

論文 / 著書情報
Article / Book Information

題目(和文)	システムの信頼性特性に関する統計学的研究
Title(English)	
著者(和文)	宮川雅巳
Author(English)	MASAMI MIYAKAWA
出典(和文)	学位:工学博士, 学位授与機関:東京工業大学, 報告番号:乙第1502号, 授与年月日:1985年12月31日, 学位の種別:論文博士, 審査員:
Citation(English)	Degree:Doctor of Engineering, Conferring organization: Tokyo Institute of Technology, Report number:乙第1502号, Conferred date:1985/12/31, Degree Type:Thesis doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

a-42

システムの信頼性特性に関する統計学的研究

宮川 雅巳

目 次

第1章 序 論	1
1.1 緒 言	1
1.2 本研究の背景	3
1.3 本研究の目的と論文構成	6
第2章 指数分布を仮定した競合モデル	11
2.1 緒 言	11
2.2 完全な競合データに対する統計的方法	12
2.3 特定モードによる故障の寿命時間のみが観測される 場合	21
2.4 故障原因不明データを含む場合	26
2.5 結 言	32
第3章 ワイブル分布を仮定した競合モデル	33
3.1 緒 言	33
3.2 ワイブル分布を仮定した競合モデルの基本的 性質	34
3.3 不完全な競合データに対する推定法の導出	42
3.4 漸近的性質とシミュレーション結果	47
3.5 結 言	54

第4章 ノンパラメトリックな競合モデル	55
4.1 緒言	55
4.2 モデルの背景と定式化	56
4.3 不完全な競合データに対する推定法の導出	64
4.4 シミュレーション実験による推定量の比較	73
4.5 適用例と考察	79
4.6 結言	91
第5章 コンポーネントの寿命データに基づくシステムの 寿命特性の推定	92
5.1 緒言	92
5.2 コヒーレントシステムの信頼度の推定	93
5.3 コヒーレントシステムのMTTFの推定	104
5.4 競合データとコンポーネントデータの併合	116
5.5 結言	118
第6章 故障メカニズムモデルと構造関数の推定	120
6.1 緒言	120
6.2 故障メカニズムモデルと信頼性データ	121
6.3 確率的コヒーレントシステムの定義とその性質	125
6.4 構造関数の推定	133
6.5 結言	136

第7章	ワイブル確率紙による母数の簡易推定法	138
7.1	緒言	138
7.2	ワイブル確率紙の数学的構造	139
7.3	極値分布の位置母数と尺度母数の推定	143
7.4	ワイブル分布の形状母数の推定	149
7.5	結言	154
第8章	結論	155
	参考文献	160
	付録	164

第1章 序 論

1. 1 緒 言

システムが要求された機能を意図する期間果たし続けていくことは、システムの設計および製造における重要な課題である。一般にシステムには様々な故障モードがあり、その影響も多様である。またその故障メカニズムも確率的法則に支配される部分が多く、問題を極めて難しいものになっている。さらに今日、システムを構成する部品、コンポーネントは、その種類も数も増加し、これらが複雑な構造をなしている。このようなシステムの信頼性を確保するためには、総合的な工学技術の体系が不可欠である。信頼性工学はこのような必要性から生まれてきたものである。

信頼性の研究は、第2次世界大戦の中頃、アメリカ空軍がレーダーなどの各種電子システムの故障に悩まされ続けたことから始められたが、その工学的基盤が確立されたのは1960年前後である。1952年、アメリカ国防省の中にAGREE (Advisory Group on Reliability of Electronic Equipment) が設置され、1957年にはAGREE報告が提出された。この報告書は、信頼性を定量的に表わす尺度を明確にし、この尺度で測られた信頼性を満足するシステムや機器を作り出すための手順を示している。さらに1960年代中頃よりこの分野の研究が増し、1970年以降は信頼性工学というひとつの工学体系が

出来上がったと言えよう。

信頼性の工学的手法の体系は次のA)、B)、C)の3つに大別して整理することができる。

A) システムの企画、設計の段階で用いられる手法で、FMEA (failure mode and effects analysis)、FTA (fault tree analysis)、設計審査などがある。

FMEAはグラマン社で海軍機の操縦システム開発の手法として、またFTAは、バルシステムにおける安全性開発の手法として、それぞれ実務面から開発された手法であるが、システム信頼性の理論的研究の成果がその後の発展に大きく貢献している。

B) システムの試作あるいはサブシステムの組立が終わった段階で用いられる手法で、故障解析(これには故障物理、再現試験などが含まれる)や寿命試験(これには限界試験、加速寿命試験などが含まれる)などがある。

故障解析、寿命試験においては、材料や部品の劣化、摩耗に対して確率過程を応用した理論研究により様々な寿命分布が導出され、さらにここで導出された寿命分布をもとに寿命データからその寿命特性を推定、検定する統計的方法が整備されてきた。

C) システムの実使用段階で得られる多くのデータを基にした手法で、品質情報収集システムの活用やワイブル解析などがある。

システムの実使用段階での工学的手法は

- 1) アメリカの自動車メーカーを中心としたマイル保証、消費者主義の台頭、リコール問題の発生などにより品質保証の中でとくに信頼性が重視されるようになった。
- 2) 人工衛星の打ち上げや、複雑なシステムによって構成される耐久消費財が普及した。

などの社会的背景からも、その重要性が認識されてきたが、そのための理論的研究は A)、B) に比べ立ち遅れている。実使用段階での観測対象はシステム自身であるが、フィールドにおける複雑なシステムの信頼性に関するデータから、有効な情報を引き出すためには統計学的な理論体系が不可欠である。

A)、B)、C) の手法は、それぞれ予防型、再発防止型、現状分析型と目的を異にするが、これらは本来独立に用いられるべきものではなく、相互の関連が重要である。フィールドにおけるシステムの寿命データ解析においても、単にシステムの故障時間に関する現状分析ではなく、故障モードやコンポーネントのレベルでの詳細な解析がなされてこそ、FMEAやFTA、さらに寿命試験へと有用な情報を展開することが可能になる。本研究は、このような点を踏まえて、実使用段階でのシステムの寿命情報を、システムの設計あるいは試作といった開発段階における工学的手法の中で活用させるための統計的方法論の一端について、それらの理論的考察を含めて論じている。

1. 2 本研究の背景

信頼性における寿命データは、

- 1) 故障以前に観測が打切られた中途打ち切りデータ(センサードデータとも呼ぶ)を含むことが多い。
- 2) 正規分布にあてはまらないことが多く、正規分布を基にした多くの統計的手法が適用できない。

という顕著な特徴をもっている。信頼性における統計的方法はこの点を踏まえて発展してきた。

信頼性工学が1950年代後半に開花する以前にも1935年 Gumbel が、また1937年 Weibull が極値分布およびワイブル分布をそれぞれ信頼性独自の分布として生み出している。また1952年 Davis が指数分布を信頼性の分野に導入し、1953年頃より Epstein と Sobel らにより指数分布に基づく統計的方法の理論的研究が進んだ。当時から寿命試験では試験時間短縮のために定時打ち切り方式、定数打ち切り方式が考案されていたが、それらによって得られるセンサードデータに対する解析法もこの中に含まれ、その結果は AGR E E 報告の一部になっている。センサードデータの問題は、寿命データの不完全性としてこのように早くから取り上げられ、寿命データ解析論の骨子をなしたものである。

1958年 Kao^[21] が真空管の寿命データに対してワイブル解析を行い、初期故障型のものと摩耗故障型のものに層別し、混合ワイブル分布をあてはめたことは、ワイブル確率紙とともにワイブル分布の普及に大きく貢献した。また、1956年 Lieblein と Zelen^[26] がボールベアリングの疲労寿命に対し対数変換を施し、順序統計量による線形推定量を導いたことは、ワイブル分布に対する統計理論の推進に役立った。一方、1961年 Zelen と Dannewiller^[54] によって指数分布を基にした統計手法が頑健性に乏しいことが指摘され、いっそうワイブル分布が普及し、センサードデータを含めてそのための推測法が研究されてきた。本論文でも実用面から、とくに簡便性を重視したワイブル確率紙上での

推定法を提示している。

一方、センサードデータを対象にしたノンパラメトリック手法や図的解析法は、むしろ臨床医学の生存時間研究の分野で発展してきた。この背景には、ある程度の数のデータがとれることと信頼性の寿命試験における定時打ち切り方式、定数打ち切り方式によって得られるセンサードデータとは異なるタイプのデータ、すなわちランダムセンサードデータがこの分野で存在していたためである。1958年 Kaplan と Meier^[22]によって Kaplan・マイヤー推定量 (product-limit 推定量とも呼ばれる) が提唱され、信頼性の分野では1960年 Herd が、また1964年 Johnson がそれぞれ類似した方法を提案したが、広く普及したのは1972年 Nelson が累積ハザード法を提案してからである。これらは寿命試験データよりもむしろフィールドデータの解析上有用となるもので、本論文でもこれらの手法を基にした解析法を提案している。

信頼性における寿命データの特徴として、上記2点にさらに加えるならば

- 3) 複数の故障モード、故障原因が存在する
- 4) 使用環境、使われ方に違いがある
- 5) 修理系では同一アイテムの故障を繰り返し観測する

などが挙げられよう。しかし信頼性の統計的方法において、これらに対する方法論の研究が確立されたのは1970年代になってからである。この背景には、FMEA、FTAといった手法の普及と品質情報収集システムの整備などがあり、信頼性データの型も {故障時間、故障モード、故障原因、使用環境、---} といった多次元データとして収集されるようになり、それに伴ないこの種のデータに対する解析法が必要不可欠になったことが挙げられる。たとえば、最も簡単な {故障時間、故障モード} という2次元データに対する1つのアプローチとして競合モデルの適用がある。競合モデルはもともと医学統計の分野で開

発されたものであるが、本研究で極めて重要な役割も果たすものであり、第2章以降で詳しく述べる。一方、使用環境や使われ方といった要因解析としては、1972年Coxが生存時間研究の分野で提唱した比例ハザードモデルとそれに基づく推定法が、信頼性の分野にも適用されつつあるが、この推定法には、複雑な数値計算を要するため、一般への普及にはまだ時間がかかると思われる。比例ハザードモデルについては本論文では取り扱わない。また5)に関しては、とくに履歴を考慮したモデルと解析法の重要性が今日強く叫ばれているが、これについても本研究では触れない。

本研究の背景として、もうひとつの重要な研究分野は、コヒーレントシステムの構造解析に代表されるシステム信頼性の理論である。この分野は1960年頃よりアメリカ西海岸派の研究者によって進められ、その主要な成果はBarlow and Proschan^[3]にまとめられている。本論文ではシステムの故障メカニズムの確率的モデルとして、コヒーレントシステムを拡張したものを提示している。またこの分野に関連して、コンポーネントの寿命データからシステムの寿命特性を推定する問題がある。これは1960年代中頃から始められたが、本研究では従来の結果を統一的に表現し、さらに適用範囲を広げている。

最後に、本研究で提唱する統計的方法の思想的背景のひとつとして、ジャックナイフ、ブートストラップおよび再配分といった「サンプルからの再抽出」(resampling from sample)の考え方を挙げる事ができる。本研究ではこれをノンパラメトリック手法に適用している。

1.3 本研究の目的と論文構成

寿命データの解析手法は、信頼性工学よりもむしろ医学統計の分野で意欲的に研究され 1970 年代から大きな進展も遂げてきた。しかしそこで開発された手法をそのまま信頼性の分野のフィールドデータに適用することは必ずしも得策でない。信頼性工学は、もともと固有技術との結びつきが強く、故障メカニズムに対する事前情報は一般に生物の死亡メカニズムに対するそれよりも豊富である。また反面、データ収集体制は臨床データに比べ劣り、よりデータが不完全であるという面も見逃せない。

本研究で目指す信頼性における統計的方法とは、データを獲得する場、取り扱うデータの量と質および型、さらにその時点で有する事前情報に応じた柔軟かつ多様な統計手法の体系である。その際、データからモデルを構築する「帰納」の段階と、モデルにデータをあてはめ知見を得る「演繹」の段階とが存在するわけだが、「帰納」の段階においても、故障メカニズムや分布形に関する情報を生かし、「演繹」の段階においては、念頭とするモデルからのズレに対しても頑健で、しかもそれぞれの段階で様々な不完全データの処理にも耐えられることが、本研究で提示する寿命データ解析法の特徴である。このようなアプローチは、固有技術と強い関連をもち、反面不完全データの多い信頼性の寿命データに適していると考えられる。そしてこのような一連の統計的手法を基に、フィールドでの寿命情報を新製品の設計・開発段階へ活用させる方法論を導くことが本研究の目的である。この目的のために本論文では以下のような論文構成を行っている。

第1章では、信頼性工学における本研究の位置付けと本研究の背景を述べるとともに、本研究の目的と論文構成を述べている。

第2章から第4章までは、フィールドにおけるシステムの寿命データ解析を

目的として、競合モデルを用いることにより、システムの故障時間と故障モードもしくは故障原因が観測されるというデータから故障モードごとの寿命特性を推定する方法について述べている。ここではモデルの基本的性質を明らかにし、様々な不完全データ、とくに故障原因不明データを含むデータに基づく寿命特性の推定法を与えている。これらの方法はフィールドでのシステムの寿命情報を、FMEA、FTAおよび寿命試験といった開発段階の工学手法において活用させる上で、有効なものと考えられる。

第2章では、指数分布を仮定した競合モデルを取り上げ、完全な競合データに基づく若干の統計的手法を提示するとともに、新たに2つのタイプの不完全な競合データを考案し、そのモデル化と解析手法を与える。ここではとくに不完全データを競合モデルの枠組みの中で定式化するという工夫をしている。この解析手法により、不完全な競合データに基づく各故障モードの故障率に対する不偏かつ有効な推定量が得られている。

第3章ではワイブル分布を仮定した競合モデルを扱っている、まずこのモデルの基本的性質を明らかにし、これよりシステムの寿命データから複数モードの競合性を検知するための指針を与えている。次に故障原因データの有無から3種類の競合データを考え、それぞれに基づく推定法を与え、これらの推定精度を比較することから母数推定におよぼす故障原因データの役割を定量的に評価している。また、故障原因不明データを含む場合については指数分布の場合と同様の方法では計算が極めて複雑になることから新たに、EMアルゴリズムを応用した効率的な推定アルゴリズムを与えている。

第4章ではノンパラメトリックな競合モデルを扱い、第2章、第3章と同様に、故障原因不明データを含む不完全な競合データに対する寿命特性の推定法を考察している。まず分布型に何ら制約を置かない完全なノンパラメトリック

推定において、十分統計量に基づく信頼度関数の一致推定量を与えるとともに、この場合の問題点を指摘する。次にこの問題点を踏まえて付加的仮定のもとでEMアルゴリズムを応用したいくつかの推定法を提示している。さらに故障原因データの役割についてはシミュレーション実験によりパラメトリックな場合と類似した結果を確認している。ここで得られたノンパラメトリック推定法は第7章で述べるワイブル確率紙による推定を含め、第2章、第3章で提示した方法の前処理段階の手法として極めて有効である。

第5章では、第2章から第4章までに述べた一連の統計的方法により、フィールドにおけるシステムの寿命データから抽出した各故障モード、各コンポーネントに関する寿命情報を、システムの開発段階における寿命試験の結果に展開し活用させる方法論について述べている。そのためにまず、複雑な構造をもつシステムのモデルとしてコヒーレントシステムを取り上げ、コンポーネントの寿命データからコヒーレントシステム全体の信頼度ならびに平均寿命を効率よく推定する方法を与えている。このとき最尤推定量、一樣最小分散不偏推定量、さらにノンパラメトリック推定量がいずれもシステムの構造関数の標準形を通じて統一的に表現されている。また、様々な見地からこれら推定量を比較した結果、ノンパラメトリック推定量が優れた性質をもっていることが明らかにされる。さらにここで与えた推定法の枠組みの中でフィールドにおけるシステムの寿命データを合わせて活用する方法について述べている。

第6章では、故障メカニズムモデルとしてより広い範囲も取り扱い、それに伴う信頼性データについて考察している。まず第2章から第4章までに論じた競合モデルをコヒーレント構造モデルに拡張する上で必要となる信頼性データの型を明らかにしている。次にコヒーレント構造モデルに代表される決定的故障メカニズムモデルから確率的故障メカニズムモデルまでを統一的に表現する

ものとして、新たに確率的コヒーレントシステムを定義し、その性質を述べている。さらに第5章に述べたシステムの寿命特性推定問題を故障メカニズムが不明な場合にも適用させるために、故障原因データによるコヒーレントシステムの構造関数の推定法ならびに確率的コヒーレントシステムの故障メカニズムに対する探索の指針を与えている。

第7章では、第5章に述べた推定問題において、コンポーネントの寿命分布としてワイブル分布を仮定したときに、極めて有効な手段となるワイブル確率紙による母数の簡易推定法を論じている。まずワイブル確率紙の数理的構造を明らかにし、線形推定の立場から、極値分布の位置母数と尺度母数およびワイブル分布の形状母数に対する新たな2点推定法を与えている。この2点推定量により小標本においては既存の様々な推定量に劣らぬ推定精度が得られることが明らかにされる。またここで与えたワイブル確率紙による推定は、第4章で得られたノンパラメトリック推定法を用いることにより、第3章におけるワイブル分布を仮定した競合モデルの母数推定にも適用可能である。

第8章は結論として、本研究で得られた知見をまとめ、残された問題、今後の展望について考察している。

第2章 指数分布を仮定した競合モデル

2.1 緒言

競合モデル (Competing Risks Model) は医学統計の分野で発達したものであり、生物とくに人間の死亡メカニズムとして、複数の死亡原因が存在しそれらが互いに競合し結果として、その中の1つの原因により死に至るというモデルである。これを信頼性における故障解析の立場に置き換えると、複数の故障原因が存在し、それらが競合する直列システムモデルと同等であることがわかる。いま、複数のコンポーネントからなる直列システムを考えれば、この直列システムについてはシステムの故障時間と故障原因 (どのコンポーネントの故障によりシステムが故障したかを示す) のデータが得られることになる。このとき、システムを構成するコンポーネントを故障モードと対応させて考えることもできる。

信頼性の分野においてフィールドデータから寿命特性を推定する際、観測される対象は必然的に製品としてのシステムであるから、個々のコンポーネントもしくは故障モードの寿命特性を推定したいときには、競合モデルを仮定してシステムの寿命データから推測を行うことが多い。競合モデルの起源は18世紀の物理学者 D. Bernoulli にさかのぼり、彼は「もしこの世から天然痘がなくなれば、人間の寿命はどのくらい延びるであろうか」という問いに対し、上記モデルの原型を構築した ((13) の付録Aを参照)。これからわかるように、

競合モデルのほとほとのねらいは個々の原因の影響の定量的把握と原因を排除したときの効果の予測にあった。

一方、競合モデルはフィールドにおいて多く見られるランダムセンサリングされたデータ（以下、ランダムセンサードデータと呼ぶ）の定式化にも用いることができ、極めて汎用なモデルと言える。本章では、2.2節で競合モデルの一般的な定式化といくつかの問題点を述べると共に、指数分布を仮定した競合モデルにおいて完全な競合データが得られる場合に対しては、故障率の点推定量、および逆抽出による正確な区間推定法を与える。さらに2.3節と2.4節では、指数分布を仮定した競合モデルにおいて新たな2つの不完全データのモデル化とそれらに基づく母数の推定法を示し、完全な競合データが得られる場合との比較を行う。

2.2 完全な競合データに対する統計的方法

2.2.1 問題の定式化

故障時間とともに故障モードもしくは故障原因が観測される状況も考える。故障時間、故障モードを表わす確率変数をそれぞれ X, S で表わせ、 (X, S) を観測するものとする。通常、 X は非負の実数値もとり、 k 種の故障モードを考えれば S は $\{1, 2, \dots, k\}$ 上の値をとる。

