

特殊な裾形状を有する確率分布の推定の精度の検証

兼清 泰明* 河原 宏紀** 田中 秀弦**
丸山 収*** 須藤 敦史**** 佐藤 京*****

Verification of Accuracy in Statistical Estimations for Probability Distributions Having Tails of Special Shape

by

Hiroaki KANEKIYO*, Koki KAWAHARA**, Hidenori TANAKA**,
Osamu MARUYAMA***, Atsushi SUTOH**** and Takashi SATOH*****

Statistical estimations for probability distributions having tails of special shape, such as a double-mode distribution as well as the so-called heavy-tailed or fat-tailed distribution, are quantitatively discussed through virtual experiments using computer simulations. In this paper, a probability distribution describing the damage degree of concrete liners of tunnels in cold region is examined as an example of the double-mode probability distribution.

First, virtual data sets of observations are generated by the use of quasi-random numbers for a Pareto distribution as a typical example of the fat-tailed distribution, whereas the actual data set obtained for the damage degree of concrete liners is used for generating virtual data set for the double-mode distribution. Next, statistical estimations are executed by the use of probability papers to identify the probability distribution showing the “best” fitting among supposed plural candidates for probability distributions. Finally, the accuracy of the estimation is quantified by applying the coefficient of determination. The results show that the accuracy of the estimation in the tail region is scarcely improved even if the number of data is increased.

Key words:

Probability distribution, Heavy-tailed distribution, Fat-tailed distribution, Double-mode probability distribution, Statistical estimation

1 緒 言

ばらつきが認められる材料物性値等の確率分布を推定する場合、確率分布の形状に関する何らかの予備的な知識がない限り、ばらつきを示すデータを確率紙上にプロットして、最もフィットがよい分布形を採用するという手法が用いられることが多い。この場合、対象となる確率分布の裾(テイル)部分の収束が非常に遅い分布では、このような推定における分布形状の特定が非常に困難となることが起こり得る。近年、リスク解析や通信ネットワークのトラフィック解析の分野において、ヘヴィーテイル^{1),2)}あるいはファットテイル^{3),4)}と呼ばれる特性を持つ確率分布に関心が高まってきているが、ぜい性の顕著な材料の強度などのばらつきを記述する確率分布にはそのような裾形状が認められることが少なくない。こういった特徴を持つ分布、いわゆる裾が重いと称される分布の推定では、ある程度の数のデータが得られていたと

しても、裾の重くない分布形と誤って推定してしまう可能性が考えられる。特に、こういった分布を限られた個数の実測データから推定する際、危険側の推定を避けるためにはどの程度のサンプル数が必要であるのか、という定量的観点からの研究事例はほとんど報告されていない。

また、複数分布特性が混在するような多重モード型の確率分布においても、分布の裾野の形状が裾の重い分布と似た形状となることがあり、著者らによる先行研究⁵⁾では、寒冷地トンネル覆工コンクリートの損傷度^{6),7)}のばらつきの特性が二重モード分布により裾部分を含めてうまく説明できることが明らかとされている。このような分布においては、データ数が十分でない場合には二種類含まれるモードのうちの一種類のみからなる確率分布と誤って推定してしまう危険性があり、結果として裾の重い分布における分布形の誤推定と同じような危険側の

† 原稿受理 令和4年8月9日 Received Aug. 9, 2022 ©2023 The Society of Materials Science, Japan

* 正会員 関西大学 環境都市工学部 都市システム工学科 〒564-8680 吹田市山手町

Department of Civil, Environment and Applied Systems Engineering, Kansai University, Yamate-cho, Suita, 564-8680.

** 関西大学大学院 理工学研究科 環境都市工学専攻 〒564-8680 吹田市山手町

Graduate School of Science and Engineering, Kansai University, Yamate-cho, Suita, 564-8680.

*** 東京都市大学 工学部 都市工学科 〒158-8557 東京都世田谷区玉堤

Department of Urban & Civil Engineering, Tokyo City University, Tamazutsumi, Setagaya-ku, Tokyo 158-8557.

