0.0
2年	0.500	0.0	0.0	0.0						
					12,466.5	0.300	3,740.0	3,740.0	2,779.6	2,779.6
5年	0.200	24,933.0	0.0	24,933.0						
					34,488.5	0.100	3,448.9	7,188.8	3,348.9	6,128.5
10年	0.100	44,044.0	0.0	44,044.0						
					53,209.8	0.050	2,660.5	9,849.3	2,610.0	8,738.5
20年	0.050	62,375.5	0.0	62,375.5						
					67,648.3	0.017	1,127.5	10,976.8	1,121.7	9,860.2
30年	0.033	72,921.0	0.0	72,921.0						
					79,512.3	0.013	1,060.2	12,036.9	1,052.8	10,913.0
50年	0.020	86,103.5	0.0	86,103.5						
					94,994.0	0.010	949.9	12,986.9	939.8	11,852.8
100年	0.010	103,884.5	0.0	103,884.5					1,294.2	13,147.0

表-4 Case B に関する年被害軽減期待額算出表

流量規模	年平均超過確率(N)	被害額(百万円) (I) 事業を実施しない場合	(II) 事業を実施した場合	(III) 被害軽減額 ((I)－(II))	区間平均被害額（百万円）	区間確率	年平均被害額（百万円）	年平均被害軽減期待額（百万円）	積分値	積分値の合計
1.1年	0.909	0.0	0.0	0.0						
					0.0	0.409	0.0	0.0	0.0	0.0
2年	0.500	0.0	0.0	0.0						
					0.0	0.300	0.0	0.0	0.0	0.0
5年	0.200	0.0	0.0	0.0						
					0.0	0.100	0.0	0.0	0.0	0.0
10年	0.100	0.0	0.0	0.0						
					6,187.7	0.050	309.4	309.4	171.7	171.7
20年	0.050	12375.5	0.0	12375.5						
					17,648.3	0.017	294.1	603.5	288.3	460.0
30年	0.033	22921.0	0.0	22921.0						
					29,512.3	0.013	393.5	997.0	386.1	846.1
50年	0.020	36103.5	0.0	36103.5						
					44,994.0	0.010	449.9	1,447.0	439.8	1,285.9
100年	0.010	53884.5	0.0	53884.5					794.2	2,080.1

(2) マニュアルによれば，流域をブロックごとに分割した場合の，流域全体の被害額は，ブロックごとに独立して計算し，合計すればよいので，ブロックごとに異なる $f(x)$ を用いることが可能である．

(3) ブロック内で，T が異なると破堤点が異なり，異なる降雨強度式を用いる必要がある場合がある．その場合は途中で $f(x)$ が変わるが，複数の $\int f(x)h(x)dx$ を求めて合成すればよい．

4. 結論

治水計画策定の際に用いられる確率・統計的手法のうち，宝・高棹の SLSC と，費用便益計算における，便益の計算手法について検討を行った．前者について，SLSC は非常に便利な客観的規準ではあるが，それ自体が，確率変数であることを示した．また，サンプル数に応じて異なる SLSC の閾値を用いるべきであり，確率分