上記データに対して2つの解析アプローチがある。1つは原因別ハザード (cause-specific hazard) 分析と呼ばれるもので、興味の対象は原因別ハザード関数

$$\lambda_j(t) = \lim_{\Delta t \rightarrow 0} \frac{Pr\{t \leq X < t + \Delta t, \delta = j \mid X \geq t\}}{\Delta t} \quad (2.1)$$

にある。すなわちこのアプローチは、観測される X と δ の同時分布から導出される諸特性の推測を目的とした言わば現象の叙述の立場である。

これに対し、いま1つのアプローチが競合モデルに基づく解析である。競合モデルでは、故障原因 j のみが存在したときの仮想的な故障時間 X_j を考え ($j = 1, \dots, k$)、アイテムの故障時間ならびに故障原因がそれぞれ

$$\begin{aligned} X &= \min(X_1, X_2, \dots, X_k) \\ \delta &= \{j \mid X = X_j\} \end{aligned} \quad (2.2)$$

で与えられるとするものであり、観測される現象を説明するための1つの構造モデルと言える。しかしながら、競合モデルが広く受け入れられるには次の2つの障害があった。

- 1) X_1, \dots, X_k の物理的意味が不明確なことが多い。
- 2) X_1, \dots, X_k の独立性・従属性が同定できない。

1) は医学統計の分野ではその通りであろうが、信頼性においては直列システムのその構成コンポーネントのように X_1, \dots, X_k の物理的意味が明確であることも少なくない。一方、2) は Tsiantis [49] らによって指摘されたことだが、任意の X_1, \dots, X_k に対して、 $X = \min(X_1, \dots, X_k)$ 、 $\delta = \{j \mid X = X_j\}$ と同一の同時分布をもつ X' 、 δ' がたがいに独立な X'_1, \dots, X'_k によって $X' = \min(X'_1, \dots, X'_k)$ 、 $\delta' = \{j \mid X' = X'_j\}$ として構成されてしまう。そのため X_1, \dots, X_k の間に特定の従属モデルを仮定しない限り、

X_1, \dots, X_k の同時分布の推定は不可能になる (たとえば [13] の 4 章や [47] を参照): 一方, X_1, \dots, X_k に独立性を仮定し, X_j の故障率関数を

$$r_j(t) = \lim_{\Delta t \rightarrow 0} \frac{Pr \{ t \leq X_j < t + \Delta t \mid X_j \geq t \}}{\Delta t} \quad (2.3)$$

とすれば, (X, δ) の観測データより $r_j(t)$ が推定され得るが, このとき $r_j(t)$ は式 (2.1) の $\lambda_j(t)$ に等しい. すなわち, 独立性を仮定した競合モデルによる解析と原因別ハザード分析とは形式上は全く同等なのである. したがって現象の叙述の立場にたてば, 競合モデルで独立性を仮定することはいっように差しつかえなく, また個々の原因の影響の定量的把握という競合モデル本来の立場においても, 一般の従属モデルの諸特性に対して独立な競合モデルにより上下限を与えることができるため (たとえば [42]), 独立性を仮定した競合モデルの研究は有意義なものと言える.

一方, フィールドで多く観測されるランダムセンサーデータの定式化にも競合モデルが適用される. たとえば, 図 2.1 に示すような使用開始時点がそれぞれ異なるアイテムにおいて, 現時点で未故障のものは中途打ち切りになるので使用開始時点をそろえ小さい順に並べれば図 2.2 のように故障と打ち切りの入り混じったデータとなる. このようなデータをランダムセンサーデータと呼ぶ. 使用開始時点が暦時間において等しくても実稼動時間がそれぞれ異なれば同様のデータを得る. ランダムセンサーデータは競合モデルにより次のように定式化される. 寿命を表わす確率変数 X_1 とこれに独立な打ち切り変数 X_2 とすれば

$$X = \min (X_1, X_2)$$

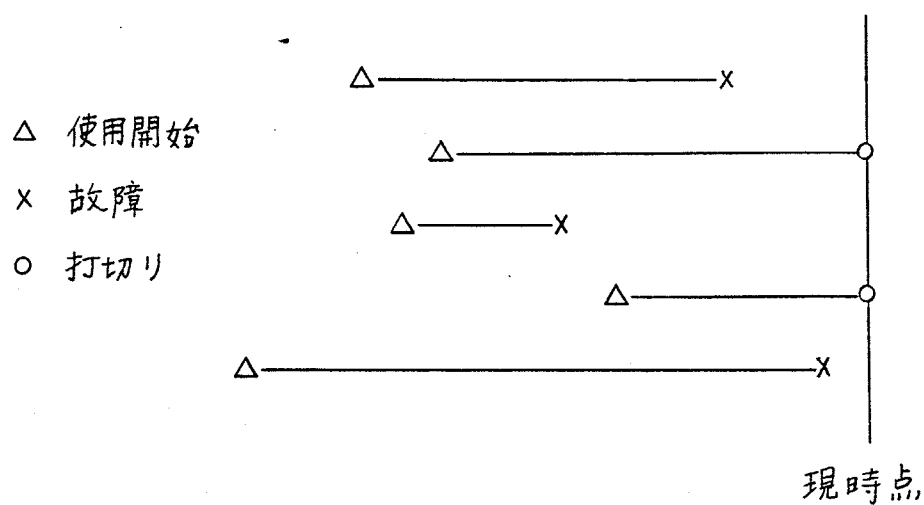


図 2.1 使用開始時点の異なるデータ

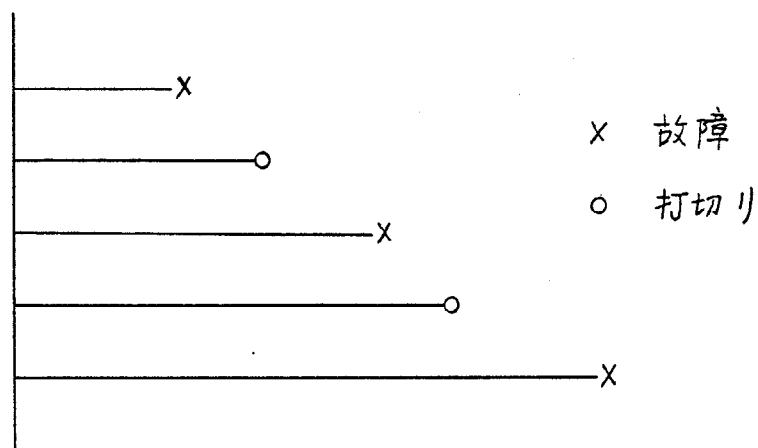


図 2.2 ランダムセンサーデータ

$$\delta = \begin{cases} 1 & , \quad X = X_1 \text{ のとき} \\ 2 & , \quad X = X_2 \text{ のとき} \end{cases} \quad (2.4)$$

が観測されることになり、 $k=2$ の競合モデルに帰着する。ここで $\delta=1$ は故障、 $\delta=2$ は打切りを表わしている。図2.1の例では、使用開始時点から現時点までの時間を打切り変数 X_2 と考えればよい。

また一般に、 k 種の故障モードが存在する競合モデルにおいても、特定の故障モードの寿命特性にのみ興味があるならば、その他のモードをまとめて1つのモードとすればこの場合にも $k=2$ の競合モデルに帰着することがわかる。以下、独立性を仮定した $k=2$ の競合モデルにおいて議論を展開する。

2.2.2 基本的性質と点推定

この節では、各モードの寿命分布として指数分布を仮定した $k=2$ の競合モデルを取り扱う。いま、モード j の寿命を確率変数 X_j で表わし、これは故障率 λ_j の指数分布

$$f_j(x) = \lambda_j e^{-\lambda_j x}$$

に従うものとする($j=1, 2$)。また X_1 と X_2 はたがいに独立とする。さらに観測されるものは式(2.4)で定められる (X, δ) とする。いま、 n 個のシステムを観測してデータ (X_i, δ_i) ($i=1, \dots, n$)が得られているとする。ここで、 n 個のシステムの寿命はたがいに独立と仮定する。このように n 個のシステムについて故障時間と故障原因がすべて得られるとき、得られるデ

- タを完全な競合データと呼ぶ。 (X_i, δ_i) ($i=1, \dots, n$) に関しては次のような性質がわかっている [13]。

- 1) X は故障率 $(\lambda_1 + \lambda_2)$ の指数分布に従う。
- 2) モード 1 による故障数 $r = \sum_{i=1}^n \delta_i$ は 2 項分布 $B(n, p)$ に従う。ただし $p = \lambda_1 / (\lambda_1 + \lambda_2)$ である。本章では以降、 p はすべてこの意味で用いる。
- 3) X と δ は独立である。すなわちシステムの故障時間と故障原因は独立である。

さてそこで、完全な競合データが得られる場合の故障率 λ_1 の点推定について考える。 (X_i, δ_i) ($i=1, \dots, n$) の同時密度は

$$\begin{aligned} f(x_1, \dots, x_n, \delta_1, \dots, \delta_n) &= (\lambda_1 + \lambda_2)^n e^{-(\lambda_1 + \lambda_2) \sum_{i=1}^n x_i} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^{\sum_{i=1}^n \delta_i} \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n - \sum_{i=1}^n \delta_i} \\ &= \lambda_1^r \lambda_2^{n-r} e^{-(\lambda_1 + \lambda_2) \sum_{i=1}^n x_i} \end{aligned} \quad (2.5)$$

となる。ここで $r = \sum_{i=1}^n \delta_i$ である。したがって λ_1 の最尤推定量は

$$\hat{\lambda}_1 = \frac{r}{\sum_{i=1}^n x_i} \quad (2.6)$$

となる [13]。また、式 (2.5) より $(\sum_{i=1}^n X_i, r)$ が (λ_1, λ_2) の完備十分統計量であり

$$E\left(\frac{1}{\sum_{i=1}^n X_i}\right) = \frac{\lambda_1 + \lambda_2}{n-1}$$

となることから

$$\lambda_1^* = \frac{n-1}{n} \cdot \frac{r}{\sum_{i=1}^n X_i} \quad (2.7)$$

が λ_1 の一様最小分散不偏推定量である [13] . λ_1^* の分散は $n > 2$ で

$$\text{Var}(\lambda_1^*) = \frac{n \lambda_1^2 + (n-1) \lambda_1 \lambda_2}{n(n-2)} \quad (2.8)$$

となる. $\hat{\lambda}_1$ の平均二乗誤差は

$$\text{MSE}(\hat{\lambda}_1) = \frac{(n+2) \lambda_1^2 + n \lambda_1 \lambda_2}{(n-1)(n-2)}$$

であるから, $n > 2$ では一様に $\text{MSE}(\hat{\lambda}_1) > \text{MSE}(\lambda_1^*) = \text{Var}(\lambda_1^*)$ の関係がある. また, $n > 2$ では λ_1^* をさらに縮小した

$$\hat{\lambda}_1^* = \frac{n-2}{n} \cdot \frac{r}{\sum_{i=1}^n X_i}$$

の平均二乗誤差が

$$\text{MSE}(\hat{\lambda}_1^*) = \frac{n \lambda_1^2 + (n-2) \lambda_1 \lambda_2}{n(n-1)}$$

となり, 一様に $\text{MSE}(\lambda_1^*) > \text{MSE}(\hat{\lambda}_1^*)$ の関係がある. ここで $\rho = \lambda_1 / \lambda_2$ と n がともに小さい場合には, $\hat{\lambda}_1^*$ を用いることにより平均二乗誤差はかなり縮小される. しかし $\hat{\lambda}_1^*$ は λ_1^* よりも故障率 λ_1 を小さ目に推定するので, ど

ちらを用いるかは推定の目的によると思われる。

2.2.3 逆抽出による区間推定

次に、 λ_1 の区間推定について考える。完全な競合データが得られるとき、 $2(\lambda_1 + \lambda_2) \sum_{i=1}^n X_i$ が自由度 $2n$ の χ^2 分布に従う。しかし、この性質を用いて λ_1 の信頼区間を求めることは、局外母数 λ_2 が含まれているので難しい。しかしながら、次の逆抽出が可能な場合には、以下に示すように正確な信頼区間を求めることができる。この逆抽出とは、モード 1 による故障が一定数 r になるまで観測を続ける方法であって、このとき総観測数を N で表わせば、これは確率変数となる。

ところで定時打ち切り方式の寿命試験では、打ち切り時間を τ とすれば、 $\min\{X, \tau\}$ が観測されるが、このときの総試験時間は定数打ち切りの場合と異なつてガンマ分布に従わず、その分布は複雑な形になり、正確な信頼区間を求めることは簡単ではない。しかしこの場合に、逆抽出により一定の故障数が観測されるまで試験を続けると、そのときの総観測時間はガンマ分布に従うことが、Nadler^[41] によって示されている。しかし、この性質をランダムセンサリングのなされる場合に直接適用することはできない。

いま、観測数 N は負の二項分布 $NB(r, p)$ に従い、また $N = n$ の条件のもとでは $\sum_{i=1}^n X_i$ はガンマ分布 $\Gamma(\lambda_1 + \lambda_2, n)$ に従う。ゆえに $Z = \sum_{i=1}^n X_i$ の確率密度関数は

$$f(z) = \sum_{n=r}^{\infty} \binom{n-1}{r-1} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^r \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-r} \frac{(\lambda_1 + \lambda_2)^n}{(n-1)!} z^{n-1} e^{-(\lambda_1 + \lambda_2)z}$$

$$\begin{aligned}
&= \frac{\lambda_1^r}{(r-1)!} z^{r-1} e^{-\lambda_1 z} \sum_{n=r}^{\infty} \frac{(\lambda_2 z)^{n-r}}{(n-r)!} e^{-\lambda_2 z} \\
&= \frac{\lambda_1^r}{(r-1)!} z^{r-1} e^{-\lambda_1 z}
\end{aligned}$$

となる。これより、「モード 1 が原因の故障が一定数 r になるまで観測を続けるとき、総観測時間 $\sum_{i=1}^N X_i$ はガンマ分布 $[(\lambda_1, r)]$ に従う」という Nadler が導いたのと同様の結果が導ける。したがって、逆抽出のもとでは、信頼水準 $1-\varepsilon$ の λ_1 の信頼区間は

$$\left[\frac{\chi_{2r}^2(1-\frac{\varepsilon}{2})}{2 \sum_{i=1}^N X_i}, \frac{\chi_{2r}^2(\frac{\varepsilon}{2})}{2 \sum_{i=1}^N X_i} \right]$$

で与えられる。ただし、 $\chi_{\phi}^2(\varepsilon)$ は自由度 ϕ の χ^2 分布の上側 ε 点とする。
またこのとき、 λ_1 の不偏推定量は

$$\hat{\lambda}_1^{\circ} = \frac{r-1}{\sum_{i=1}^N X_i}$$

であり、その分散は $r > 2$ では

$$\text{Var}(\hat{\lambda}_1^{\circ}) = \frac{\lambda_1^2}{r-2}$$

となる。そしてこのときの期待総観測時間は

$$E\left\{ \sum_{i=1}^N X_i \right\} = \frac{r}{\lambda_1}$$

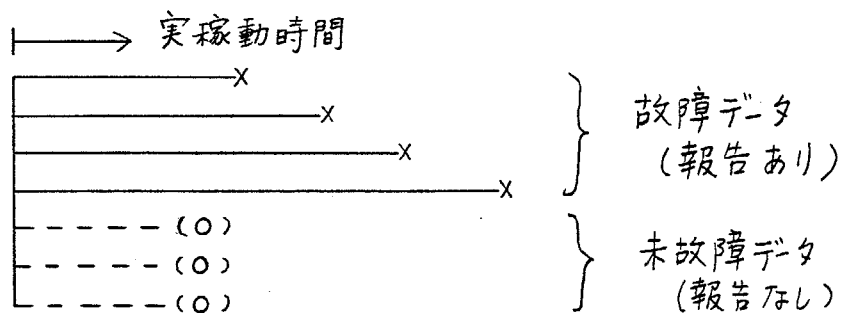
となる。以上の考察より、フィールドデータから λ_1 を推定する際、逆抽出が可能であるならば、これを用いることにより推定精度が λ_2 に依存しなくなるのがわかる。

2.3 特定モードによる故障の寿命時間のみが観測される場合

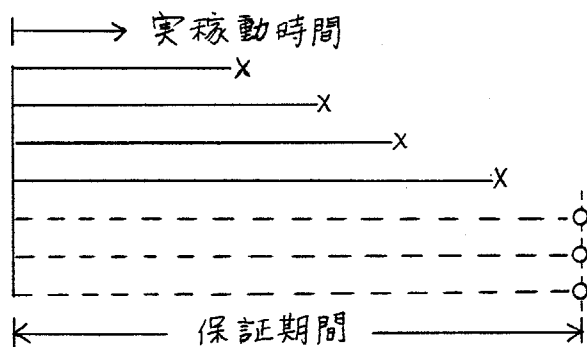
2.3.1 問題の背景と定式化

一般に、市場よりメーカーが収集するフィールドデータでは、故障データのみがクレーム情報として収集され、正常な未故障データに関しては、どこまで使用して未故障なのかの情報は入手されない場合がある。いま製品の保証期間を考えれば、ユーザーからメーカーへの修理の請求などにより故障データの収集は可能である。この際、故障データのみで信頼性特性を推定することは、たとえば現時点までに死亡した人の寿命のみをもって人間全体の寿命特性を推定するに等しく、寿命は過小評価されてしまう。したがって現時点で稼働している未故障品をも考慮して解析する必要がある。

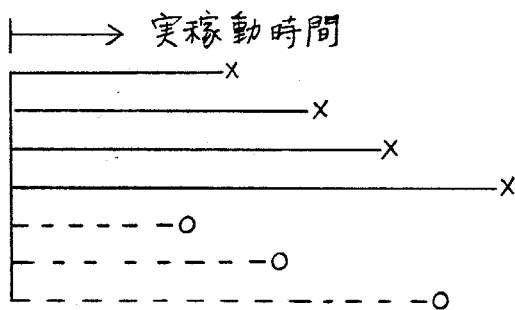
上述のデータは図2.3(a)のようになる。このとき、実稼働時間と暦時間とが等しい製品に対しては、たとえば一年間の保証期間中、修理の請求がなければその間正常であったとみなし、図2.3(b)のようなデータを考え、従来の定時打切りデータの解析法を用いればよい。ところが、自動車、建設機械、複写機などのように実稼働時間(車では実走行距離)と暦時間とが一致しない製品に対しては上記の考え方は適用できない。使用頻度や一日当たりの実稼働時間がユーザーによって異なるものに対しては、実稼働時間は千差万別であり、図



(a) X-カーが入手するデータ



(b) 実稼動時間と暦時間が等しい場合



(c) 実稼動時間と暦時間が等しくない場合

図 2.3 保証期間中のフィールドデータ [39]

2.3(c)のように故障データ(\times 印)と未故障データ(\circ 印)が入り混じっているからである。しかも、その未故障品の稼働時間が観測されていない点に問題がある。このようなデータを競合モデルの枠組みで考えると、特定のモードによる故障時間だけが観測され、その他のモードによる故障時間は観測されない場合に相当する。部品メーカーについて言えば、自社の部品による故障データだけが入手され、他社の部品による故障データや未故障データは入手されない場合である。

いま、故障までの時間(もしくは特定のモードによる故障時間)を X_1 、保証期間中の実稼働時間(もしくはその他のモードによる故障時間)を X_2 とすれば、 $k=2$ の競合モデルに帰着し、式(2.4)の表現を用いれば、 $\{X=x, \delta=1\}$ および $\{X>0, \delta=2\}$ が観測されることになる。すなわち、未故障($\delta=2$)の製品の打ち切り時間が欠測値となるわけである。このように、 (X_i, δ_i) ($i=1, \dots, n$)の一部が観測されないデータを本論文では不完全な競合データと呼ぶことにする。上記データはそのうちの1つである。

2.3.2 推定量の導出とその性質

いま、 n 個のシステムに対して、 r 個の $\{X_i, \delta_i=1\}$ ($i=1, \dots, r$)と $n-r$ 個の $\{\delta_i=2\}$ ($i=r+1, \dots, n$)が得られたとき、尤度は一般に

$$L \propto \left[\prod_{i=1}^r f_1(x_i) R_2(x_i) \right] \cdot \left[P_r(X_1 > X_2) \right]^{n-r} \quad (2.9)$$

と書ける。ただし $f_j(\cdot)$ 、 $R_j(\cdot)$ はそれぞれ X_j の確率密度関数、信頼度関

数である ($j=1, 2$). X_1, X_2 がそれぞれ故障率 λ_1, λ_2 の指数分布に従うときは, 式(2.9)は

$$L = \binom{n}{r} \left(\frac{-\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-r} \lambda_1^r e^{-(\lambda_1 + \lambda_2) \sum_{i=1}^r x_i} \quad (2.10)$$

と書ける. また

$$\frac{\partial}{\partial \lambda_1} \ln L = \frac{r}{\lambda_1} - \frac{n-r}{\lambda_1 + \lambda_2} - \sum_{i=1}^r x_i \quad (2.11)$$

$$\frac{\partial}{\partial \lambda_2} \ln L = \frac{n-r}{\lambda_2} - \frac{n-r}{\lambda_1 + \lambda_2} - \sum_{i=1}^r x_i$$

となるから, 式(2.11)の値を0とおくことにより, λ_1 の最尤推定量 $\tilde{\lambda}_1$ は $r > 0$ のとき

$$\tilde{\lambda}_1 = \frac{r^2}{n \sum_{i=1}^r x_i} \quad (2.12)$$

で与えられることがわかる ($r=0$ のときは λ_1 の最尤推定量は存在しない).