**** 東北工業大学 工学部 都市マネジメント学科 〒982-8577 仙台市太白区八木山香澄町

Department of Civil Engineering & Management, Tohoku Institute of Technology, Taihaku-ku, Sendai, 982-8577.

***** (国独) 寒地土木研究所 〒062-8602 札幌市豊平区平岸

Civil Engineering Research Institute for Cold Region, Hiragaishi, Toyohira-ku, Sapporo, 062-8602.

評価を与えてしまう可能性がある。

そこで本研究では、裾の重い分布を含めて分布裾野が特殊な形状を有する確率分布形を推定するに際して、推定に用いるデータ数と推定精度への影響、また裾の危険側の評価に対する影響について、計算機実験を通じてその特徴を明らかにする。

2 本研究の直接的背景

2.1 裾の重い確率分布

材料・構造システムの劣化や不健全性を表す指標、あるいはシステムに負荷される荷重値など、値が大きくなるほど構造健全性の観点からのリスクが増大するような量を X とし、実数値を取る確率変数であるものとする。その確率分布関数を $F_X(x)$ 、さらに $F_X(x)$ が微分できるものとして、その微分、すなわち X の確率密度関数を $f_X(x)$ と表す。信頼性解析上は、 x の値が非常に大きな値を取る領域での $F_X(x)$ あるいは $f_X(x)$ の挙動が重要となる。この部分を右裾野、あるいは簡単に裾と呼ぶこととする。なお、材料の余寿命のように分布の左裾の挙動が重要となる場合についても本論文と同様のアプローチは適用し得ると考えられる。

確率密度関数 $f_X(x)$ の裾が、指数関数よりも緩やかにゼロに収束していくような分布は、ヘヴィータイルを有すると言われている。数学的には、どのような正の実数 a を取っても、

$$\lim_{x \rightarrow \infty} e^{ax} \{1 - F_X(x)\} = \infty \quad (1)$$

となる分布をヘヴィータイルを有すると定義する^{1),2)}。この定義から、指数分布や正規分布はヘヴィータイルを持たないのに対して、信頼性解析において広く用いられている確率分布の中で、対数正規分布や、形状母数が1より小さいWeibull分布はヘヴィータイルを持つ。

ヘヴィータイルを持つ分布の中で、さらに裾のゼロへの収束が遅く、 $x \rightarrow \infty$ の極限下での確率密度関数のゼロへの収束、あるいは、確率分布関数の1への収束がべき関数的であるような分布は、特にファットテイルを有すると呼ばれている^{3),4)}。あるいは、両者を総称して裾の重い分布ともよばれる。ファットテイルを持つ確率分布の代表例は、Pareto分布⁸⁾であり、その確率分布関数は次式で与えられる。

$$F_X(x) = \begin{cases} 1 - \left(1 + \frac{x}{\beta}\right)^{-\alpha} & (x \geq 0) \\ 0 & (x < 0) \end{cases} \quad (2)$$

ここで、 α, β は正のパラメーターである。

2.2 裾の重い確率分布に対する統計的推定の問題点

対象量の確率分布を推定するために収集するデータの個数が十分でない場合、当然推定の精度は上がらないが、裾の重くない分布と誤って推定してしまう危険性がある。

一例として、 $\alpha = 3, \beta = 20$ として、式(2)で与えられる Pareto 分布に従う乱数 10 個を計算機上で発生させ、

正規確率紙上にプロットした結果を Fig.1 に示す。ただし、確率分布関数の推定値はミーンランク法⁹⁾により算出してある。Fig.1 のフィッティングより得られた正規分布の確率密度関数と、設定した Pareto 分布の確率密度関数のテイル部分を対数目盛でプロットしたものを Fig.2 に示す。なお Fig.2 では相違を見やすくするために縦軸を対数としていることに加え、Pareto 分布の分布域である正の領域のみ表示していることから横軸との間の面積を比較しにくいのが、両確率密度関数ともに分布全域での横軸との間の面積は1となっている。Fig.1 では、危険率5%のKolmogorov-Smirnov検定(KS検定)では棄却できないことが確認できており、それゆえに確率紙上では比較的良好なフィッティングが得られていると考えられるが、裾の部分では桁違いの大きな誤差が生じてしまっている。本研究は、こういった問題における推定精度を定量的に論じる必要性を直接的な動機とするものである。