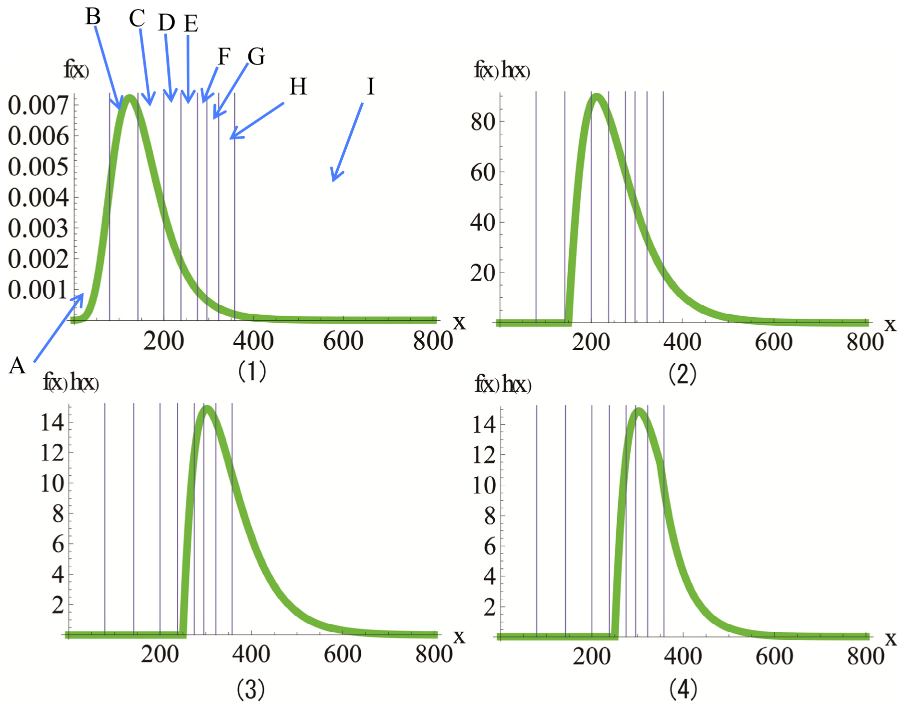

図-7 (1)降水量の確率密度関数 $f(x)$ (2),(3),(4)は，それぞれ Cases A,B,C の $f(x)h(x)$．図中の A はリターンピリオドが 1.1 年以下の区間，B は，1.1～2 年の区間．その他は，表-2, 3, 4 の流域規模の列を参照されたい．

布の適合度を評価結果を報告する際には，使った閾値がSLSC 分布のどの程度のパーセント点にあるかを示すことを提案した．

後者については，治水経済調査マニュアルに記述してある費用便益分析法に関し，最近のソフトウェアを用いた定積分を用いる方法で算定した被害額の期待値と，マニュアルで求めたものがかなり異なる場合があることを示した．

謝辞：国土交通省三重河川国道事務所と三重県県土整備部の関係部署の方々には，データをいただくとともに，種々ご教示いただいた．また，国土技術研究センターの湧川勝己氏には資料をお送りいただいた．ここに記して深謝いたします．なお，本論文は，三重大学大学院生物資源学研究科の「紀伊半島における自然災害のモニタリングと予測に関する研究」と三重大学自然災害対策室における風水害研究の成果をまとめたものである．

参考文献

1) 宝馨，高棹琢馬：水文頻度解析における確率分布モデルの評価規準，土木学会論文集，No.393/II-9, pp.151-160, 1988.
2) 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，京都大学防災研究所年報，Vl.29, B-2, pp.1-15, 1986.
3) 国土交通省河川局：治水経済調査マニュアル（案），107pp., 2005.
4) 劉瑀：治水経済調査の便益計算に関わる問題点, OGI Technical Reports, Vol.14, 応用技術, 2003.
5) 湧川勝己：治水経済調査における新たな洪水リスクの評価と費用便益分析, JICE REPORT, 第 2 号, pp.13-19, 2001.
6) 森寛典，高木朗義：堤防の破堤確率を考慮した洪水被害額の算定方法に関する基礎的考察，河川技術論文集, Vol.13, 2007.
7) 田中茂信，宝馨：河川流量の頻度解析における適合度と安定性の評価，水工学論文集, Vol.43, pp.127-132, 1999.
8) 国土技術研究センター：中小河川計画の手引（改訂案），技術資料編, 165pp., 2006.
9) Hosking, J. R. M. and Wallis, J. R.: *Regional Frequency Analysis*, 224pp., Cambridge, 1997.
10) Stuart, A. and Ord, K. : *Distribution Theory, Kendall's Advanced Theory of Statistics*, Volume 1 of Six Edition, 676pp., Hodder Arnold, 1994.
11) 亀田弘之，池淵周一，春名攻：確率・統計解析，新体系土木工学第 2 巻, 307pp., 技報堂出版, 1981.

(2009. 8. 11 受付)

CONSIDERATIONS OF STATISTICAL METHOD IN FLOOD-CONTROL PLANNING - SLSC AND COST BENEFIT ANALYSIS -

Yasuhisa KUZUHA

The author investigated an issue of the standard least-square criterion (SLSC) which has been used in Japan. It is most important that SLSC is in itself a random variable and it has a probability distribution, and the distribution depends on the number of sample. We have to notice it in flood-control planning and use different thresholds of SLSC (namely, e.g. SLSC=0.04), depending sample sizes. Thus, the author proposes that we indicate the threthold's cumulative probability like "the threshold correspons to 16% -percentile in the probability distribution". Moreover, the author investigated the way which has been used for estimating the expected value of damaged amount by floods, which is used in Japan.