またここでは, フィールドデータの解析を対象とし, データが多数得られることを仮定しており

$$P_r \{ r=0 \} \longrightarrow 0 \quad (n \longrightarrow \infty)$$

となることに注意すれば, 最尤推定量の一般論が適用可能になる. そこで, 推定量 $\tilde{\lambda}_1$ をその漸近分布の分散で評価し, この漸近分散をもとに 2.2 節のように完全な競争データが観測される場合と, この節のモデルのようにモード 1

が原因の故障時間だけが観測される場合の λ_1 の推定量の比較を行う。

完全な競合データに基づく最尤推定量 $\hat{\lambda}_1$ の漸近分散は

$$A. \text{Var}(\hat{\lambda}_1) = \frac{\lambda_1^2 + \lambda_1 \lambda_2}{n} \quad (2.13)$$

である。一方、 $\tilde{\lambda}_1$ の漸近分散は

$$A. \text{Var}(\tilde{\lambda}_1) = \frac{\lambda_1^2 + 2\lambda_1 \lambda_2}{n} \quad (2.14)$$

である。式(2.13), (2.14)の導出については付録1を参照されたい。これより $A. \text{Var}(\hat{\lambda}_1) \leq A. \text{Var}(\tilde{\lambda}_1)$ の関係があることがわかるが、これは $\tilde{\lambda}_1$ についてはモード2が原因の故障時間が観測されていないのだから当然の関係と言える。また、 $\tilde{\lambda}_1$ の $\hat{\lambda}_1$ に対する漸近相対効率

$$A. R. E.(\tilde{\lambda}_1 : \hat{\lambda}_1) = \frac{\lambda_1^2 + \lambda_1 \lambda_2}{\lambda_1^2 + 2\lambda_1 \lambda_2} \quad (2.15)$$

となる。 $\rho = \lambda_1 / \lambda_2$ の値を変えたときの式(2.15)の値を表2.1に示す。

表 2.1 $\tilde{\lambda}_1$ の $\hat{\lambda}_1$ に対する漸近相対効率

$\rho = \lambda_1 / \lambda_2$	0.20	0.50	1.00	2.00	5.00
A. R. E.	0.55	0.60	0.67	0.75	0.86

2.4 故障原因不明データを含む場合

2.4.1 問題の背景と定式化

一般に、システムには複数の故障モード、故障原因が存在し、それらがたがいに競合し結果としてその中の1つの原因で故障に至る。競合モデルでは観測対象となるシステムに対して、{故障時間, 故障原因}という2次元データが得られるものとしているが、多くの場合データの中にはシステムの故障以前に観測が打切られたものも含まれ、その場合には必然的に故障原因は観測されないものとしている。このデータは式(2.4)の表現で表わせば、 $\{X > \infty, \delta = ?\}$ と書けるが、既存の競合理論の枠組みの中でこのデータは取り扱うことができる。

ところが一般に、故障原因の究明には多くの時間と費用を要するため、フィールドでは故障データにおいてもその故障原因が観測されないことが多い。とくに、故障した時点で直ちに廃棄される消費財においては、すべての対象製品に対しその故障原因も完全に把握することは不可能と言えよう。このとき、得られるデータは $\{X = \infty, \delta = ?\}$ と書ける。

いま、 n 個のシステムについて、故障時間 X_1, \dots, X_n だけが観測される場合を考えると、これらの同時密度は

$$f(x_1, \dots, x_n) = (\lambda_1 + \lambda_2)^n e^{-(\lambda_1 + \lambda_2) \sum_{i=1}^n x_i}$$

となるから、 $\sum_{i=1}^n X_i$ が $\Lambda = \lambda_1 + \lambda_2$ の完備十分統計量となり、 Λ の一様最小分

散不偏推定量は

$$\hat{\Lambda} = \frac{n-1}{\sum_{i=1}^n X_i}$$

で与えられる。しかし、この場合に λ_1 の最尤推定量や一様最小分散不偏推定量を導くことはできない。すなわち、 $\sum_{i=1}^n X_i$ が $\Lambda = \lambda_1 + \lambda_2$ に関する情報はもっているが、 $p = \lambda_1 / (\lambda_1 + \lambda_2)$ に関する情報をもっていないためである。したがって、指数分布を仮定した競合モデルにおいては、一部のシステムについては故障原因を究明することが不可欠と言える。

そこで、故障原因の観測されたデータ $\{X=x, s=j\}$ と観測されていないデータ $\{X=x, s=?\}$ が混入した場合を考える。すなわち、図2.4に示すように n 個のシステムの寿命データの中で m 個 ($m \leq n$) については故障時間ならびに故障原因が観測され、残りの $n-m$ 個については故障時間だけが観測されているようなデータを考える。ここで m は定数とする。このとき、「故障原因が観測されるか否かは、故障時間や真の故障原因に無関係である」と仮定

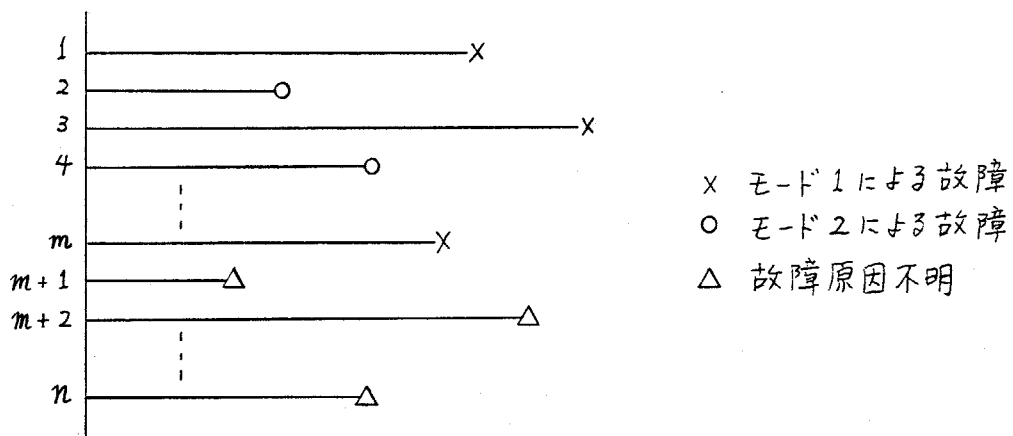


図 2.4 故障原因不明を含む競合データ

する。すなわち、どちらも同一の母集団からの無作為標本であるとする。これはたとえば、ある時期に市場に出す製品のうち、無作為に選んだ何台かの製品については、モニター制度あるいは追跡調査により故障時間と故障原因データも収集し、残りについては故障時間だけが収集される状況を考えればよい。

2.4.2 推定量の導出とその性質

そこで、上記のような故障原因不明データを含むデータに基づく指数分布を仮定した競合モデルの母数推定の問題を考える。上記仮定のもとで、 (x_i, δ_i) ($i=1, \dots, m$), x_i ($i=m+1, \dots, n$) のデータが得られたとすると尤度は

$$L = \lambda_1^r \lambda_2^{m-r} (\lambda_1 + \lambda_2)^{n-m} e^{-(\lambda_1 + \lambda_2) \sum_{i=1}^n x_i} \quad (2.16)$$

と表わせる。ここで $r = \sum_{i=1}^m \delta_i$ である。そして

$$\frac{\partial}{\partial \lambda_1} \ln L = \frac{r}{\lambda_1} + \frac{n-m}{\lambda_1 + \lambda_2} - \sum_{i=1}^n x_i \quad (2.17)$$

$$\frac{\partial}{\partial \lambda_2} \ln L = \frac{m-r}{\lambda_2} + \frac{n-m}{\lambda_1 + \lambda_2} - \sum_{i=1}^n x_i$$

となるから、式(2.17)の値を0とおくことにより λ_1 の最尤推定量は

$$\hat{\lambda}_1 = \frac{r}{m} \cdot \frac{n}{\sum_{i=1}^n x_i}$$

で与えられることがわかる。また、式(2.16)は

$$\lambda_2^n (\lambda_1 + \lambda_2)^{n-m} \exp \left\{ -(\lambda_1 + \lambda_2) \sum_{i=1}^n X_i + r \ln \left(\frac{\lambda_1}{\lambda_2} \right) \right\}$$

と書き改められ、指数型分布族に属することがわかる。よって $(\sum_{i=1}^n X_i, r = \sum_{i=1}^m \delta_i)$ が (λ_1, λ_2) の完備十分統計量になる (Lehmann [25] p. 132)。また、 $\sum_{i=1}^n X_i$ はガンマ分布 $\Gamma(\lambda_1 + \lambda_2, n)$ に従い、 r は2項分布 $B(m, p)$ に従う。さらに、 $\sum_{i=1}^n X_i$ と r は独立である。これより λ_1 の一様最小分散不偏推定量は

$$\bar{\lambda}_1 = \frac{r}{m} \cdot \frac{n-1}{\sum_{i=1}^n X_i}$$

で与えられることになる。

次に、故障原因のわかるデータ数 m と推定量の分散との関係を調べてみる。 $m = n$ が2.2節の完全な競合データの場合に相当していることは言うまでもない。 $\bar{\lambda}_1$ の分散は

$$\text{Var}(\bar{\lambda}_1) = \frac{m \lambda_1^2 + (n-1) \lambda_1 \lambda_2}{m(n-2)} = \left(1 + \frac{n-1}{m \rho} \right) \frac{\lambda_1^2}{n-2} \quad (2.18)$$

となる。ここで $\rho = \lambda_1 / \lambda_2$ である。一方、完全な競合データが得られる場合の λ_1^* の分散は

$$\text{Var}(\lambda_1^*) = \left(1 + \frac{n-1}{n \rho} \right) \frac{\lambda_1^2}{n-2} \quad (2.19)$$

と書き直せる。よって $\bar{\lambda}_1$ ($m < n$) の λ_1^* ($m = n$) に対する相対効率は、式(2.18), (2.19)より

$$\text{eff.}(m; n, \rho) = \frac{(n-1+n\rho)m}{(n-1+m\rho)n} \quad (2.20)$$

となる。ここで n と ρ を固定し、仮に m を連続変数とみなせば、式(2.20)が

$$\frac{\partial}{\partial m} \text{eff.}(m; n, \rho) > 0$$

ならびに

$$\frac{\partial^2}{\partial m^2} \text{eff.}(m; n, \rho) < 0$$

という性質をもつことが容易に導かれる。 $n = 100$ とし、 ρ を固定し式(2.20)を m の関数としてみたときのグラフを図2.5に示す。図2.5より、たとえば $m = 80$ では $\rho = 1.0$ の場合にも 88% 以上の相対効率がおり、 $m = 30$ でも $\rho = 5.0$ と大きい場合には 72% 以上の相対効率があることがわかる。

また、このような不完全な競合データにおいて、故障原因のわかっているデータ (X_i, δ_i) ($i = 1, \dots, m$) だけから λ_1 を推定する場合と、故障原因不明データも合わせて用いる場合の推定精度の比較のために、 $m = 10$ としたときの $n - m$ の増加に伴う $\text{Var}(\bar{\lambda}_1)$ の変化の様子を図2.6に示す。図2.6より、たとえば $\rho = 4.0$ では 13 個、 $\rho = 2.0$ では 20 個の故障原因不明のデータを合わせて使うことで分散の値を半分以下に縮小できることがわかる。ただし、 m を固定したときには、 $n - m$ をいくら増加させても $\text{Var}(\bar{\lambda}_1)$ の下限は $(\lambda_1 \lambda_2) / m$ である。

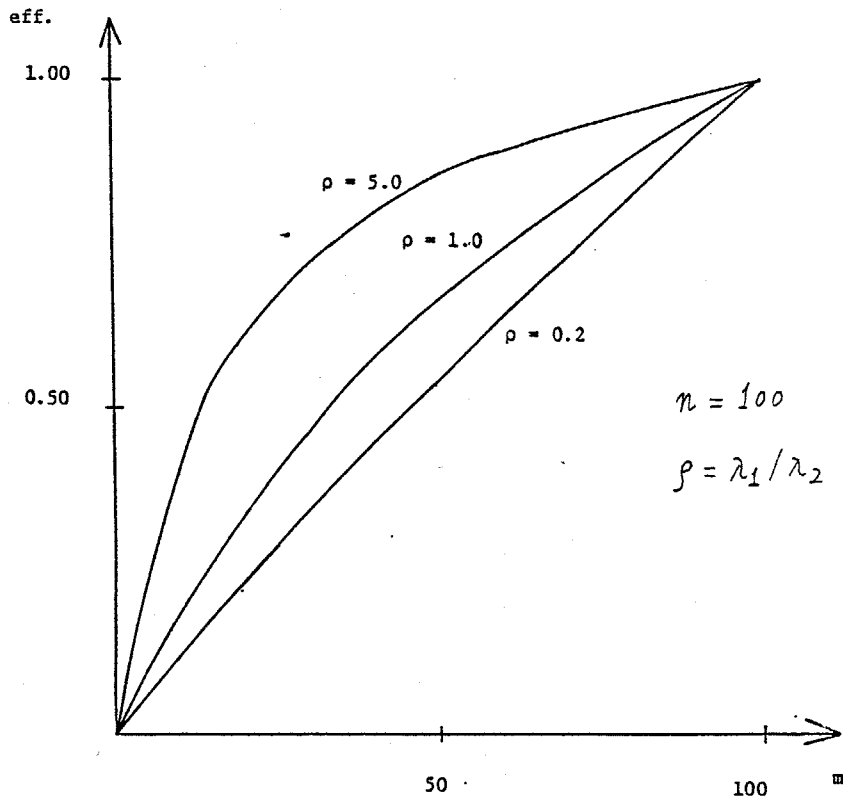


図 2.5 $\bar{\lambda}_1$ の λ_1^* に対する相対効率

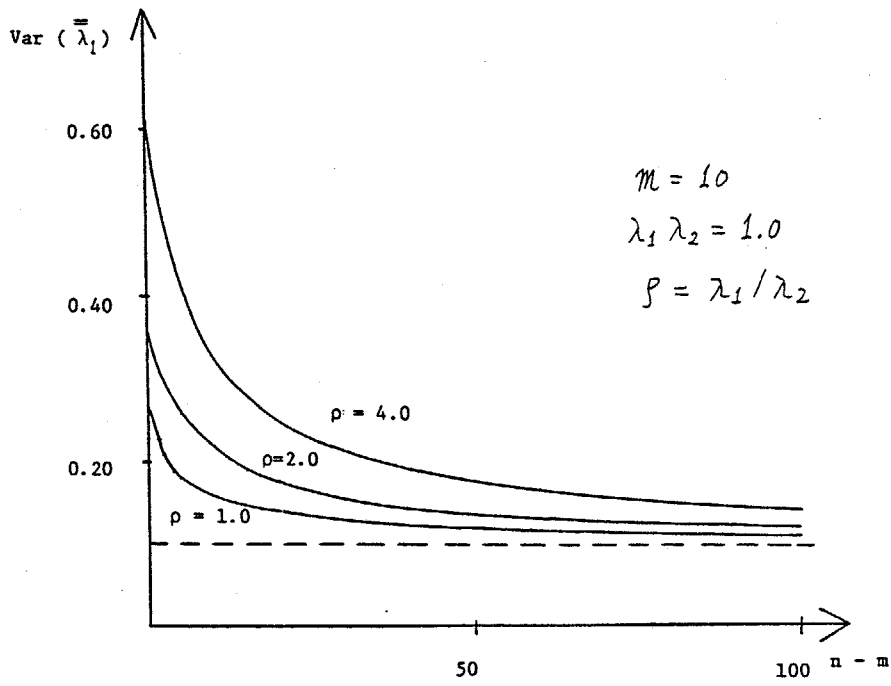


図 2.6 $\bar{\lambda}_1$ の分散 (m を固定)

2.5 結 言

本章では、競合モデルの目的と意義を述べるとともに、指数分布を仮定した競合モデルについて考察し、次の結果を得た [31]。

- 1) 完全な競合データに対して、逆抽出を行うことによって、局外母数の影響を受けない点推定ならびに区間推定の方法を導いた。
- 2) 特定のモードによる故障時間のみが観測される不完全な競合データの定式化を行い、これに基づく推定法の導出とその評価を行った。これより、データの偏りを除去でき、しかも比較的効率よく母数が推定されることが明らかになった。ここでの議論は、フィールドにおける保証期間中のクレームデータにも適用される。
- 3) 故障原因不明データが存在する不完全な競合データの定式化を行い、これに基づく推定法の導出とその評価を行った。これより、たとえば $\rho = \lambda_1 / \lambda_2 = 5.0$ のときには、約半数の故障原因を観測すれば、完全な競合データの場合に比べ約80%の推定効率が得られることがわかった。

なお、指数分布を仮定した競合モデルでは、モード数 $k=2$ における上述の結果は、推定量の形ならびにその性質に関して、 $k > 2$ でもそのまま成立する。

第3章 ワイブル分布を仮定した競合モデル

3.1 緒言

ワイブル分布はもともとスウェーデンの物理学者 Weibull が材料の破壊強度の分布として提唱したものである。今日、セラミックスをはじめとする新素材が開発され注目を集めているが、構造用セラミックスなどのぜい性材料は、「強度のばらつき」という重大な欠点をもっている。この欠点を克服するために、材料科学に基づく微構造のコントロール、非破壊検査を中心とするスクリーニング、保証試験などが試みられており、確率論と破壊力学に基づく強度解析がこれらと結びつき、構造物の設計に応用されつつある。しかしながら、これらの研究の大部分は、1種類の欠陥のみを取り扱っている。実際の構造用セラミックスには、内因欠陥（気孔、介在物、粗大結晶粒、き裂---）と外因欠陥（機械加工による損傷、環境に起因する損傷---）に大別される数多くの欠陥を内包しており、その強度に対してはこれら多種類の欠陥が破壊源となる。