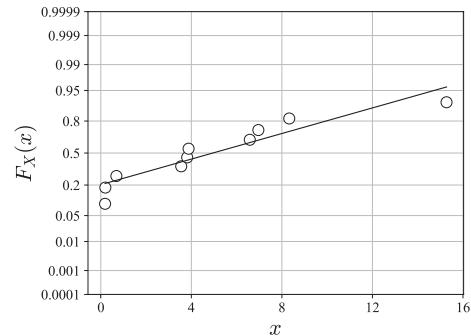


Fig. 1 Plot on a Gaussian probability paper for generated 10 random numbers obeying a Pareto distribution.

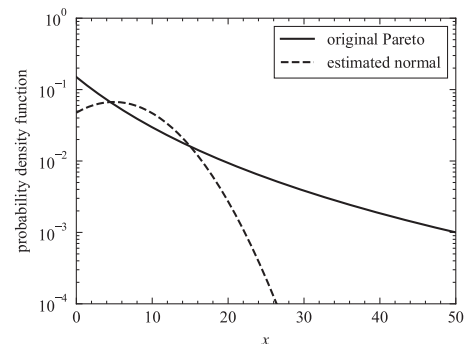


Fig. 2 Comparison of tails between the original Pareto distribution and the estimated Gaussina distribution.

3 裾が重い分布と多重モード分布との類似性

3.1 多重モード分布

材料や構造物の破壊の発生形態が複数のモードから成る場合、その劣化や損傷の指標となる X の確率分布は、多重モード型の分布となる。例えば、繊維強化複合材の破壊では、繊維部分の断裂に起因する破壊形態と、マトリックス部分の損傷に起因する破壊形態が共存するため、二重モード的な特徴を示すことが多い。

多重モード分布において、異なるモード間の分布特性の差異が大きい場合、裾の挙動がヘヴィーテイルあるいはファットテイルと類似することがある。分布に対する予備的知識が希薄であると、データのばらつきを見ただけでは、裾の重い分布であるのか、モードが複数の分布であるのか、についての判断が難しい。

3.2 寒冷地トンネル覆工コンクリートの損傷度の分布

著者らによる先行研究⁵⁾では、寒冷地トンネル覆工コンクリートの損傷度の不規則な時間成長を記述する確率モデルが提案され、北海道地区のトンネルにおける観測データのばらつきの様子をうまく再現できることが報告されている。ここで用いる「損傷度」とは、コンクリートのき裂損傷の大きさだけでなく、き裂の位置関係やはく離などの他の要因も考慮し、打音検査の結果も取り入れて総合的に数値化したものであり、15程度を上回ると至急の補修が必要とされている^{6),7)}。

文献⁵⁾では、竣工後時間 t [years] 経過後の損傷度 X の確率密度関数として、次式で与えられるような二重モード型の確率密度関数が提案されている。

$$f_X(x, t) = \sum_{i=1}^2 \zeta_i f_X(x, t; \mu_i) \quad (3)$$

$$f_X(x, t; \mu) = \frac{e^{-\lambda t}}{x} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} f_W^{n*} \left(\log \frac{x}{x_0} - \mu t \right) \quad (4)$$

$$f_W^{n*}(w) = \frac{(\rho w)^{n-1}}{(n-1)!} \rho e^{-\rho w} \quad (n = 1, 2, \dots) \quad (5)$$

ここで、 μ_1, μ_2 は各モードでの平均進展抵抗、 x_0 は初期平均損傷度、 λ は雑音として導入されている複合 Poisson 過程の単位時間当たりの平均ジャンプ発生回数、 ρ は複合 Poisson 過程のジャンプの確率分布に関するパラメーター、 ζ_1, ζ_2 は各モードの分配比率で、 $\zeta_1 + \zeta_2 = 1$ であり、 f_W^{0*} はデルタ関数を表す。

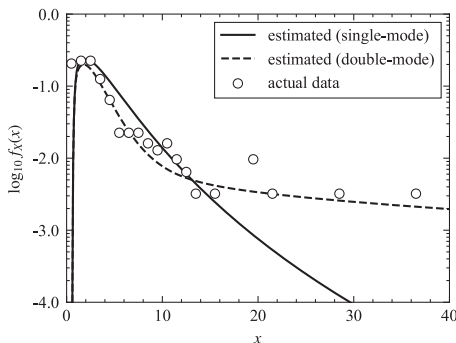


Fig. 3 Normalized frequency distribution of the damage degree of concrete liners and the theoretically derived probability density function as a double-mode distribution as well as a single-mode distribution.

Fig.3 は、竣工からの経過年数が 11[years] から 49[years] までのすべての損傷度のデータについて、規格化された相対度数値を「○」印で表し、これと比較する形で、式

(3)~式 (5) で与えられる確率密度関数のグラフを破線で描いたものである。ただし、分布裾野部分の挙動を拡大するために、縦軸は常用対数で表示してあり、デルタ関数の項は除いてある。ここでパラメーター値は、 $\lambda = 1.0$ [years⁻¹], $x_0 = 0.6$, $\rho = 21$, $\mu_1 = 0.045$ [years⁻¹], $\mu_2 = 0.15$ [years⁻¹], $\zeta_1 = 0.6$, $\zeta_2 = 0.4$ である。さらに比較のため、 $\zeta_1 = 1.0$, $\zeta_2 = 0$ とした単一モードでの確率密度関数を図中に実線で示す。このとき、いずれのモードとも確率密度関数の全区間での積分は 1 となることは確認してある。Fig.3 からわかるように、裾の部分に裾の重い分布に類似するような収束の遅い挙動が認められ、単一モード型の密度関数ではこのような裾の挙動を過小評価してしまうのに対して、二重モード型の確率密度関数とすることによりテイルを含めて全体の度数分布の傾向をうまく再現することができる。

これに対して Fig.4 は、Fig.3 に示したトンネル覆工コンクリートの損傷度の実データ全体からなる度数分布と、それと同じ平均・分散を持つように式 (2) 中の α, β を定めた Pareto 分布から抽出したサンプルの規格化された相対度数分布とを比較したものである。ただし、Pareto 分布からのデータの抽出個数は、損傷度の全データ数 (311 個) と同一としてある。この結果からわかるように、文献⁵⁾ で提案された二重モード型の確率分布の裾の振舞はファットテイルの代表例である Pareto 分布と類似しており、平均近傍の度数分布がほぼゼロに収束する付近よりさらに右側に断続的にサンプルが出現するという特徴が認められる。したがって、2.2 節で述べた分布形の推定における誤推定の危険性が二重モード型の確率分布に対しても存在すると考えられる。

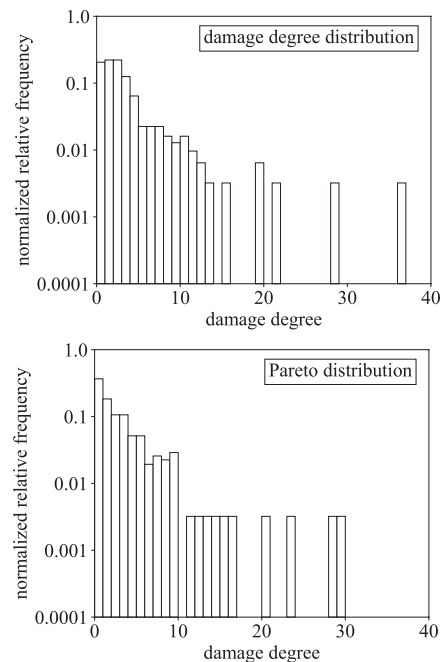


Fig. 4 Comparison of normalized frequency distributions between the damage degree distribution and a Pareto distribution with the same mean and variance.