これに対し、Bolotin [6]、松尾ら [28] は、多種類の破壊源をもつぜい性材料の強度に関する研究を行い、Weibullと同様に最弱リンク説に基づいて多重モードワイブル分布を導出した。この多重モードワイブル分布は、ワイブル分布を仮定した競合モデルと等価である。ここで破壊源は、故障モードもしくは故障原因と対応させて考えることができる。しかしながら、導出された多重モードワイブル分布の性質や、これに基づく統計的方法についてはほとんど研究

がなされていない。本章では、ワイブル分布を仮定した競合モデルにおけるいくつかの基本的性質を明らかにして多重モードワイブル分布を適用する上での指針を与える。また一方、破壊強度試験において、破壊源が観測されないことが多いという事実を踏まえて、故障原因データの有無により3種類の競合データを与え、それぞれに対する推定法とその推定精度について考察する。

なお、本章の議論は、材料の破壊試験における統計解析のみならず、個々のコンポーネントの寿命がそれぞれ異なるワイブル分布に従うような直列システムに対しても適用可能であるので、フィールドデータ解析上極めて有用なものと言える。

3.2 ワイブル分布を仮定した競合モデルの基本的性質

3.2.1 ワイブル確率紙上での性質

k 種類の故障モードが存在する競合モデルを考え、モード j の寿命 X_j に対して、2母数ワイブル分布

$$f_j(x) = \frac{m_j}{\theta_j} \left(\frac{x}{\theta_j}\right)^{m_j-1} \exp\left\{-\left(\frac{x}{\theta_j}\right)^{m_j}\right\} \quad (3.1)$$

を仮定する ($j=1, \dots, k$)。ここで、 m_j, θ_j はそれぞれ形状母数、尺度母数と呼ばれるものである。さらに、 X_1, \dots, X_k はたがいに独立と仮定する。このとき、 X_1, \dots, X_k の最小値

$$X = \min(X_1, \dots, X_k)$$

の確率密度関数は

$$f(x) = \left[\prod_{j=1}^k \left(\frac{m_j}{\theta_j} \right) \left(\frac{x}{\theta_j} \right)^{m_j-1} \right] \exp \left\{ - \sum_{j=1}^k \left(\frac{x}{\theta_j} \right)^{m_j} \right\} \quad (3.2)$$

で与えられる。これが Bolotin, 松尾らが導いた多重モードワイブル分布の最も単純な形であり, ここではこれを長モード競合ワイブル分布と呼ぶ。

そこでまず, 寿命試験や破壊強度試験などで得られるデータが長モード競合ワイブル分布にあてはまるかどうかを簡便的に判断する方法について考える。通常の2母数ワイブル分布に対しては, ワイブル確率紙に打点をし, ほぼ直線上に点が並ぶかどうかで判定できる。ワイブル確率紙の数理的構造は第7章で述べるが, 2母数ワイブル分布の信頼度関数の式

$$R(x) = \exp \left\{ - \left(\frac{x}{\theta} \right)^m \right\}$$

の両辺に対して, 対数を2回とると

$$\ln \ln \frac{1}{R(x)} = m \ln x - m \ln \theta$$

の関係を満足することから, $\ln x$ と $\ln \ln [1/R(x)]$ との間に形状母数 m を傾きとする線形関係が成り立つことを利用している。そこで式(3.2)の長モード競合ワイブル分布において, $\ln x$ と $\ln \ln [1/R(x)]$ がいかなる関係にあるかを調べてみる。

いま, X_j が式(3.1)の2母数ワイブル分布に従うとき, これを対数変換した $Y_j = \ln X_j$ は極値分布(2重指数分布とも呼ぶ)

$$G_j(y) = 1 - \exp \left[- \exp \left(\frac{y - u_j}{b_j} \right) \right] \quad (3.3)$$

に従う。ここで $G_j(\cdot)$ は Y_j の分布関数であり、位置母数 u_j と尺度母数 b_j はそれぞれもとの母数 m_j, θ_j と、 $u_j = \ln \theta_j, b_j = 1/m_j$ の関係にある。また、 $X = \min(X_1, \dots, X_k)$ の対数変換 $Y = \ln X$ と $\min(Y_1, \dots, Y_k)$ は等しいので、 Y の分布関数は

$$G(y) = 1 - \exp \left\{ - \sum_{j=1}^k \exp \left(\frac{y - u_j}{b_j} \right) \right\} \quad (3.4)$$

で与えられる。さらに、 $Y_j = \ln X_j$ のとき ($j = 1, \dots, k$)、 $\bar{G}(y) = 1 - G(y)$ と $R(x)$ の値は等しいので式(3.2)において $\ln x$ と $\ln \ln [1/R(x)]$ の関係を見ることは、式(3.4)において y と $\ln \ln [1/\bar{G}(y)]$ の関係を見ることに帰着する。式(3.4)で、移項した後、対数を2回とれば

$$\begin{aligned} \ln \ln \frac{1}{\bar{G}(y)} &= \ln \ln \left[\sum_{j=1}^k \exp \left(-\frac{u_j}{b_j} \right) \exp \left(\frac{y}{b_j} \right) \right] \\ &= \ln \ln \left[\sum_{j=1}^k a_j \exp \left(\frac{y}{b_j} \right) \right] \end{aligned} \quad (3.5)$$

を得る。ここで $a_j = \exp(-u_j/b_j)$ で、正の係数である。

式(3.5)を y で微分すれば

$$\frac{d}{dy} \left[\ln \ln \frac{1}{\bar{G}(y)} \right] = \frac{\sum_{j=1}^k \frac{a_j}{b_j} e^{y/b_j}}{\sum_{j=1}^k a_j e^{y/b_j}} \quad (3.6)$$

となり，さらに2回微分をとれば

$$\frac{d^2}{dy^2} \left[\ln \ln \frac{1}{G(y)} \right] = \frac{\sum_{i < j} a_i a_j \exp\left(\frac{y}{b_i} + \frac{y}{b_j}\right) (b_i - b_j)^2 / b_i^2 b_j^2}{\left[\sum_{j=1}^k a_j e^{y/b_j} \right]^2} \quad (3.7)$$

と計算され，この値は $b_1 = b_2 = \dots = b_k$ ，すなわち $m_1 = m_2 = \dots = m_k$ のときには0となり，その他の場合には正の値になることがわかる。ゆえに，形状母数が異なるワイブル分布に従う2つ以上のモードが競合した場合，そのときの競合ワイブル分布はワイブル確率紙上では下に凸の曲線になることがわかる。したがって，ワイブル確率紙上で直線ではなく，下に凸の曲線になったときには，異なるタイプの故障モードが競合している可能性を示している。

さらに， $k=2$ に限って，傾きの挙動を細かく調べてみると以下のことがわかる。式(3.6)を $k=2$ の場合で書き改めると

$$\begin{aligned} \frac{d}{dy} \left[\ln \ln \frac{1}{G(y)} \right] &= \frac{1}{b_1} + \frac{a_2 \left(\frac{1}{b_2} - \frac{1}{b_1} \right) e^{y/b_2}}{a_1 e^{y/b_1} + a_2 e^{y/b_2}} \\ &= \frac{1}{b_2} - \frac{a_1 \left(\frac{1}{b_2} - \frac{1}{b_1} \right) e^{y/b_1}}{a_1 e^{y/b_1} + a_2 e^{y/b_2}} \end{aligned}$$

を得る。いま， $b_1 > b_2$ （すなわち $m_1 < m_2$ ）とすれば，傾きは $1/b_1$ と $1/b_2$ の間をとることがわかる。また， y は $(-\infty, \infty)$ を動くので，傾きは y の十分小さいところでは $1/b_1$ に，逆に y の十分大きなところでは $1/b_2$ に近づくことが定性的にわかる。

Kao^[21] は，真空管の寿命データをワイブルプロットした結果，かなり顕著な凸状を示すことから，初期故障を起こす真空管と摩耗故障を起こす真空管と

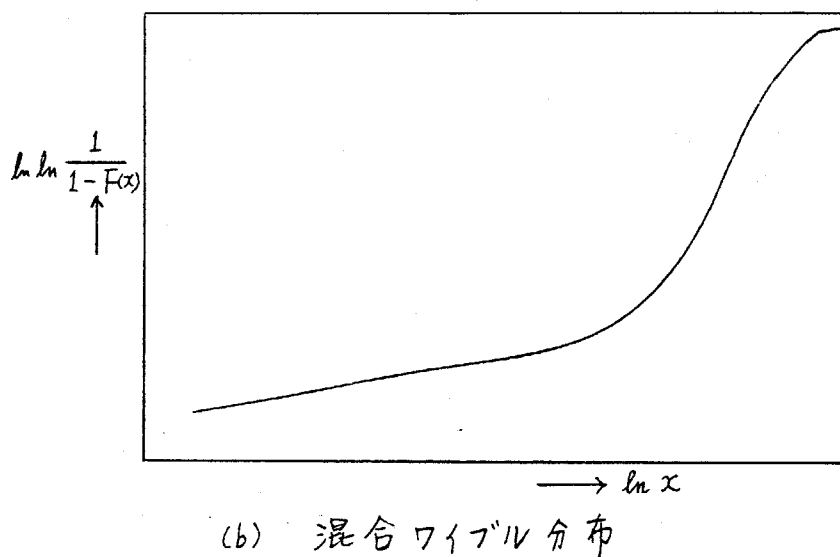
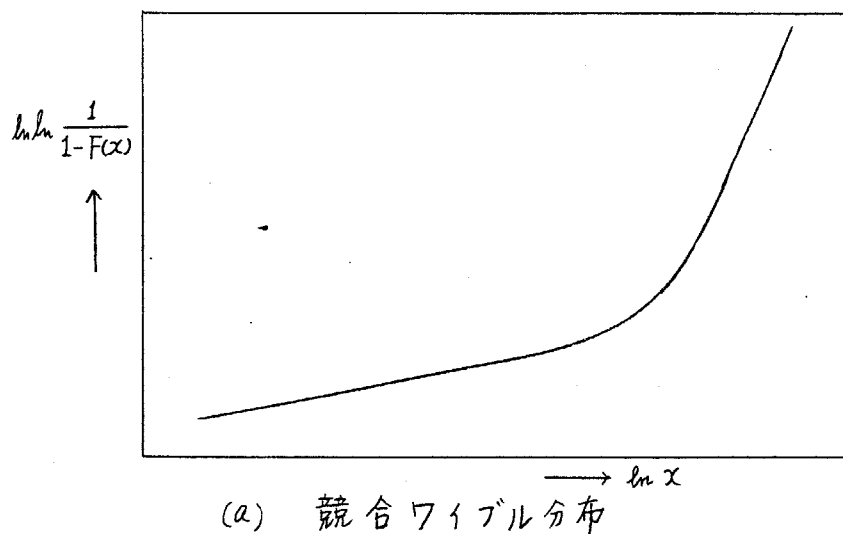


図 3.1 ワイブル確率紙上での競合ワイブル分布と混合ワイブル分布

摩耗故障を起こす真空管との混合モデルを想定し、母数推定も行っている。しかしながら、初期故障を起こす真空管に摩耗故障の要因が存在しないとは考えにくいので、競合性を加味したモデルをあてはめるべきと思われる。事実、初期故障型と摩耗故障型の2つのモードの混合モデルでは、ワイブル確率紙上において、厳密には下に凸でなく、 x の十分大きいところでは再び傾きが小さくなるS字状になることが知られている(図3.1参照)。蟹江と野中^[20]は、

Kao のデータに対して競合モデルに基づき推定した曲線と混合モデルに基づき推定した曲線のあてはまり具合を比較しているが、 $F(x) > 0.99$ のデータがないため、このままではどちらのモデルがより適合しているかの判定はできないとしている。

このようにデータ数が少ないときには、競合モデルと混合モデルともアイテムの故障時間データのみから判定することは、顕著に S 字状を示す場合以外は難しい。しかしながら、この 2 つのモデルが意味する故障メカニズムは全く異なっている。すなわち、競合モデルでは、個々のアイテムがすべての故障モードを潜在的に有し、どの故障モードによっても故障する可能性があるのに対し、混合モデルでは、それぞれのアイテムが固有の故障モードを有し、それによってのみ故障するとしている。したがって、2 つのモデルの判定は重要な意義をもつ。本節で明らかにしたワイブル確率紙上での長モード競合ワイブル分布の性質は、データ数が十分大きい場合に、この判定に 1 つの指針を与えるものと言える。

3.2.2 競合データの分類と推定問題

従来、多重モードワイブル分布に基づく破壊強度試験の統計解析においては、破壊強度データのみから母数推定が行われる場合が多かった^[6]。しかしながら、破壊強度データとともに破壊原因データが観測される場合にはこれを併用することが、推定精度上有利であることは明らかである。ここでは、一般的な競合モデルの枠組みの中で、2.3 節と同様に、故障原因データが観測されるか否かにより、競合データを次の 3 つに分類する。

- i) 完全データ ----- すべてのアイテムについて故障時間ならびに故障原因が観測されたもの
- ii) 故障原因不明データ ----- すべてのアイテムについて故障時間だけが観測されたもの
- iii) 一部故障原因不明データ ----- 一部のアイテムについては故障時間と故障原因が観測され、残りのアイテムについては故障時間だけが観測されたもの

ここで、iii) については、2.3 節と同様「故障原因が観測されるか否かは、その故障時間や真の故障原因に無関係である」という仮定を満足しているものとする。また、ワイブル分布を仮定した競合モデルではモード数 k により、推定値を求める数値計算の計算量、推定量の性質などが異なるので、 $k = 2$ に限定せずに以下の議論を展開する。

まず、i) の完全データが得られる場合の母数推定問題について考える。いま、全アイテム数を n (定数) としたとき、モード j が原因の n_j 個の故障時間 x_{j1}, \dots, x_{jn_j} が観測されたとする ($j = 1, \dots, k, n = \sum_{j=1}^k n_j$)。このとき、尤度は一般に

$$L \propto \prod_{j=1}^k \prod_{i=1}^{n_j} \left\{ f_j(x_{ji}) \prod_{\substack{l=1 \\ l \neq j}}^k R_l(x_{ji}) \right\} \quad (3.8)$$

と書ける。ここで $f_j(\cdot)$, $R_j(\cdot)$ はそれぞれモード j の確率密度関数、信頼度関数である。式(3.8)を変形すると、次の有用な式を得る [13]。

$$L \propto \prod_{j=1}^k L_j \quad (3.9)$$

$$L_j = \prod_{i=1}^{n_j} f_j(x_{ji}) \prod_{\substack{l=1 \\ l \neq j}}^k \prod_{i=1}^{n_l} R_j(x_{li})$$

式(3.9)で注目すべきことは、 L_j がモード j の分布のみから構成されていることである。ここで、一般性を失うことなく各モードの寿命分布の母数が関数的に独立であるとするれば、式(3.8)の L を最大化することは、式(3.9)の各々の L_j を最大化することに帰着し、尤度方程式は各モードで独立したものになる。式(3.1)を式(3.9)に代入すれば、次のような簡単な尤度方程式を得る。

$$\frac{n_j}{m_j} + \sum_{i=1}^{n_j} \ln x_{ji} - \frac{n_j \sum_{l=1}^k \sum_{i=1}^{n_j} x_{li}^{m_j} \ln x_{li}}{\sum_{l=1}^k \sum_{i=1}^{n_j} x_{li}^{m_j}} = 0 \quad (3.10)$$

$$\theta_j = \left(\frac{\sum_{l=1}^k \sum_{i=1}^{n_j} x_{li}^{m_j}}{n_j} \right)^{1/m_j}$$

m_j, θ_j の最尤推定値は式(3.10)の解として定まり、これは Newton-Raphson 法により容易に求まる。このように最尤推定値は局外母数とは無関係に定まるが、最尤推定量の性質は局外母数の値に依存する。

次に、ii) の故障原因不明データに基づく母数推定について考える。観測される故障時間 X の確率密度関数は式(3.2)であるが、このオ1項に和の形があるため、尤度方程式は極めて複雑な非線形連立方程式となる。完全データの場合の式(3.10)では、まずモードごとに尤度方程式が分離し、しかも式(3.10)のオ1式には未知母数として m_j しか含んでいない。そのため数値計算は極めて容易だが、ii) の場合にはモードごとに分離せず、それぞれの尤度方程式に未

知母数のすべてが含まれる。松尾と村田^[30]は、 $k=2$ の競合ワイブル分布において、Nelder-Mead法（シンプレックス法）を用いて直接対数尤度の最大化を試みることで、母数の推定を行っているが、モンテカルロ・シミュレーションで確認した結果、真値に比べ異常に大きな推定値が得られることがあることを報告している。また、蟹江と野中^[20]は、やはり $k=2$ の場合にワイブル確率紙上において、故障時間データの巧みなグループ分けによって線形回帰式により $(m_1, \theta_1, m_2, \theta_2)$ を同時に推定する方法を提唱している。この方法も m_1 と m_2 がかなり離れていないと極めて不安定であることが実験的に確かめられている。 $k>2$ の場合の解析例は未だ報告されていない。

最後に、iii)の一部故障原因不明データについてであるが、これについても通常の方法で尤度を構成し、尤度方程式を解こうとすると、ii)の故障原因不明データの場合と同様な数値計算上の困難が生じる。そこで、この場合における数値計算上の工夫を次節で述べる。

3.3 不完全な競合データに対する推定法の導出

3.3.1 問題の定式化

n 個のアイテムを観測し、 n_j 個のモード j による故障時間 x_{j1}, \dots, x_{jn_j} ($j=1, \dots, k$)と、 n_0 個の故障原因不明の故障時間 x_{01}, \dots, x_{0n_0} を得たものとする。ここで、 $n = n_0 + \sum_{j=1}^k n_j$ である。このとき尤度は

$$L \propto \prod_{j=1}^k \prod_{i=1}^{n_j} \left\{ f_j(x_{ji}) \prod_{\substack{l=1 \\ l \neq j}}^k R_l(x_{ji}) \right\} \times \prod_{i=1}^{n_0} f(x_{0i}) \quad (3.11)$$

と書ける。ここで、 $f_j(\cdot)$, $f(\cdot)$ はそれぞれ式(3.1), (3.2)で与えられるものである。このとき式(3.11)のオ1項は、式(3.8)と同様に $\prod_{j=1}^k L_j$ の形に書き改められるが、オ2項の $f(\cdot)$ のために L 全体としてモード j ごとに分離することはできない。そのため L を直接最大化しようとするれば、3.2.2節に述べた ii) 故障原因不明データの場合と同様の数値計算上の困難が生じ、とくに $k > 2$ では最尤推定値を求めることは不可能に近い。

そこでここでは、このような困難を避けるために、EMアルゴリズムを応用した推定アルゴリズムを新たに提案する。EMアルゴリズムは Dempster's^[4] により 1977年提案されたものであるが、その基本的考え方はもっと古くからあったようである。EMアルゴリズムとは、各種不完全データから最尤推定値を求めるための数値計算の技法であり、

- 1) 不完全データと母数の推定値を与えたときの、十分統計量の条件付期待値を求める (Eステップ)
- 2) 十分統計量の条件付期待値から、最大尤度を与える母数の値を求める (Mステップ)

からなり、1), 2) を繰り返すことで最尤推定値を求めるものである。EMアルゴリズムはこれまでにセンサードデータ、とくにセンサードデータも含んだ回帰分析や、欠測値のある多元配置データの解析などに適用されてきた。これらの問題の多くは、対象とする不完全データの尤度が閉じた形で求められない場合もしくは極めて複雑な場合であるが、そのときにも計算機による繰り返し計算で効率的に最尤解が得られている。

EMアルゴリズムにおけるEステップとは、不完全データから形式的に完全な擬似データを発生させるステップと考えることができる。ただしこの発生の方は、条件付期待値というプロセスでなされねばならない。そこで、ここで

の故障原因不明データに適用させるため、次の条件付故障原因確率を定義する。

$$p_j(x) = \Pr \{ \delta = j \mid X = x \} \quad (3.12)$$

これは、アイテムの故障時間が x であるという条件のもとでの、真の故障原因がモード j である条件付確率である。ワイブル分布を仮定した競合モデルではモード j の故障率関数を

$$\lambda_j(x) = \frac{m_j}{\theta_j} \left(\frac{x}{\theta_j} \right)^{m_j-1} \quad (3.13)$$

と記せば ($j = 1, \dots, k$),

$$p_j(x) = \frac{\lambda_j(x)}{\sum_{j=1}^k \lambda_j(x)} \quad (3.14)$$

と書ける。

3.3.2 EMアルゴリズムを応用した推定法

ここで提案する推定法は次の手順から構成される。

1) 初期推定値の作成

n_0 個の故障原因不明データを無視して、 $\sum_{j=1}^k n_i$ 個の完全データより式(3.10)を Newton-Raphson 法で解き、初期推定値 (1回目の推定値) ($m_j^{(1)}$),

$\theta_j^{(1)}$ ($j=1, \dots, k$) を得る.

2) Eステップ

r 回目 ($r \geq 1$) の推定値 ($m_j^{(r)}, \theta_j^{(r)}$) ($j=1, \dots, k$) より, n_0 個の故障原因不明の故障時間 x_{01}, \dots, x_{0n_0} における各モードの故障率関数値をそれぞれ

$$\lambda_j^{(r)}(x_{0i}) = \frac{m_j^{(r)}}{\theta_j^{(r)}} \left(\frac{x_{0i}}{\theta_j^{(r)}} \right)^{m_j^{(r)} - 1} \quad (3.15)$$

$(i=1, \dots, n_0, j=1, \dots, k)$

で推定し, x_{01}, \dots, x_{0n_0} における条件付故障原因確率の推定値

$$p_j^{(r)}(x_{0i}) = \frac{\lambda_j^{(r)}(x_{0i})}{\sum_{j=1}^k \lambda_j^{(r)}(x_{0i})} \quad (3.16)$$

を求める. $p_j^{(r)}(x_{0i})$ は x_{0i} で故障したアイテムの真の故障原因がモード j である確率の推定値であるから, これをモード j が原因である「重み」と考え, 擬似的に $p_j^{(r)}(x_{0i})$ 個のアイテムが x_{0i} でモード j により故障したとみなす. そこで, すべての故障時間 x_{li} ($l=0, 1, \dots, k, i=1, \dots, n_l$) に対して

$$\delta_j^{(r)}(x_{li}) = \begin{cases} 1 & \text{---- } l=j, \text{ すなわちモード } j \text{ が原因のとき} \\ 0 & \text{---- } l \neq j, l \geq 1, \text{ すなわちモード } j \text{ 以外が原因のとき} \\ p_j^{(r)}(x_{li}) & \text{---- } l=0, \text{ すなわち故障原因不明のとき} \end{cases} \quad (3.17)$$

と定義する. $\delta_j^{(r)}(x_{li})$ は, r 回目の母数の推定値に基づき, x_{li} で故障したアイテムの故障原因がモード j である「重み」である.

3) Mステップ

モード j については、擬似的な完全データ $(x_{li}, \delta_j^{(r)}(x_{li}))$ ($l=0, 1, \dots, k, i=1, \dots, n_l$) を基に

$$\sum_{l=0}^k \sum_{i=1}^{n_l} \frac{\delta_j^{(r)}(x_{li})}{m_j} + \sum_{l=0}^k \sum_{i=1}^{n_l} \delta_j^{(r)}(x_{li}) \ln x_{li}$$

$$\frac{\left[\sum_{l=0}^k \sum_{i=1}^{n_l} \delta_j^{(r)}(x_{li}) \right] \cdot \left[\sum_{l=0}^k \sum_{i=1}^{n_l} x_{li}^{m_j} \ln x_{li} \right]}{\sum_{l=0}^k \sum_{i=1}^{n_l} x_{li}^{m_j}} = 0$$

(3.18)

$$\theta_j = \left(\frac{\sum_{l=0}^k \sum_{i=1}^{n_l} x_{li}^{m_j}}{\sum_{l=0}^k \sum_{i=1}^{n_l} \delta_j^{(r)}(x_{li})} \right)^{1/m_j}$$

を解き、この解を $(r+1)$ 回目の推定値 $(m_j^{(r+1)}, \theta_j^{(r+1)})$ ($j=1, \dots, k$) とする。ここで式(3.18)は、式(3.10)に簡単な係数の修正を施すだけでよいので、Newton-Raphson法で容易に解ける。

4) m_j, θ_j ($j=1, \dots, k$) の推定値が所定の精度で収束するまで、2)と3)のステップを繰り返す。

以上のアルゴリズムにより、完全な競合データの場合と同様な数値計算で推定値を求めることができる。指数型分布族を含む広範な分布族では、特殊な場合を除いて、EMアルゴリズムは収束しその収束先は最尤推定値になることが

証明されている^[7]。また、次節のシミュレーション実験においては、この推定アルゴリズムを用いているが、様々な条件のもとで、このアルゴリズムの収束は極めて速いことが実験的に確認された。

3.4 漸近的性質とシミュレーション結果

3.4.1 情報量行列に基づく漸近的評価

競合ワイブル分布における母数推定の評価は、これまでには、松尾と村田^[30]がモンテカルロ・シミュレーションによって行ったものしかない。ここでは、3.2.2節で分類した i) 完全データ、ii) 故障原因不明データの情報量行列を数値的に求め、これの逆行列をとることにより各母数に対する最尤推定量の漸近分散を求める。iii) 一部故障原因不明データの情報は、i)、ii) の情報量の線形結合で与えられるので、とくにここでは求めない。また、小標本に対する評価としては、i)、iii) を中心に、モンテカルロ・シミュレーションで調べた結果を3.4.2節で示す。

まず、i) の完全データにおいては、推定量の形は局外母数に無関係だが、その性質は局外母数の値に依存するため、統一的な評価は困難である。そこで、以下 $k=2$ に限定して数値的に調べることにする。いま、モード1による故障時間を x_{11}, \dots, x_{1n_1} 、モード2による故障時間を x_{21}, \dots, x_{2n_2} とすれば尤度は

$$L = \prod_{j=1}^2 \left(\frac{m_j}{\theta_j} \right)^{n_j} \times \prod_{j=1}^2 \prod_{i=1}^{n_j} \left(\frac{x_{ji}}{\theta_j} \right)^{m_j-1} \times \exp \left\{ - \sum_{j=1}^2 \sum_{i=1}^{n_j} \left(\frac{x_{ji}}{\theta_j} \right)^{m_j} - \sum_{j=1}^2 \sum_{i=1}^{n_j} \left(\frac{x_{ji}}{\theta_j} \right)^{m_j} \right\}$$

となり, 対数尤度を 1 回微分すれば

$$\frac{\partial}{\partial m_1} \ln L = \frac{n_1}{m_1} + \sum_{i=1}^2 \ln x_{1i} - n_1 \ln \theta_1 - \sum_{j=1}^2 \sum_{i=1}^{n_j} \left(\frac{x_{ji}}{\theta_1} \right)^{m_1} \ln \left(\frac{x_{ji}}{\theta_1} \right) \quad (3.19)$$

$$\frac{\partial}{\partial \theta_1} \ln L = -\frac{n_1 m_1}{\theta_1} + \sum_{j=1}^2 \sum_{i=1}^{n_j} \left(\frac{x_{ji}}{\theta_1} \right)^{m_1} \left(-\frac{1}{\theta_1} \right)$$

となる. $(\partial/\partial m_2) \ln L$, $(\partial/\partial \theta_2) \ln L$ も同様だが, ここで注意すべきことは式(3.19)に (m_2, θ_2) が含まれていないことである. すなわち, 情報量行列 (4×4) は, (m_1, θ_1) の部分と (m_2, θ_2) の部分に分離される. これより漸近的には $(\hat{m}_1, \hat{\theta}_1)$ と $(\hat{m}_2, \hat{\theta}_2)$ は独立になることを知る.

情報量行列の要素は, (m_1, θ_1) については, それぞれ

$$E \left[\left(\frac{\partial}{\partial m_1} \ln L \right)^2 \right], E \left[\left(\frac{\partial}{\partial m_1} \ln L \right) \left(\frac{\partial}{\partial \theta_1} \ln L \right) \right], E \left[\left(\frac{\partial}{\partial \theta_1} \ln L \right)^2 \right]$$

で求まるが, ここで x_{1i} と x_{2i} は, それぞれモード 1, モード 2 が故障原因である条件付分布に従うものであり, 分布が異なるのでこのままでは計算は極めて複雑になる. そこで

$$E \left[\left(\frac{\partial}{\partial m_1} \ln L \right)^2 \right] = E \left[-\frac{\partial^2}{\partial m_1^2} \ln L \right]$$

$$E \left[\left(\frac{\partial}{\partial m_1} \ln L \right) \left(\frac{\partial}{\partial \theta_1} \ln L \right) \right] = E \left[-\frac{\partial^2}{\partial m_1 \partial \theta_1} \ln L \right]$$

$$E \left[\left(\frac{\partial}{\partial \theta_1} \ln L \right)^2 \right] = E \left[-\frac{\partial^2}{\partial \theta_1^2} \ln L \right]$$

の性質を利用し, 2 回微分から求めると

$$\begin{aligned} \frac{\partial^2}{\partial m_1^2} \ln L &= -\frac{n_1}{m_1^2} - \sum_{j=1}^2 \sum_{i=1}^{n_j} \left(\frac{x_{ji}}{\theta_1}\right)^{m_1} \left[\ln\left(\frac{x_{ji}}{\theta_1}\right)\right]^2 \\ \frac{\partial^2}{\partial m_1 \partial \theta_1} \ln L &= -\frac{n_1}{\theta_1} + \sum_{j=1}^2 \sum_{i=1}^{n_j} \left[\frac{1}{\theta_1} \left(\frac{x_{ji}}{\theta_1}\right)^{m_1} + \frac{m_1}{\theta_1} \left(\frac{x_{ji}}{\theta_1}\right)^{m_1} \ln\left(\frac{x_{ji}}{\theta_1}\right) \right] \quad (3.20) \\ \frac{\partial^2}{\partial \theta_1^2} \ln L &= \frac{n_1 m_1}{\theta_1^2} - \sum_{j=1}^2 \sum_{i=1}^{n_j} \frac{m_1 (m_1 + 1) x_{ji}^{m_1}}{\theta_1^{m_1 + 2}} \end{aligned}$$

を得るが (m_2, θ_2 についても同様), ここで x_{ji} は式(3.20)において, すべて

$$\sum_{j=1}^2 \sum_{i=1}^{n_j} \varphi(x_{ji})$$

の形で入っていることを知る. ただし $\varphi(\cdot)$ は任意の関数である. 此より, 問題は $n=1$ の場合に帰着し, 式(3.20)の各式を $n=1$ として, 式(3.2)の $f(x)$ で積分すれば, 情報量行列の要素が求まることになる. 式(3.20)の θ_1 式については

$$E \left[-\frac{\partial^2}{\partial m_1^2} \ln L \right] = \frac{1}{m_1^2} \Pr\{X_1 < X_2\} + \int_0^{\infty} \left(\frac{x}{\theta_1}\right)^{m_1} \left[\ln\left(\frac{x}{\theta_1}\right)\right]^2 f(x) dx$$

で求まる (他も同様). また, 局外母数の影響は $f(x)$ の中に表われることになる.

一方, ii) の故障原因不明データに対しては, 式(3.2)の $f(x)$ を基に, スコアベクトル

$$S = \left(\frac{\partial}{\partial m_1} \ln L, \frac{\partial}{\partial \theta_1} \ln L, \frac{\partial}{\partial m_2} \ln L, \frac{\partial}{\partial \theta_2} \ln L \right)'$$

に対して SS の要素を直接 $f(x)$ で数値積分した。

さらに、情報量行列より漸近分散共分散行列を求め、その対角成分である漸近分散を算出した。その結果の一部を表3.1に示す。これよりわかることは、 $m_1 = 1.0$, $m_2 > 1.0$ の場合にも、ii) 故障原因不明データに基づく推定量の漸近分散は、i) 完全データに基づく推定量の漸近分散の10倍程度であり、 $m_1, m_2 > 1.0$ の場合には、 $10^2 \sim 10^3$ 倍と極めて大きな差を知ることができる。したがって、ii) の故障原因不明データから母数推定を行うことは極めて危険なことと言える。表3.1のケース3は、松尾と村田^[30] がシミュレーション実験を行ったときの母数の設定値であり、彼らのシミュレーション結果とここでの漸近分散を比較すると、小標本での漸近分散に対する効率においても、ii) は i) に比べ、約 $1/40 \sim 1/90$ と極めて悪いことがわかる。なお、次節のシミュレーション実験でも、この母数の設定値を用いている。

表3.1 $\hat{m}_1, \hat{\theta}_1, \hat{m}_2, \hat{\theta}_2$ の漸近分散

ケース	$A. Var(\hat{m}_1)$	$A. Var(\hat{\theta}_1)$	$A. Var(\hat{m}_2)$	$A. Var(\hat{\theta}_2)$
$m_1=1.0, m_2=2.0$	1.317	1.988	4.602	5.518×10^{-1}
$\theta_1=\theta_2=1.0$	5.599	2.559×10	2.559×10	4.485
$m_1=2.0, m_2=5.0$	4.895×10^{-1}	5.346	2.914×10	9.059×10^{-2}
$\theta_1=\theta_2=1.0$	7.029×10^2	1.399×10^3	3.833×10^3	5.523×10
$m_1=21.3, m_2=22.5$	5.597×10^2	8.840×10^2	6.609×10^2	2.447×10^3
$\theta_1=163.3$	1.805×10^4	8.111×10^5	1.228×10^4	9.382×10^5
$\theta_2=186.0$				

上段が i) 完全データでの値

下段が ii) 故障原因不明データでの値

3.4.2 シミュレーション結果と考察

松尾と村田^[30]は、松尾と木村ら^[29]の行ったステアタイト・セラミックスの4点曲げ試験および松尾らの実施した3点曲げ試験のデータより、ステアタイト・セラミックスがもつ内部き裂、表面き裂という2つの故障モードの母数の推定を行っている。このとき、試験片を観測しても故障原因（破壊源）がはっきりしないものが少なくないという理由から、故障原因データを無視し、ii)の故障原因不明データとして、これより *Nelder-Mead* 法（シンプレックス法）で推定値を求めている。同時に、すべての故障原因が観測された完全データとの比較を目的として、得られた推定値を設定値としたシミュレーション実験も行っている。

ここでは、上記実験に対して、新たに iii) の一部故障原因不明データを加え、故障原因不明データの割合が推定精度に及ぼす影響を調べることを主目的としたモンテカルロ・シミュレーション実験を行った。

設定した母数の値は、[30]と同様に

$$\begin{aligned} m_1 &= 21.3 & m_2 &= 22.5 \\ \theta_1 &= 163.3 & \theta_2 &= 186.0 \end{aligned}$$

とした。ここで、モード1として内部き裂、モード2として表面き裂をそれぞれ想定している。実験の因子としては、サンプルサイズ（10, 20, 30, 40, 50の5水準）と、故障原因不明データの割合（0～70%まで10%ごとの8水準）の2つを取り上げ、各実験で500回の繰り返しを行った。

得られた結果は以下の通りである。図3.2は、故障原因不明データの割合が0%、すなわち完全データのときの、サンプルサイズと推定値の平均および標

標準偏差との関係を示したものであるが、サンプルサイズが10から20へ増えるとき、大きく改善されることがわかる。また、図3.3は、故障原因不明データの割合と推定値の平均および標準偏差との関係を示したもののだが、故障原因不明データの割合が40%ぐらいまでは、0%（完全データ）の場合とほとんど変わらないことがわかる。図3.4は、平均2乗誤差との関係を示しているが、これらは第2章で与えた指数分布の場合と類似した傾向をみせている。また、松尾と村田のシミュレーション結果と比較すれば、割合100%、すなわちii)の故障原因不明データに対し、30%の故障原因データが加わると、 $n=20$ では、約1/1000になるという顕著な結果が得られた。

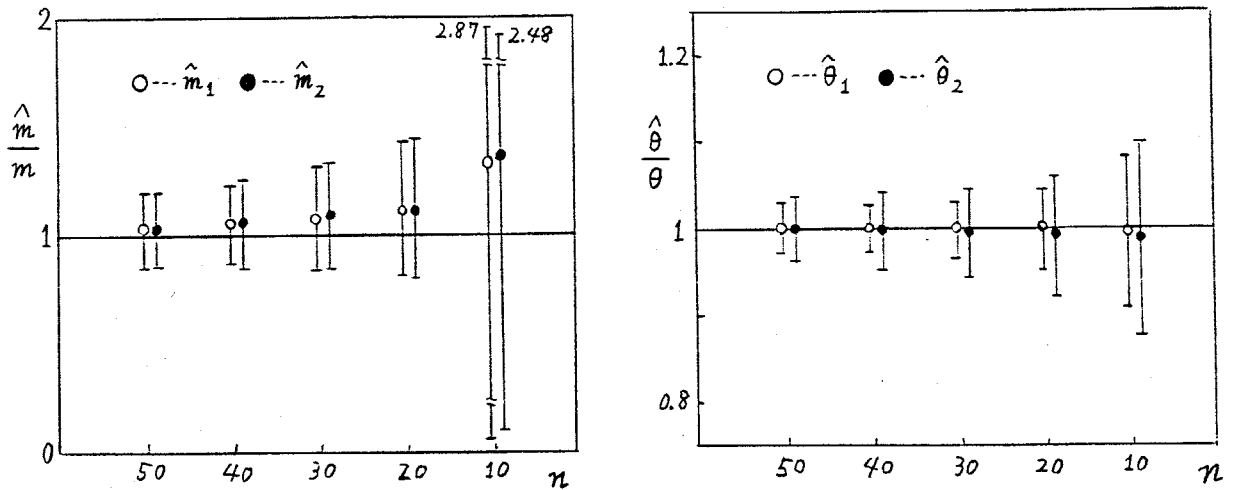


図3.2 サンプルサイズと推定値の平均、標準偏差の関係

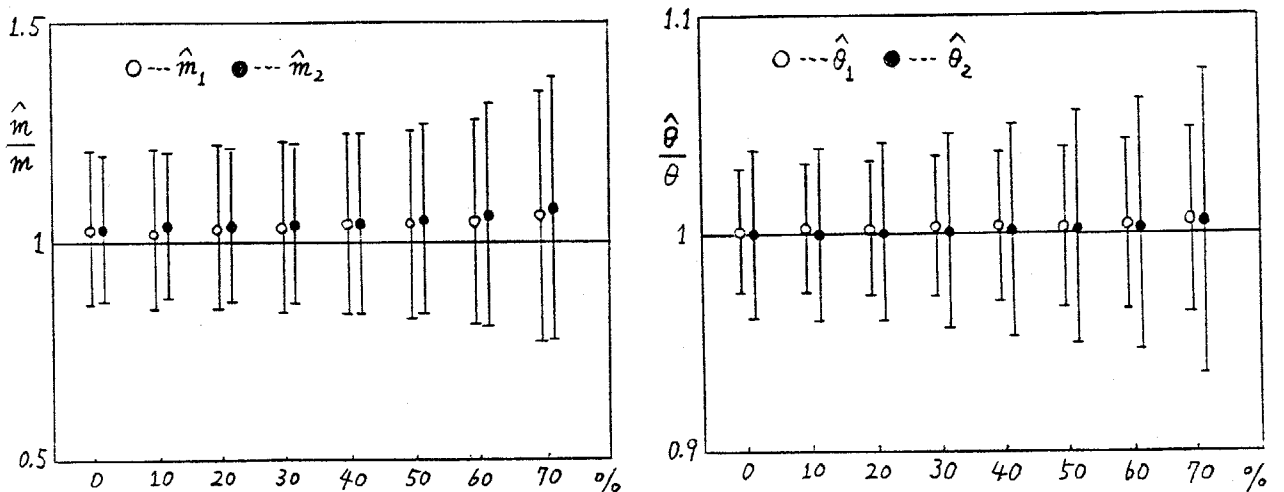


図3.3 故障原因不明データの割合と推定値の平均、標準偏差の関係 (n=50)