4 特殊な裾形状を有する確率分布の推定に関する計算機実験

4.1 計算機実験の概要

2章および3章で述べたような性質を有する確率分布を、ここでは総称として特殊裾形状を持つ分布と呼ぶこととする。このような分布に対して、観測されたデータから確率分布形を確率紙を用いて推定する場合に、裾部分の特性を再現できるような分布形の推定が可能であるかどうか、またそのような推定にどの程度のサンプル数が必要となるのかについて、計算機実験を行って検証する。

4.1.1 実験の対象とする分布 対象とする特殊裾形状を持つ分布は、ファットテイル型の代表例である Pareto 分布と、二重モード型の良好なフィットが確認されているトンネル覆工コンクリートの損傷度の確率分布とする。ただし、比較の対象とするため、特殊裾形状を有しない分布として正規分布についても同じ実験を試みる。

4.1.2 仮想的なデータの生成法 計算機上で生成する仮想的な観測データは、あらかじめ設定した確率分布形に従うような乱数により生成する。ベースとなる区間 $[0, 1]$ の一様乱数はメルセンヌ・ツイスター¹⁰⁾を用いて生成し、Pareto 分布については逆関数法を、正規分布については Box-Muller 法を用いてそれぞれの分布に従う疑似乱数列となるように変換する。

損傷度の確率分布については実測データの中から疑似乱数を用いて抽出することによって仮想的観測データを生成する。データセット中から無作為に抽出すると、実際の観測における出現確率が反映されない抽出となるため、データセットから算出された平均・分散と同じ平均・分散を有する正規乱数を発生させ、生成された乱数値に最も近い値をデータセット中から抽出するという方法を用いる。

4.1.3 データからの確率分布の推定法 得られた仮想的な観測データから、確率紙を用いてフィッティングを行い、最もフィッティングが良いと判定された分布形を推定された分布形とする。本研究では、推定された分布の確率分布関数によるデータの累積度数分布に対する決定係数¹¹⁾ (Pearson の積率相関係数の2乗; R^2 と表す) が最も1に近い場合を最もフィッティングが良いと判定することとする。確率紙の種類としては、正規確率紙、対数正規確率紙、2母数 Weibull 確率紙、Gumbel 確率紙を用い¹²⁾、これらの中から推定分布形を選択する。

4.2 実験結果と考察

4.2.1 正規分布の場合 まず比較対象として、標準正規分布に従う母集団から抽出したデータに対して行った推定結果を Fig.5 に示す。Fig.5 は、横軸に抽出したデータの個数を、縦軸にそのデータから最もフィッティングが良いと判定された分布について、分布の裾部分のみで算出した決定係数 R^2 の値を1から減じた $1 - R^2$ を常用対数プロットしたものである。図中の「×」印は、4.1.3 項で述べた方法で最良と選択された分布形が正規分布で

あることを示している。なお、本研究では数値評価上の裾部分を確率分布の上位5%部分と設定しており、分布形の推定に用いた決定係数とは値が異なる点に注意する必要がある。

この結果からわかるように、正規分布に従う母集団から抽出されたデータについては、データ個数が少なくても分布形の誤推定の可能性が非常に低く、データ個数の増加に対応して裾部分の推定精度は順調に向上していく。本実験結果では、200~300 サンプル程度で、裾部分を含めてほぼ十分な推定精度が得られている。

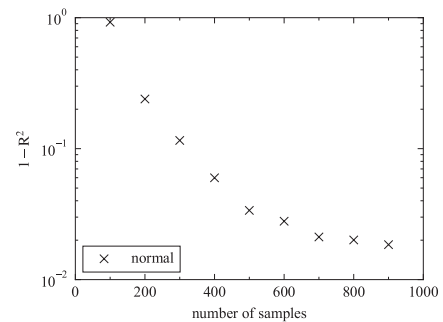


Fig. 5 Effect of the number of samples on the accuracy of the estimation for a data set sampled from a Gaussian distribution.