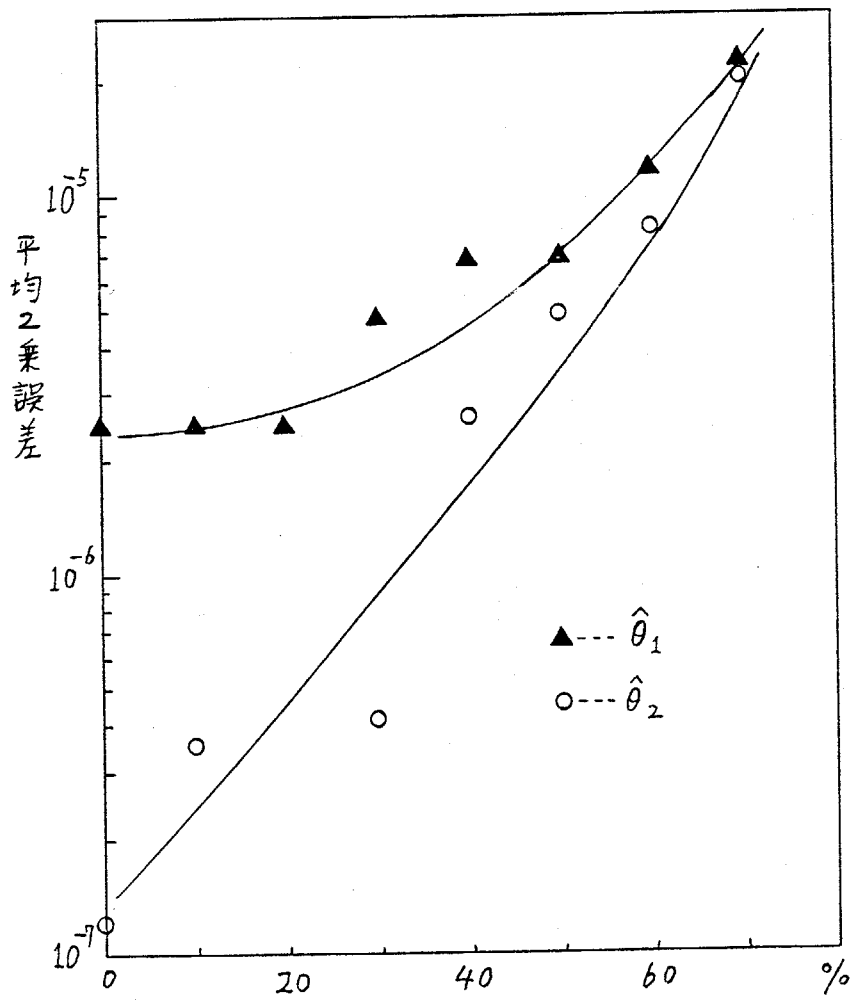
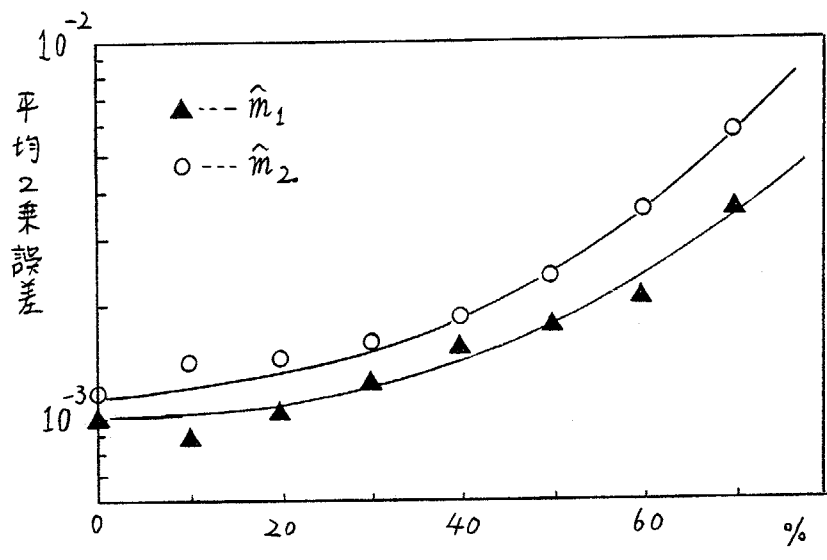


図 3.4 故障原因不明データの割合と平均二乗誤差 ($n=50$)

3.5 結言

本章では、ワイブル分布を仮定した競合モデルについて考察し、長モード競合ワイブル分布の基本的性質を明らかにするとともに、故障原因データの有無により3種類の競合データを考え、それらに基づく母数推定法とその推定精度について考察した。ここで得られた主な結果は次の通りである[39][40]。

- 1) 長モード競合ワイブル分布がワイブル確率紙上で下に凸の曲線になることを証明した。この性質により、寿命データから複数故障モードの競合性を検知するための指針を得た。
- 2) 故障原因不明データを一部に含む不完全な競合データに基づく効率的な母数推定法として、EMアルゴリズムを応用した新たな推定法を導いた。これにより完全データの場合とほぼ同様の数値計算で推定値を求めることが可能となった。
- 3) 故障原因不明データの有無による母数推定の推定精度の違いを、情報量行列による漸近的評価ならびにシミュレーション実験から検討した。その結果から、母数推定に際しては、ある程度の故障原因データを収集することが不可欠であることが明らかになった。

第4章 ノンパラメトリックな競合モデル

4.1 緒言

一般に、フィールドにおける寿命データは、寿命試験データに比べその「量」は豊富であるが、「質」については劣るのが常であるから、特定の分布形やモデルを仮定した上での母数推定よりも、確率紙へのプロットなどの図的表示を実施し、その結果から必要に応じて外れ値の検出、使用環境や使用れ方による層別、さらには複数の故障モードに対する競合モデルのあてはめなどの前処理が先決となる。たとえばワイブル確率紙は、第7章で述べるように母数推定の手段としても有用であるが、それ以上に寿命データの背後に潜む構造を探索する上で重要な役割を果たすものである。ワイブル確率紙に打点する際、ワイブル分布に従うことを仮定していても、縦軸の値として分布関数のノンパラメトリックな推定値が必要となる。フィールドデータでは、様々なタイプの不完全データが得られるので、これら不完全データからノンパラメトリックに分布関数を推定する方法が必要不可欠となる。また今日、累積ハザード関数や信頼度関数よりも敏感な特性である故障率関数もしくはハザード関数のノンパラメトリック推定の活用が重視されている。この背景には、従来の故障率関数に関する従来の単調性を仮定したパラメトリックモデルや比例ハザードモデルに対する見直しがある。また同時に、ノンパラメトリック推定を可能にするだけのデータ数を集める体制が整ってきたことも挙げられよう。

競合モデルの適用においても、観測対象であるシステムの寿命データから各コンポーネントもしくは故障モードの寿命特性をノンパラメトリックに推定すべき場面が増加している。本章では、第2章、第3章で述べたパラメトリックな方法に対する前処理的手段として、競合モデルにおけるノンパラメトリックな統計的方法の問題を取り上げる。とくに、前章までに取り扱った故障原因不明データを含む場合については、ノンパラメトリック推定固有の問題点を指摘し、付加的な仮定のもとにいくつかの方法を提案し、その性質について考察する。

4.2 モデルの背景と定式化

4.2.1 完全な競合データに対する方法

ここでは簡単のため、モード数を $k=2$ とする。モード1の寿命を X_1 、その分布関数を $F_1(t) = Pr(X_1 \leq t)$ で表わし、モード2の寿命を X_2 、その分布関数を $F_2(t) = Pr(X_2 \leq t)$ で表わす。そして X_1 と X_2 は独立とする。またモード1、モード2の信頼度関数をそれぞれ $R_1(t) = 1 - F_1(t)$ 、 $R_2(t) = 1 - F_2(t)$ で表わす。

まず、すべての観測対象であるシステムに対して故障時間と故障原因が観測される完全な競合データについて考える。このとき、 n 個の同種システムを観測し、システムとしての中途打切りはないものとする。また、 X_1 と X_2 は本来連続な確率変数と考えられるから、独立性の仮定より $Pr(X_1 = X_2) = 0$ とする。ただし、フィールドで観測される故障時間は完全な連続量ではなく、月、

週, 日, 時間といった単位をもつので, データとしてのタイは存在するものとする. 第2章, 第3章ではタイについてとくにことわらなかつた. それは, パラメトリック推定においてはタイを処理する上で特別な問題が生じないからである. ところが, ノンパラメトリック推定では後述するようにタイを処理する上でいくつかの問題が生じる. そこで, ここでは「タイの存在を許す」ことを明記する.

このような仮定のもとでは, 得られるデータは式(2.4)と同じく

$$X_i = \min(X_{1i}, X_{2i})$$

$$S_i = \begin{cases} 1 & \text{--- } X_i = X_{1i} \text{ のとき} \\ 2 & \text{--- } X_i = X_{2i} \text{ のとき} \end{cases}$$

となる. ここで X はシステムの故障時間であり, その信頼度関数は仮定より, $R(t) = R_1(t) \cdot R_2(t)$ である. また S は故障原因を示している. ここで, モード1に注目するならばこのモデルは X_2 も打切り変数とするランダムセンサリングモデルとみなせ, $R_1(t)$ のノンパラメトリック推定にはカプラン・マイヤー推定量(以下, K-M推定量と記す)が用いられる. いま, X_i ($i=1, \dots, n$) を大きさの順に並べたものを $t_1 < t_2 < \dots < t_{n'}$ と表わし(タイの存在を許すので $n' \leq n$ である), t_i における $S=1$ の故障数を d_{1i} , $S=2$ の故障数を d_{2i} とする. さらに $n_i = \sum_{l=i}^{n'} (d_{1l} + d_{2l})$ と定義したときに (n_i は, t_i の直前での残存数を意味する), K-M推定量とは

$$\hat{R}_1(t) = \prod_{t_i \leq t} \left(\frac{n_i - d_{1i}}{n_i} \right) \quad (4.1)$$

で与えられるものである. なお式(4.1)は $d_{1n'} = 0$ のときには, $t_{n'} < t$ なる t では定義されない. 式(4.1)の性質としては

- 1) ノンパラメトリック下での最尤推定量である [22].
- 2) 強一致性をもつ [43].
- 3) 正規過程に収束する [8].

などがわがっている。また, Efron^[17]は "self-consistent" 推定量という別な観点から K-M 推定量を導き, 「K-M 推定量とは, 故障時間と打ち切り時間を合わせて小さい順に並べたとき, 打ち切りデータ ($\delta = 2$) の重みも後続するデータへ均等配分した後の経験分布 (empirical distribution) である」という「再配分 (redistribution)」による解釈を与えている。「再配分」の概念は, 3.3 節で用いた EM アルゴリズムにおける擬似データの発生にも通ずるものであり, 本章でもこれを応用している。

完全な競合データに基づく信頼度関数のノンパラメトリック推定法としては, K-M 推定量の他に, Nelson の累積ハザード法, Johnson の平均順位法, Herd の方法があるが, K-M 推定量が任意の時点での信頼度を推定することも意図しているのに対し, これらは故障時点での信頼度の値を推定するものである。K-M 推定量を含め, これらの方法はいずれも「再配分」の概念から統一的に解釈される^[38]。

4.2.2 故障原因不明データを含む場合の問題点

次に, 故障原因不明データを含んだ不完全な競合データを考える。このとき前章までと同様に, 「故障原因が観測されるか否かは, そのシステムの故障時間や真の故障原因に無関係である」と仮定する。また, 全データ数 n と故障原因が観測されるデータ数 m を定数とすれば, 得られるデータは

$$X_i = \min (X_{1i}, X_{2i})$$

$$\delta_i = \begin{cases} 1 & \text{--- } X_i = X_{1i} \text{ のとき} \\ 2 & \text{--- } X_i = X_{2i} \text{ のとき} \end{cases}$$

$$(i = 1, \dots, m)$$

と故障原因不明データ

$$X_i = \min (X_{1i}, X_{2i})$$

$$\delta_i = 0 \quad \text{--- 故障原因不明のとき}$$

$$(i = m+1, \dots, n)$$

となり, 図 2.4 に示したものになる.

そこでまず, このようなデータが得られたときの尤度を構成してみる. ここでシステムの分布関数を $F(t) = 1 - R(t)$ と記す. また, 第 3 章で定義したように, システムが時刻 t で故障したという条件のもとでの故障原因がモード j である確率も

$$p_j(t) = \Pr \{ X = X_j \mid X = t \}$$

で表わす ($j=1, 2$). すると $\{X=t, \delta=1\}$ の尤度への寄与は

$$p_1(t) dF(t) = (dF_1(t)) R_2(t)$$

となる. ここで $dF(t)$ は連続分布の場合は $F'(t)$ を意味し, 離散分布の場合は $F(t) - F(t-0)$ を意味する. またここで

$$F(t-0) = \lim_{\varepsilon \rightarrow 0} F(t-\varepsilon)$$

である.

同様に, $\{X=t, \delta=2\}$ の尤度への寄与は

$$p_2(t) dF(t) = (dF_2(t)) R_1(t)$$

となる。さらに、故障原因不明データ $\{X=t, \delta=0\}$ の尤度への寄与は $dF(t)$ となる。よって n 個のデータ全体の尤度を $p_1(t), p_2(t), F(t)$ で表わすと

$$L = \prod_{i=1}^m p_1(x_i)^{I[\delta_i=1]} p_2(x_i)^{I[\delta_i=2]} \times \prod_{i=1}^n dF(x_i)$$

となり、 $p_1(t), p_2(t)$ に関する部分と $F(t)$ に関する部分に分離されることがわかる。ここで、 $I[\cdot]$ は事象 $[\cdot]$ の特性関数である。これより、たとえば $p_1(t)$ については、 $p_1(t) + p_2(t) = 1$ であるから、故障原因が観測された故障時点 x_i ($i=1, \dots, m$) についてのみ推定可能で、最尤推定量は

$$\hat{p}_1(x_i) = \frac{\sum_{l=1}^m I[x_l = x_i, \delta_l = 1]}{\sum_{l=1}^m I[x_l = x_i]} \quad (4.2)$$

で与えられ、 $F(t)$ については通常の経験分布関数が最尤推定量として導出されることがわかる。

ところで、 $F_1(t), F_2(t)$ と $p_1(t), p_2(t), F(t)$ の間には、 X_1 と X_2 が独立という仮定のもとでは、連続分布の場合には

$$F_1(t) = 1 - \exp \left\{ - \int_0^t p_1(s) \frac{dF(s)}{1-F(s)} \right\}$$

の関係、離散分布の場合には

$$F_1(t) = 1 - \exp \left\{ \sum_{\substack{s \leq t \\ s: \text{不連続点}}} p_1(s) \ln \frac{1-F(s)}{1-F(s-0)} \right\} \quad (4.3)$$

の関係があるわけだが ($F_2(t)$ についても同様), 分布形に何の制約も置かない完全なノンパラメトリック推定では, 上述のとおり $p_1(t)$ は故障原因が観測されたデータの故障時点についてのみ推定可能であり, 故障原因不明データの故障時点については同時点で故障原因の観測された故障がないかぎり推定不可能である. したがって $p_1(t), F(t)$ の最尤推定量から式(4.3)の関数関係を基に, $F_1(t), F_2(t)$ のノンパラメトリックな最尤推定量を導出することはできないことがわかる.

そこで次に, ノンパラメトリック下での十分統計量に基づくヒューリスティックな推定量を考えてみる. いま, X_1, X_2 の "subdistribution function" $F_1^*(t), F_2^*(t)$ をそれぞれ

$$F_1^*(t) = P_r \{ X \leq t, S = 1 \}$$

$$F_2^*(t) = P_r \{ X \leq t, S = 2 \}$$

と定義すれば, $F(t) = F_1^*(t) + F_2^*(t)$ を満たす. また式(4.3)を書き改めれば

$$R_1(t) = \prod_{\substack{s \leq t \\ s: F_1 \text{ の不連続点}}} \frac{R_1(s) R_2(s)}{R_1(s-0) R_2(s-0)}$$

となるが, ここで, 分子だけに "subdistribution function" を適用すれば

$$R_1(t) = \prod_{\substack{s \leq t \\ s: F_1 \text{ の不連続点}}} \frac{R(s-0) - (F_1^*(s) - F_1^*(s-0))}{R(s-0)} \quad (4.4)$$

を得る.

そこで, $R(\cdot), F_1^*(\cdot)$ の推定量として, それぞれ経験分布に基づく

$$R_e(t) = \sum_{i=1}^n \frac{1}{n} I [X_i > t]$$

$$F_{1,e}^*(t) = \sum_{i=1}^m \frac{1}{m} I [X_i \leq t, \delta_i = 1]$$

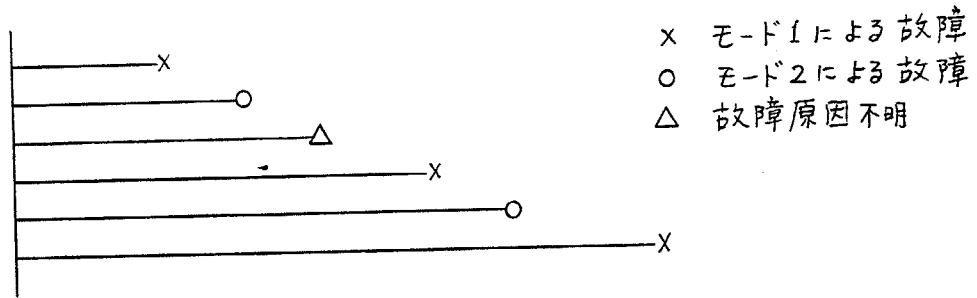
を用い、これを式(4.4)に代入した

$$R_{1,e}(t) = \prod_{\substack{s \leq t \\ s: F_{1,e}^* \text{の不連続点}}} \frac{R_e(s-0) - (F_{1,e}^*(s) - F_{1,e}^*(s-0))}{R_e(s-0)} \quad (4.5)$$

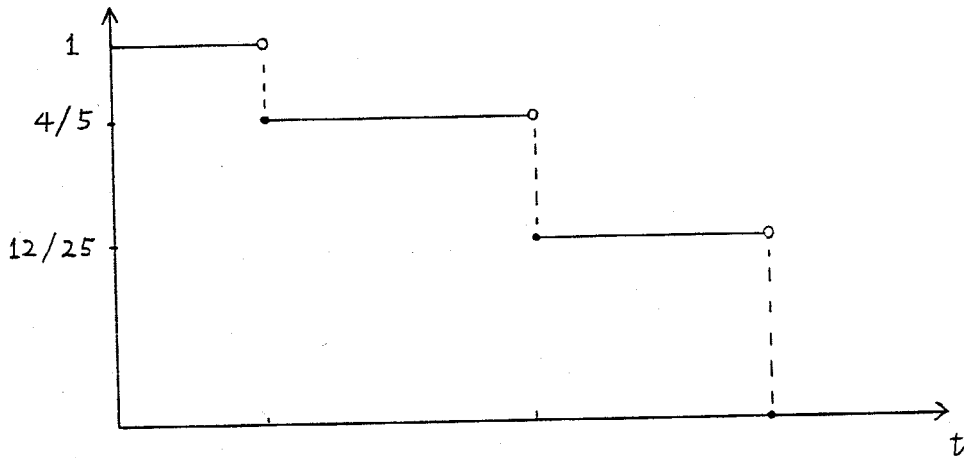
を考察することができる。この $R_{1,e}(t)$ は、システムの信頼度関数については n 個の故障時間より推定し、"subdistribution function" $F_{1,e}^*(t)$ については故障原因の観測された m 個のデータだけから推定するというある意味で合理的な推定量である。また $R_1(t)$ は式(4.4)において $R(\cdot)$ と $F_{1,e}^*(\cdot)$ の連続関数であり、式(4.5)の $R_{1,e}(t)$ はこれらの経験分布に基づく推定量であるから、Glivenko-Cantelli の定理より、 $n \rightarrow \infty, m \rightarrow \infty$ で真の $R_1(t)$ に確率1で収束する強一致推定量であることも知る。

しかしながら、 $R_{1,e}(t)$ は、本研究の目的からすれば、故障原因不明データを無視した場合に比べて本質的な改善になっていない。図4.1に $R_{1,e}(t)$ の計算例を示してある。これよりわかるように、 $R_{1,e}(t)$ はもちろん故障原因不明データを無視した場合の推定量とは異なるが、依然としてモード1による故障時点だけで不連続となる階段関数である。したがって、確率紙にプロットする際には、モード1による故障時点だけが打点の位置になってしまう。

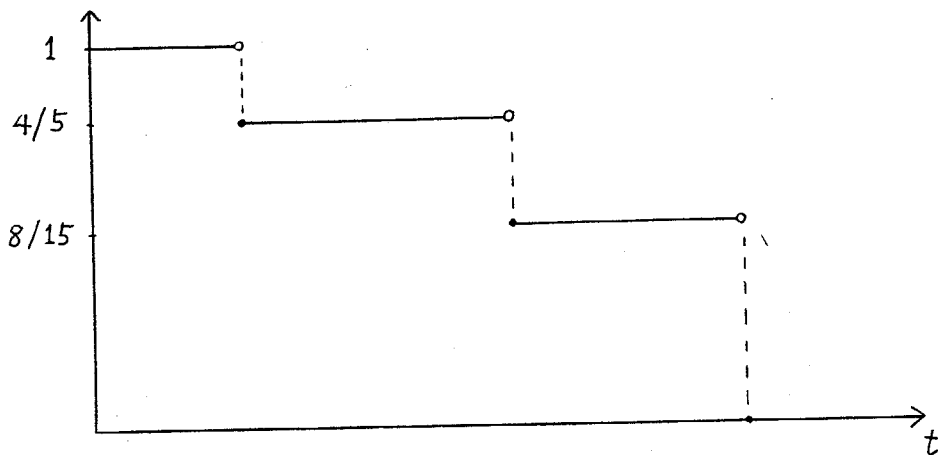
この問題の背景として、市場に出す製品のうち無作為に選んだ何台かにモニターを付け、これらについては故障時間と故障原因が報告され、残りについては故障時間だけが報告されるような状況を考えているが、このような場合モニ



(a) 故障原因不明を含んだ競合データ



(b) 式(4.5)による $R_{1,e}(t)$ の形



(c) 故障原因不明データを除いて構成した K-M 推定量

図 4.1 $R_{1,e}(t)$ の構成例

ターの割合は製品によって異なるが、高々1割~3割と思われる。このとき、わずかに得られているモード1による故障時点のみで打点することは、分布全体の傾向を知る上で極めて不都合と言える。

完全なノンパラメトリックな推定問題において、このような現象が起こるのは、得られた全データが、故障原因不明データに対してその故障時点における真の故障原因に関する情報を何らもっていないためである。しかしながら、寿命というものは本来連続な変量であり、特定の分布形を仮定できなくともその分布関数が連続であると想定することは自然であるから、データを故障時間の大きさの順に並べたとき1つの故障原因不明データの前後にモード1による故障データが多ければ、その故障原因不明データの真の故障原因がモード1である可能性が高いと推測することは極めて現実的であろう。

そこで次節では、このような状況を踏まえて、 $F_1(t)$, $F_2(t)$ が連続でしかもある程度なめらかな故障率関数 $\lambda_1(t)$, $\lambda_2(t)$ をもつと仮定する。この付加的仮定のもとで、確率紙へのプロットなどを目的として、 $R_1(t)$, $R_2(t)$ の推定量をノンパラメトリックに階段関数として構成する方法について考察する。

4.3 不完全な競合データに対する推定法の導出

4.3.1 区間分けされたデータの場合

フィールドにおいては、寿命が連続的に観測されず、たとえば1ヶ月ごとに調査をしてその間の故障数や打切り数が報告されるという場合が多い。このようなデータを区間分けされたデータと呼ぶ。医学統計の分野では生命表 (life

表4.1 故障原因不明データを含む生命表データ

区間	I_1	I_2	I_3	-----	I_N
モード1による故障数	d_{11}	$-d_{12}$	d_{13}	-----	d_{1N}
モード2による故障数	d_{21}	d_{22}	d_{23}	-----	d_{2N}
故障原因不明の故障数	d_{01}	d_{02}	d_{03}	-----	d_{0N}

table) データと呼ばれるものである。 $k=2$ の競合モデルにおいて、故障原因不明データを含んだ生命表データは表4.1のようなものになる。すなわち、予め定められた時点 $\xi_0=0 < \xi_1 < \xi_2 < \dots < \xi_N = \infty$ に対して、区間 $I_i = (\xi_{i-1}, \xi_i]$ ($i=1, \dots, N$) を定め、区間 I_i 内に発生したモード1による故障数、モード2による故障数および故障原因不明の故障数をそれぞれ d_{1i} , d_{2i} および d_{0i} と記す。

このとき、近似的に各区間内で条件付故障原因確率 $p_j(t)$ が一定とし、これを p_{ji} と記せば、 p_{ji} の自然な推定量として、3.3節の式(3.16)もしくは式(4.2)にならう

$$\hat{p}_{ji} = \frac{d_{ji}}{d_{1i} + d_{2i}} \quad (j=1, 2, i=1, \dots, N) \quad (4.6)$$

を考えることができる。すなわち、故障原因が観測されたデータから p_{ji} を推定するわけである。そこで、区間 I_i における d_{0i} 個の故障原因不明データを、それぞれ $d_{0i} \hat{p}_{1i}$ 個のモード1による擬似故障データと $d_{0i} \hat{p}_{2i}$ 個のモード2による擬似故障データとに置き換える。これより、区間 I_i でのモード j のハザード値 h_{ji} のノンパラメトリック推定量として

$$\hat{h}_{ji} = \frac{d_{ji} + d_{0i} \hat{p}_{ji}}{n_i} \quad (4.7)$$

を得る。こゝに $n_i = \sum_{l=i}^N \sum_{j=0}^2 d_{jl}$ で、区間 I_i の初めでの残存数を意味している。式(4.7)より、区間 I_i の終りでのモード j の累積ハザード値 H_{ji} のノンパラメトリック推定値として

$$\hat{H}_{ji} = \sum_{l=1}^i \hat{h}_{jl} \quad (4.8)$$

も導くことができる。式(4.8)により累積ハザード紙へのプロットが可能になる。さらに、信頼度関数のノンパラメトリック推定量としては、K-M推定量に基づき

$$\hat{R}_j(t) = \prod_{i: \xi_i \leq t} \left(\frac{n_i - d_{ji} - d_{0i} \hat{p}_{ji}}{n_i} \right) \quad (4.9)$$

を得る。

4.3.2 区間分けされていないデータの場合

寿命が連続的に観測されていても、その数が十分大きければ区間分けして解析することも考えられる。一方、それほどデータの数が得られないときには、個々の故障時間に基づき、なるべくなめらかに信頼度関数を推定することが肝要と言える。ここでは、3.3節と同様にEMアルゴリズムを応用した推定法を

提示するが、「 $F_1(t), F_2(t)$ が連続である程度なめらかな故障率関数 $\lambda_1(t), \lambda_2(t)$ をもつ」ことを仮定した上でのノンパラメトリック推定であるため、問題そのものが離散分布と連続分布の両方の側面をもっている。そのため、EM アルゴリズムにおける E ステップ (擬似データの発生) では連続分布を前提とし、M ステップ (尤度最大化) では離散分布を前提とすることになり、対象とするモデルがステップごとに異なることになる。これは言い換えれば、付加的仮定のもとでのノンパラメトリック推定において、この付加的仮定も E ステップの中で利用するということである。また E ステップにおいては、上記仮定の性質上 $\lambda_1(t), \lambda_2(t)$ の推定には任意性が生じてしまう。以下、推定アルゴリズムを 2 つのステップに分けて述べる。

1) $p_j(t)$ の推定値が得られたもとでの $R_j(t)$ の推定

4.2.2 節に述べたような故障原因不明データを含んだ競合データに対し、故障時間 x_i ($i=1, \dots, n$) を大きさの順に並べたものを $t_1 < t_2 < \dots < t_{n'}$ ($n' \leq n$) と表わし、 t_i におけるモード 1 の故障数を d_{1i} 、モード 2 の故障数を d_{2i} 、故障原因不明の故障数を d_{0i} と記す。さらに $d_i = d_{1i} + d_{2i} + d_{0i}$ とし、 $n_i = \sum_{l=1}^{n'} d_l$ とする。

いま、 $x = t_i$ なる 1 つの故障原因不明データに対して $p_j(t_i)$ の推定値 $p_j^*(t_i)$ が得られたものとする。このとき、4.3.1 節と同様に、 t_i での 1 つの故障原因不明データを $p_1^*(t_i)$ の重みのモード 1 による故障と $p_2^*(t_i)$ の重みのモード 2 による故障とに置き換えることを考える。すなわち、 t_i でのモード 1、モード 2 の擬似故障データも、条件付故障原因確率の推定値もこれらの重みとして発生させるわけである。このとき、K-M 推定量の拡張形として

$$\hat{R}_j(t) = \prod_{t_i \leq t} \left(\frac{n_i - d_{ji} - d_{oi} p_j^*(t_i)}{n_i} \right) \quad (4.10)$$

(j = 1, 2)

が擬似故障データを所与としたときの最尤推定量として導かれる。ただし式(4.10)は、 $\hat{R}_1(t) \cdot \hat{R}_2(t)$ がシステムの信頼度関数 $R(t)$ の経験分布と一致しないという不自然さをもつ。これは同一時点においてモード1の故障とモード2の故障がタイとして観測されたとき、もしくは擬似データとして発生されたとき、式(4.10)で $\hat{R}_1(t)$ においてはモード1の故障がモード2の故障の直前に生じたとみなし、逆に $\hat{R}_2(t)$ においてはモード2の故障がモード1の故障の直前に生じたとみなすためである。この不自然さは異なる故障モードによる故障がタイとして観測された場合に、K-M推定量自身をもつものである。

また一方、離散分布においては式(4.3)の関係が成り立つので、離散分布におけるシステムのハザード関数

$$-\ln [R(s)/R(s-0)]$$

はモード1, モード2それぞれ $p_1(s)$, $p_2(s)$ の重みで割り振られるものである。したがって、 $p_1(s)$, $p_2(s)$, $R(s)$ の推定値を式(4.3)に代入すれば $R_1(t)$, $R_2(t)$ の推定値が得られるという考え方ができる。この考え方は

$$\tilde{R}_j(t) = \prod_{t_i \leq t} \left(\frac{n_i - d_i}{n_i} \right)^{\frac{d_{ji} + d_{oi} p_j^*(t_i)}{d_i}} \quad (4.11)$$

(j = 1, 2)

が導かれる。式(4.11)では、 $\tilde{R}_1(t) \cdot \tilde{R}_2(t)$ がシステムの信頼度関数 $R(t)$ の経験分布と一致する。またここで注意すべきことは、故障原因不明データが

ないときにも，同一時点でモード1の故障とモード2の故障がタイとして観測された場合には，式(4.11)はK-M推定量と一致しないということである。

なお，式(4.10)，(4.11)のいずれを用いる場合にも， $p_j(t)$ の推定値の得られていない初期状態においては，3.3.2節に示した推定アルゴリズムと同様に，故障原因不明データを除いて構成したK-M推定量も初期推定値として次のステップに進めばよい。ただし，次に述べるように $p_j(t)$ の推定法によってはこのような操作は不要になる。

また，確率紙上にプロットする際に必要となる故障時点 t_i ($i=1, \dots, n'$)での信頼度関数のノンパラメトリック推定値としては，式(4.10)，(4.11)もそれぞれ修正した

$$\hat{R}_j^*(t_i) = \prod_{t_\ell \leq t_i} \left(\frac{n_\ell - d_{j\ell} - d_{0\ell} p_j^*(t_\ell) + 1}{n_\ell + 1} \right)$$

$$\tilde{R}_j^*(t_i) = \prod_{t_\ell \leq t_i} \left(\frac{n_\ell - d_\ell + 1}{n_\ell + 1} \right)^{\frac{d_{j\ell} + d_{0\ell} p_j^*(t_\ell)}{d_\ell}} \quad (j=1, 2)$$

が考えられる。これはHerdの推定量に基づくもので，完全データの場合の平均ランク $(n-i+1)/(n+1)$ をランダムセンシングされたデータへ拡張したものとみなせる。

2) $R_j(t)$ の推定値が得られたもとでの $p_j(t)$ の推定

$R_1(t)$ ， $R_2(t)$ の推定値が得られれば，連続である程度なめらかな故障率関数をもつという条件のもとでは，何らかの方法で $\lambda_1(t)$ ， $\lambda_2(t)$ を推定することが可能になり，これに基づき $p_1(t)$ ， $p_2(t)$ が推定される。ただし先に述べ

たように、この条件下での故障率関数の最尤推定量はすべての t に対しては一
意に定まらない。一般には、ある時点 t での故障率の推定においては t の前後
 $k(n)$ 個の故障データが用いられる。ここで $k(n)$ は n の適当な非減少関数で
ある。

いま、1つの故障原因不明データ ($x = t_i, s = 0$) に対してその故障時点
 t_i におけるモード j の故障率を推定するとき

- i) $l \leq i$ なるいくつかの $(t_l, d_{jl} > 0)$ を用いる
- ii) $l \geq i$ なるいくつかの $(t_l, d_{jl} > 0)$ を用いる
- iii) いくつかの $(t_l, d_{jl} > 0)$ を用いる

の3通りがあるが、ここで ii), iii) を用いた場合には、 $R_j(t)$ および $p_j(t)$ の
推定値が収束するまで「 $p_j(t)$ の推定値が得られたもとでの $R_j(t)$ の推定」と
「 $R_j(t)$ の推定値が得られたもとでの $p_j(t)$ の推定」の2つのステップを計算
機により繰り返し計算しなければならない。これは3.3.2節にみたように、不
完全データから最尤推定値を求める多くの問題で、EMアルゴリズムを用いる
際に繰り返し計算が必要になるのと同じ状況である。ところが i) を用いると
このような繰り返し計算は不要になる。 $R_j(t)$ の推定値を $R_j^*(t)$ で表わせば ($j = 1, 2$), $d_{1i} = d_{2i} = 0$ の場合に、たとえば最も簡単なものとして直前の
故障データだけを用いた $\lambda_j(t_i)$ の推定値は

$$\hat{\lambda}_j(t_i) = \frac{R_j^*(t_{-ji} - 0) - R_j^*(t_{-ji})}{R_j^*(t_{-ji}) (t_i - t_{-ji})} \quad (j = 1, 2) \quad (4.12)$$

ここで t_{-ji} は $l < i$, $d_{jl} > 0$ なる t_l の最大値
で与えられる。ただし、それぞれ対応する t_{-1i} , t_{-2i} が無いときには、0と
おく。これより $p_j(t_i)$ の推定値は

$$\hat{p}_j(t_i) = \frac{\hat{\lambda}_j(t_i)}{\hat{\lambda}_1(t_i) + \hat{\lambda}_2(t_i)} \quad (j=1, 2) \quad (4.13)$$

で与えられる。なお、 $\hat{\lambda}_1(t_i) = \hat{\lambda}_2(t_i) = 0$ のときには、 $\hat{p}_j(t_i) = 1/2$ とおく。また $d_{1i} > 0$ もしくは $d_{2i} > 0$ の場合には $\hat{p}_j(t_i)$ は式(4.2)で与えられるものとする。このように $\hat{p}_j(t_i)$ を定義すれば、 $R_j^*(t)$ として 1) で提案した 2組の推定量 $\hat{R}_j(t)$ と $\tilde{R}_j(t)$ のいずれを用いた場合にも、 $t_1 < t_2 < \dots < t_n$ の順に $R_j^*(t_i)$, $p_j^*(t_i)$ の値が定まり、一度定まった値は変わらないことになる。図4.2では $\hat{p}_j(t_i)$ を用いた場合の $R_j^*(t)$ の構成例を示してあるが、 $(t_1, d_1 = d_{11} = 1)$, $(t_2, d_2 = d_{22} = 1)$ のもとで、 $(t_3, d_3 = d_{03} = 1)$ における $p_j(t_3)$ の推定値は、既に定まっている $R_1^*(t_1 - 0)$, $R_1^*(t_1)$ と $(t_3 - t_1)$ より $\hat{\lambda}_1(t_3)$ が、 $R_2^*(t_2 - 0)$, $R_2^*(t_2)$ と $(t_3 - t_2)$ より $\hat{\lambda}_2(t_3)$ が式(4.12)よりそれぞれ定まり、これらを式(4.13)に代入することにより得られる。得られた $\hat{p}_j(t_3)$ から式(4.10), (4.11)により $R_1^*(t_3)$, $R_2^*(t_3)$ が新たに定まり、引き続き $t_4 < \dots < t_n$ についてもこの操作を繰り返していけばよい。

また比較のため、iii) の中で最も簡単なものとして、 $d_{1i} = d_{2i} = 0$ の場合に t_i の直前、直後の故障データを用いる推定量もあげておく。このとき t_i における $\lambda_j(t_i)$ の推定量は

$$\tilde{\lambda}_j(t_i) = \frac{R_j^*(t_{-ji} - 0) - R_j^*(t_{+ji} - 0)}{R_j^*(t_i) (t_{+ji} - t_{-ji})} \quad (j=1, 2) \quad (4.14)$$

ここで t_{-ji} は $l < i$, $d_{jl} > 0$ なる t_l の最大値
 t_{+ji} は $l > i$, $d_{jl} > 0$ なる t_l の最小値