4.2.2 Pareto 分布の場合 次に、 $\alpha = 1, \beta = 1$ の Pareto 分布に従う母集団から抽出したデータに対して行った推定結果を示す。Fig.6 は、抽出サンプル数 100 個の場合について、最も良いフィッティングが得られた対数正規確率紙上のプロットを示したものであり、Fig.7 は推定された対数正規分布と元の Pareto 分布の確率密度関数の裾の部分の比較を示したものである。ただし、確率分布関数値の推定には Fig.1 と同様にミーンランク法を適用し、確率密度関数が非常に小さな値を取る部分を拡大表示するために、Fig.7 の縦軸は常用対数でプロットしてある。

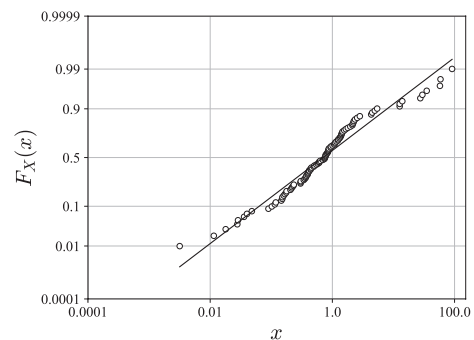


Fig. 6 Plot of a data set sampled from a Pareto distribution on a log-normal probability paper.

Fig.6 でも、Fig.1 と同様に危険率 5% の KS 検定では棄却できず、良好なフィッティングが得られているように見えるが、Fig.7 からわかるように裾の部分では正しい

数値よりもかなりの過小評価となってしまうことがわかる。

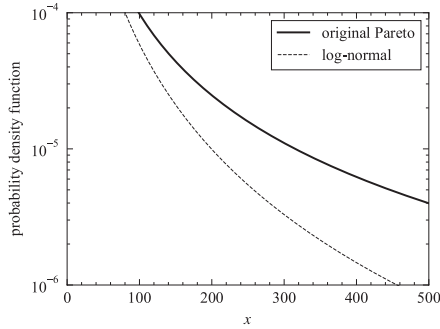


Fig. 7 Comparison of tails between the original Pareto distribution and the estimated log-normal distribution.

次に、抽出サンプル数が推定の精度、特に裾の部分の推定精度に与える影響を調べるために、Fig.5と同じ要領で裾の部分の決定係数値の1からの残差を対数プロットしたものをFig.8に示す。図中「△」の印は、4.1.3項で述べた方法で最良と選択された分布形が対数正規分布であることを示している。これより、抽出データ数を増加させても裾の部分の推定精度はなかなか上昇せず、データ数を1000程度にまで増やしても裾部分の十分な推定精度が得られているとは言い難いことがわかる。この結果をFig.5と比較すると、裾の重い分布と、裾形状が特殊でない分布では、推定精度に関する相違が顕著であると考えられる。

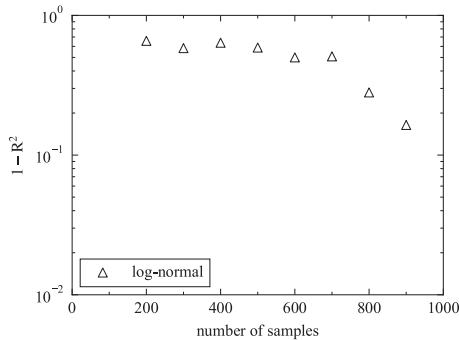


Fig. 8 Effect of the number of samples on the accuracy of the estimation for a data set sampled from a Pareto distribution.

4.2.3 トンネル損傷度分布の場合 次に、4.1.2項で述べた方法により、トンネル覆工コンクリートの損傷度データから抽出して行った計算機実験結果を示す。Fig.9は、抽出サンプル数100個の場合について、抽出されたデータセットに関して最も良いフィッティングが得られたWeibull確率紙上のプロットを示したものであり、Fig.10は推定された分布と、式(3)～式(5)で与えられる二重モード分布の確率密度関数の裾の部分の比較を示したものである。Fig.10の縦軸は常用対数でプロットしてあり、さらに比較のために単一モード ($\zeta_1 = 1, \zeta_2 = 0$) での確率密度関数も示してある。

文献⁵⁾で二重モード型の分布で裾野がよく説明できることが指摘されている損傷度データについても、Fig.9のように、少数の抽出サンプル数の場合は確率紙で比較的良好なフィッティングが得られる（この場合も危険率5%のKS検定で棄却できない）。しかし、Fig.9からわかるように、推定されたWeibull分布では単一モード型の分布よりは裾の部分をややかながら過大に評価できるが、二重モード型の分布よりは過小評価となっており、やはり分布の裾部分の推定は危険側の推定を与えてしまう。

次に、Fig.5と同じ要領で裾の部分の決定係数値の1からの残差を対数プロットしたものをFig.11に示す。図中「□」の印は、4.1.3項で述べた方法で最良と選択された分布形がGumbel分布であることを、また「○」印はWeibull分布であることを示している。これより、抽出データ数を増加させても裾部分の推定精度は上昇せず、抽出データ数を増加させると逆に裾部分の推定精度が低下するという傾向も認められる。また、最良のフィッティングを示す分布形も抽出サンプルの特性に依存して変化してしまうことがわかる。この結果をFig.5と比較すると、抽出サンプル数を増加させても、裾部分の推定精度だけでなく、分布形そのものの推定にも安定的傾向が得られないことがわかる。

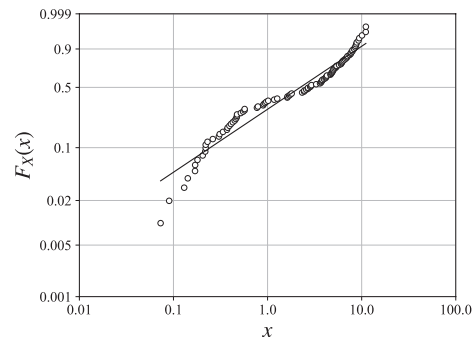


Fig. 9 Plot of a data set sampled from the actual data for the damage degree of concrete liners on a Weibull probability paper.

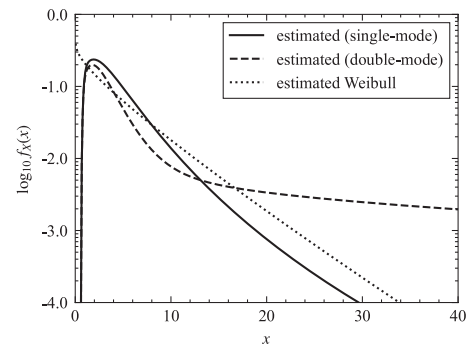


Fig. 10 Comparison of tails between the theoretically derived double-mode as well as single-mode probability density function and the estimated Weibull probability density function.

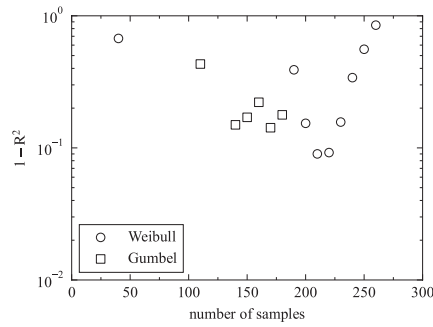


Fig. 11 Effect of the number of samples on the accuracy of the estimation for a data set sampled from actual data for the damage degree of concrete liners.

4.2.4 レベル超過確率の比較 トンネル覆工コンクリートの損傷度は数値が大きくなるほど覆工コンクリートの信頼性は低下するので、損傷度の確率分布に関して、あるレベルを損傷度が超える確率の推定が重要となる。Fig.12は、4.2.3項で推定された最良のフィッティングから得られる、所与のレベル x_c を損傷度が超過する確率 $1 - F_X(x_c)$ を、レベル x_c の関数として対数プロットしたものである。比較のため、文献⁵⁾で得られている二重モード型の分布から得られる超過確率と単一モード型の分布から得られる超過確率もプロットしてある。この結果からわかるように、確率紙から推定される最良フィットの確率分布でも、裾の部分をよく再現できる二重モード型の分布と比較して、レベル超過確率をかなり極端に過小評価、すなわち、かなり極端な危険側の評価を与えてしまうことがわかる。

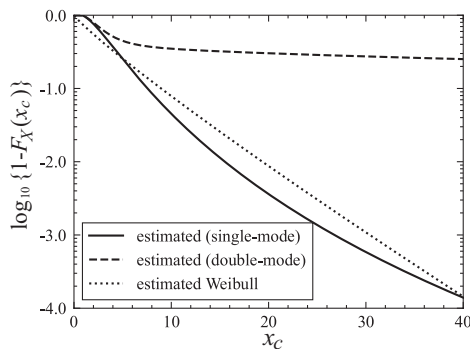


Fig. 12 Comparison of a level-exceedance probability for the estimated probability distribution, the double-mode probability distribution and the single-mode probability distribution

5 結 言

本研究では、裾の重い確率分布と、それに類似した挙動を示す多重モード型確率分布を対象として、確率紙を用いた分布形の統計的推定の精度、特に裾の部分の推定の精度に関して、推定に用いる抽出データ数との関係について計算機実験を通じて定量的な考察を行った。

この結果、裾形状が特殊でない分布である正規分布では抽出データ数を上げると裾の部分を含めて推定精度が上昇するのに対して、裾形状が特殊な分布では裾の部分の推定精度が上がらず、本研究で対象とした母集団の特性では、裾の部分について危険側の推定しか得られないことが明らかとなった。したがって、裾の部分の推定精度を上げるにはそれ以上のデータ数が必要となるが、こういった多数のデータを抽出することが困難なケースも多いことから、何らかの理論的なアプローチを駆使した分布形の推定が必要となると考えられる。

今後は、裾形状が特殊な分布の個々の特性と、必要となる裾の推定精度を得るために要するデータ抽出数についてさらに踏み込んだ考察が必要であろう。

参 考 文 献

- 1) T. Rolski, H. Schmidli, V. Schumidt and J. Teugels, “Stochastic processes for insurance & finance”, John Wiley & Sons (1999).
- 2) S. Foss, D. Korshunov and S. Zachary, “An introduction to heavy-tailed and subexponential distributions”, Springer Science & Business Media (2013).
- 3) B. Mandelbrot, “The variation of certain speculative prices”, The Journal of Business, Vol.36, No.4, pp.394-419 (1963).
- 4) T. Mikosch, “Regular variation, subexponentiality and their applications in probability theory”, Technical Report, Univ. of Groningen (1999).
- 5) H. Kanekiyo, O. Maruyama, A. Sutoh and T. Satoh, “Tail properties of probability distribution of damage degree for tunnel concrete liners in cold 4egion”, Journal of the Society of Materials Science, Japan, Vol. 65, No.4, pp.275-281 (2016).
- 6) A. Sutoh, T. Satoh and H. Nishi, “An identification of correlation between demand performances to damage of tunnel lining using AHP”, The Eleventh East Asia-Pacific Conference on Structural Engineering & Construction (Sensing & monitoring Techniques 28), pp.682-683 (2008).
- 7) T. Satoh, H. Nishi, A. Sutoh and O. Maruyama, “Investigation of evaluation and estimation of tunnel lining deterioration based on inspection data”, XIVth International Winter Road Congress Andorra 2014, Session D, p.96 (2014).
- 8) V. Pareto and Vilfredo, “Cours d’economie politique”, Journal of Political Economy, 6 (1898).
- 9) E. J. Gumbel, “Statistics of extremes”, Chapter 2, Columbia University Press (1958).
- 10) <http://www.math.sci.hiroshima-u.ac.jp/m-mat/MT/mt.html>
- 11) S.Wright, “Correlation and causation”, Journal of Agricultural Research, Vol.20, pp.557-585 (1921).
- 12) <https://www.i-juse.co.jp/statistics/product/>