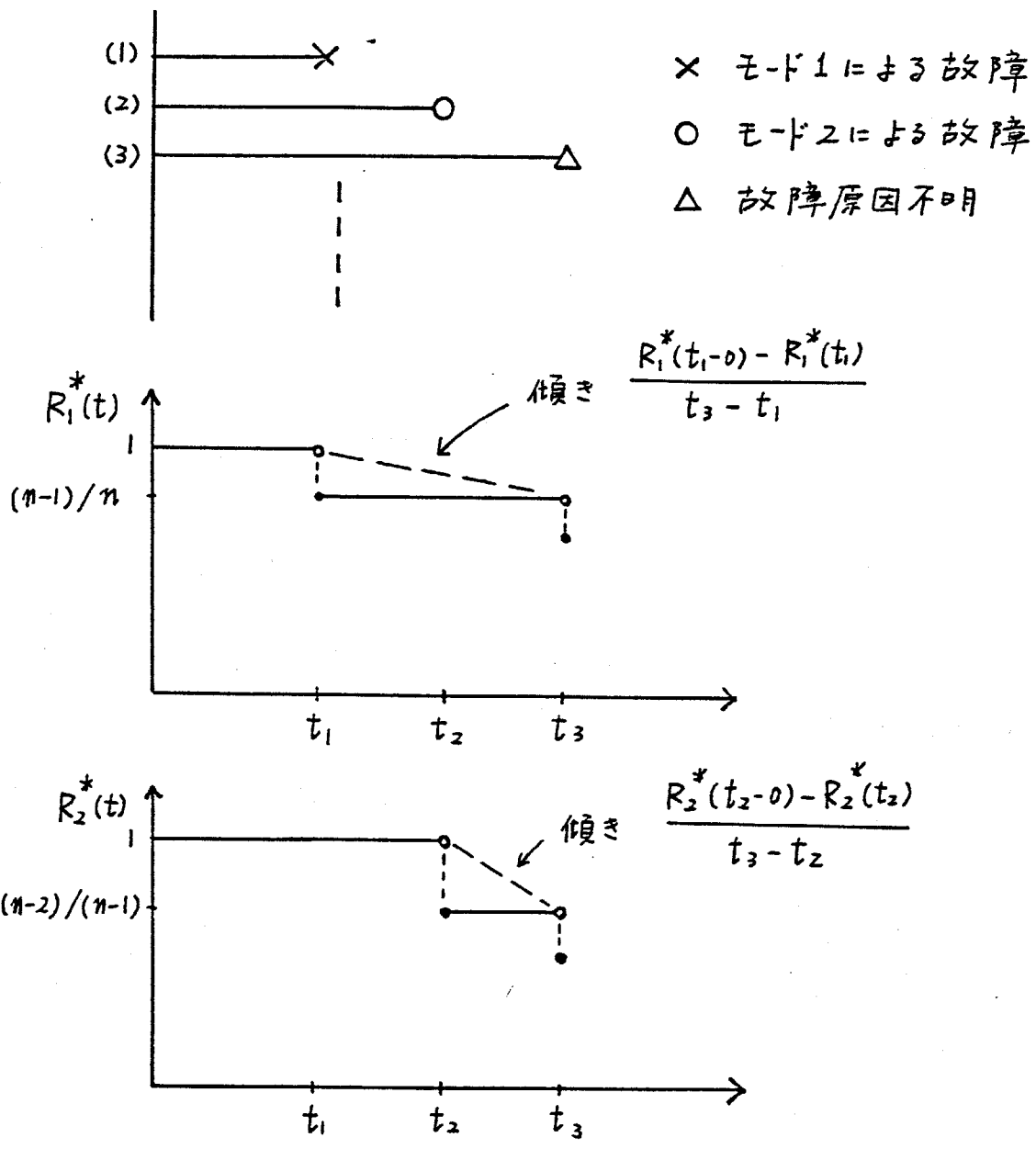


図 4.2 $\hat{p}_j(t_i)$ を用いた $R_1^*(t)$, $R_2^*(t)$ の構成例

となる。ただし対応する直前、直後の故障データがないときにはそれぞれ直後、直前の故障データだけを用いる（推定量の形は式(4.12)と同様なので省略する）。式(4.14)より得られる $p_j(t_i)$ の推定量を

$$\tilde{p}_j(t_i) = \frac{\tilde{\lambda}_j(t_i)}{\tilde{\lambda}_1(t_i) + \tilde{\lambda}_2(t_i)} \quad (4.15)$$

とする。なおこの場合も $d_{1i} > 0$ もしくは $d_{2i} > 0$ のときには $\tilde{p}_j(t_i)$ は式(4.2)で与えられるものとする。

4.4 シミュレーション実験による推定量の比較

K-M推定量の性質としては、漸近的な性質については様々なことが調べられているが、小標本での性質では Chenら^[11]により $F_1(t)$ と $F_2(t)$ が比例ハザードの関係にあるという条件のもとでの厳密なモーメントの式が求められているにすぎない。そこで、ここではシミュレーション実験により、4.3.2節で提案した区間分けされていないデータに基づくいくつかの推定法の偏りと平均2乗誤差（以下MSEと記す）を調べる。またそれとともに第2章、第3章と同様に、故障原因不明データを除いて構成した通常のK-M推定量と比べてどの程度推定精度が向上するかを検証する。

シミュレーション実験では $F_1(t)$, $F_2(t)$ としてワイブル分布

$$F_j(t) = 1 - \exp \left\{ - \left(\frac{t}{\alpha} \right)^{\beta} \right\} \quad (j=1, 2)$$

を用いた。ワイブル分布の形状母数の記号は通常 m を用いるが、ここでは故障

原因の観測されるデータ数 m と混同するので β を用いた。それに伴ない尺度母数も θ の代わりに α を用いた。この際、ワイブル分布の母数 $\alpha_1, \beta_1, \alpha_2, \beta_2$ と全データ数 n , さらに故障原因の観測されるデータ数 m が実験の因子となり得るが, ここでは $n=20, \alpha_1=\alpha_2=1.0$ に限定し, β_1, β_2 をそれぞれ $0.5, 1.0, 2.0$ の3水準, さらに m を $5, 10, 15$ の3水準にとった要因実験を行った。比較の対象とする推定量としては $R_j(t), p_j(t)$ の推定量の組合せから

- I. $\hat{R}_j(t)$ と $\hat{p}_j(t_i)$ の組合せ
- II. $\tilde{R}_j(t)$ と $\hat{p}_j(t_i)$ の組合せ
- III. $\hat{R}_j(t)$ と $\tilde{p}_j(t_i)$ の組合せ
- IV. $\tilde{R}_j(t)$ と $\tilde{p}_j(t_i)$ の組合せ

の4種類に加えて

V. 故障原因不明データを除いて構成した K-M 推定量

を取り上げ, 各条件における $R_1(t)$ に対するそれぞれの推定量の偏りと MSE $t=0.1 \sim 3.0$ (0.1) について 3,000 回のシミュレーションで求めた。また III と IV については推定値が収束するまで繰り返し計算をすることになるが, 収束の判定基準としては r 回目の $R_1(t)$ の推定値を $R_1^{(r)}(t)$ と記したとき

$$\max_{0 \leq t \leq 3.0} |R_1^{(r)}(t) - R_1^{(r-1)}(t)| < 0.005$$

を用いた。

図 4.3, 4.4 は $\beta_1 = \beta_2 = 1.0, m = 10$ の場合の各推定量の偏りと MSE である。これらより $F_1(t), F_2(t)$ が同一の指数分布の場合には, 偏りと MSE をあわせて考慮したとき IV が最も安定した推定量になっているが, 繰り返し計算

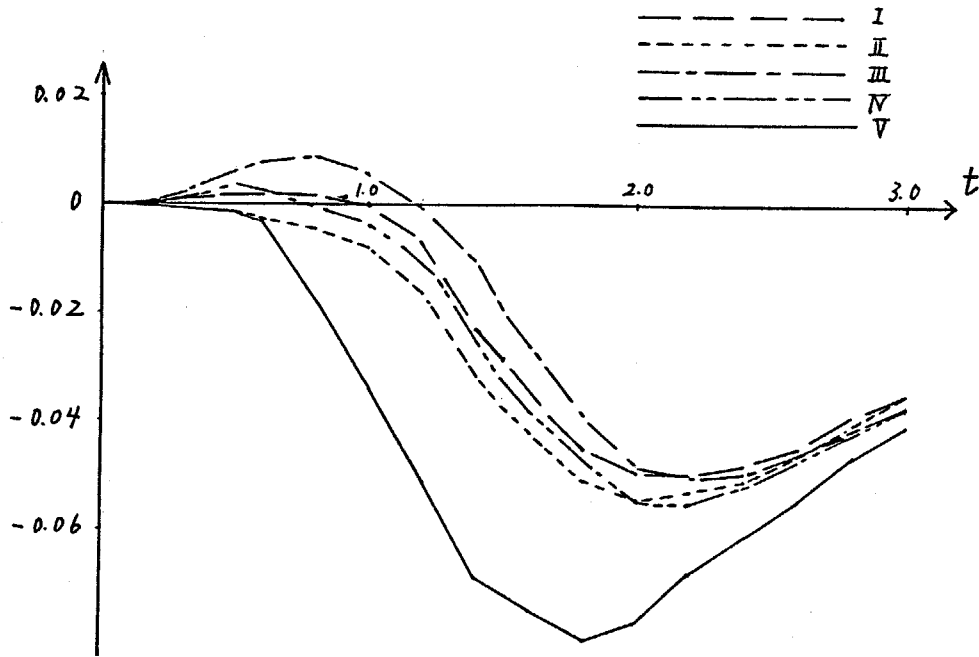


図 4.3 各推定量の偏り
 $(\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 1.0, n = 20, m = 10)$

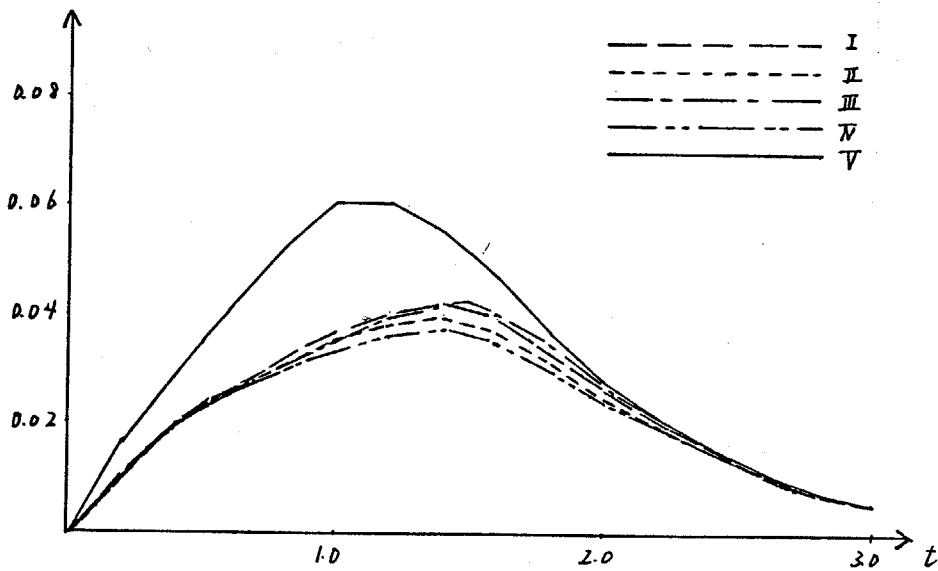


図 4.4 各推定量の MSE
 $(\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 1.0, n = 20, m = 10)$

を要しないⅡもさほど劣らないことがわかる。ⅠとⅢは推定量の形からわかるように、それぞれⅡとⅣに比べ大き目の値をとる。またⅠ、Ⅱ、Ⅲ、Ⅳのいずれを用いた場合にも、MSEはⅤよりも小さくなり、故障原因不明データを用いる効果が表われている。ⅤのMSEに対するⅡ、ⅣのMSEの比率は最も小さいところでそれぞれ0.60、0.55程度である。なお、この条件のもとではK-M推定量の偏り、MSEはChenらの与えた式より計算できるので、Ⅴに対する偏りとMSEの真値とシミュレーション値を比較したのが表4.2であるが、これよりシミュレーション値が非常によく合っているのが確かめられる。

また、詳しい実験結果は省略するが、 $\beta_1 = \beta_2$ の他の条件のもとでは次のような傾向がみられた。 $\beta_1 = \beta_2 = 0.5$ の場合、偏りはⅢが $t < 0.5$ で正の値をとるほかはいずれも負の値で、その絶対値は $\beta_1 = \beta_2 = 1.0$ の場合に比べて大きい。 $\beta_1 = \beta_2 = 2.0$ の場合には、 $t < 1.0$ では5つの推定量とも偏りはほとんどなく、それ以降は負の値をとるがその絶対値は小さい。MSEについては $\beta_1 = \beta_2 = 0.5$ 、 $\beta_1 = \beta_2 = 2.0$ の場合で山のピークはそれぞれずれるが、全般的傾向は $\beta_1 = \beta_2 = 1.0$ の場合と同様で、Ⅳが最も優れ、次にⅡが小さい値を出し

表4.2 K-M推定量の偏り、MSEの真値とシミュレーション値

($\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 1$, $n = 20$, $m = 10$)

t		0.5	1.0	1.5	2.0	2.5	3.0
偏り	真 値	-0.0017	-0.0339	-0.0710	-0.0742	-0.0588	-0.0411
	シミュレーション値	-0.0012	-0.0340	-0.0734	-0.0771	-0.0584	-0.0409
M.S.E.	真 値	0.0347	0.0605	0.0516	0.0287	0.0130	0.0054
	シミュレーション値	0.0343	0.0605	0.0517	0.0280	0.0135	0.0053

ている。

次に $m = 10$, $\beta_1 = 0.5$, $\beta_2 = 2.0$ の場合の $R_1(t)$ に対する各推定量の MSE を示したのが図 4.5 である。これは IFR の故障モードと DFR の競合モードの競合において DFR の故障モードに注目した場合であるが、この場合には $R_1(t)$ のノンパラメトリック推定は非常に精度が悪いがその中では III と IV, とくに IV が優れていることがわかる。またこのときの偏りは, I と II が最も小さいところで III と IV に比べて大きな負の偏りをもつという結果が得られた。

逆に, $\beta_1 = 2.0$, $\beta_2 = 0.5$ としたとき, すなわち IFR の故障モードに注目した場合の MSE が図 4.6 である。この場合も IV が最も優れ, 続いて III, II, I, V の順であり, IV を用いることにより V に比べ MSE はかなり改善されることがわかる。またこのときの偏りは, 逆に I と II が最も小さいところで正の偏り (絶対値は小さいが) をもつ。III と IV は $\beta_1 = \beta_2 = 2.0$ の場合とほぼ同様で, $t < 1.0$ まではほとんど偏りが無い。

このように, IFR の故障モードと DFR の故障モードの競合のような場合には, 直前の値だけを用いる I と II が偏りをもつため, 直前, 直後の値を用いる III と IV の有効性が出てくるようである。

最後に $n = m = 20$, すなわち完全な競合データに基づく $K-M$ 推定量に対する $m = 5, 10, 15$ におけるここで提案した不完全データに基づく推定量の効率も調べてみる。ここで効率としては, (完全な競合データに基づく $K-M$ 推定量の MSE) / (不完全な競合データに基づく推定量の MSE) を尺度とし, 対象とする推定量としては III と IV を用いた。 $\beta_1 = \beta_2 = 1.0$ の場合の $t = 0.5, 1.0, 2.0$ における効率の値を示したのが表 4.3 である。これより, 効率は t の値に依存するが, $m = 10$, すなわち半数程度の故障原因がわかっている場合は, かなりの推定効率が得られることがわかる。 β_1, β_2 の他の組合せの

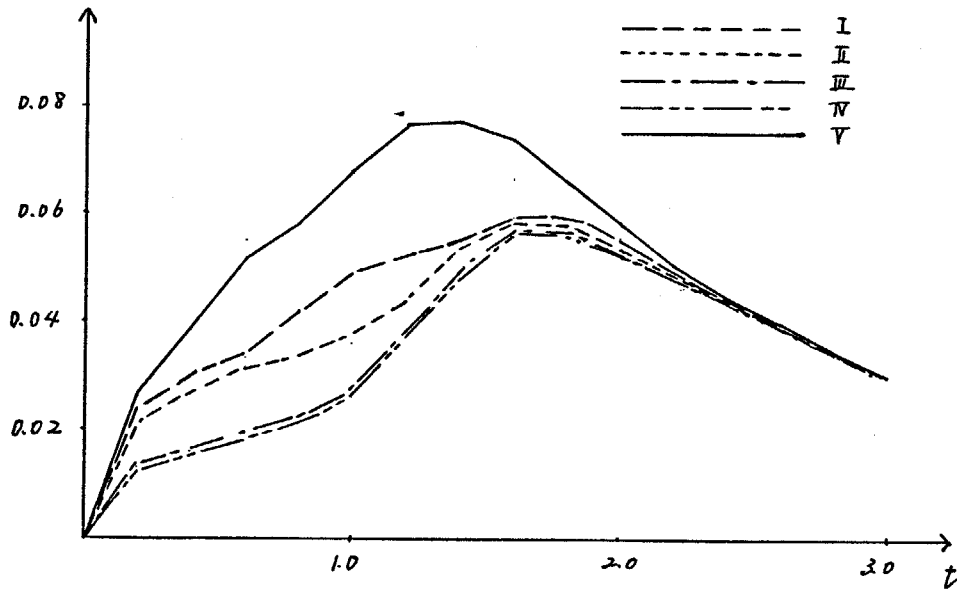


図 4.5 $R_1(t)$ に対する各推定量の MSE

$$(\alpha_1 = \alpha_2 = 1.0, \beta_1 = 0.5, \beta_2 = 2.0, n = 20, m = 10)$$

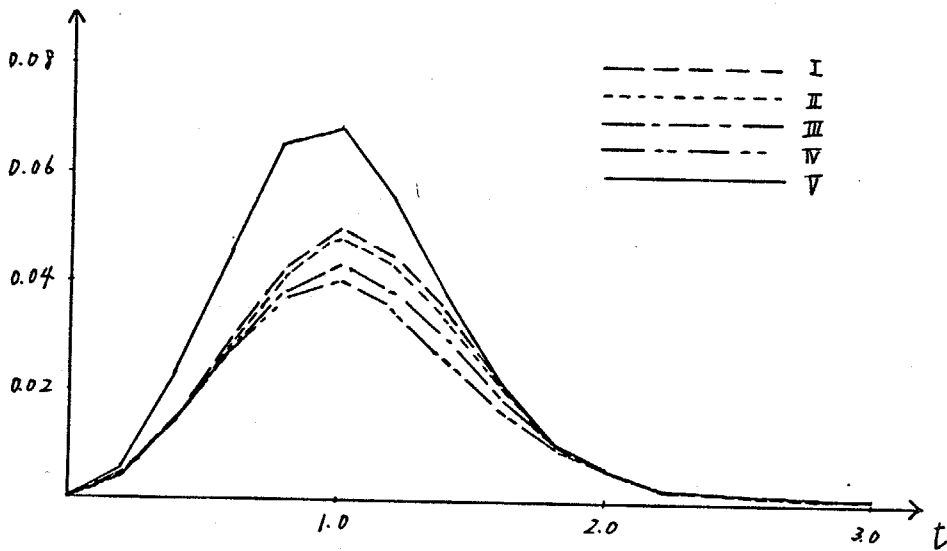


図 4.6 $R_1(t)$ に対する各推定量の MSE

$$(\alpha_1 = \alpha_2 = 1.0, \beta_1 = 2.0, \beta_2 = 0.5, n = 20, m = 10)$$

表4.3 完全な競合データ($n=m=20$)に基づくK-M推定量
に対するII,IVの効率 ($\alpha_1=\alpha_2=\beta_1=\beta_2=1, n=20$)

推定量 $\frac{t}{m}$	0.5		1.0		2.0	
	II	IV	II	IV	II	IV
5	0.369	0.385	0.493	0.524	0.762	0.768
10	0.695	0.704	0.743	0.790	0.867	0.911
15	0.865	0.870	0.846	0.903	0.939	0.945

もとでもこれに近い結果が得られている。

以上の実験結果より、ここで取り上げた推定量の中ではIVが最も優れていることがわかるが、これには計算機による繰り返し計算が必要である。また繰り返し計算を要しない簡便法としてはIIが比較的良好な性質をもっていることがわかった。一般に、データ数が多い場合にはIVを拡張して、故障原因不明データの故障時点の故障率推定において前後いくつかの故障データを用いることにより、推定量はさらに安定したものになると思われる。

4.5 適用例と考察

4.5.1 鉄道車両の故障データの解析

本節では前節までに提唱した推定法を実際の寿命データに適用した結果を示す。まず鉄道車両の保守管理委員会の報告書[48]にある昭和55年度から57年度までに本線上で起こった車両故障(A, B故障)データの解析結果を述べる。このデータは表4.4に示すように、標本数 $n=128$ で、要部検査(40万km

表 4.4 鉄道車両の故障データ [48]

No.	X	δ	No.	X	δ	No.	X	δ
1	334278	05	44	018184	16	87	221207	14
2	032581	11	45	084484	13	88	204644	01
3	268027	01	46	230433	15	89	042670	05
4	210070	15	47	114251	01	90	045004	04
5	275676	01	48	057647	13	91	231116	11
6	280519	12	49	165862	12	92	024448	01
7	359637	10	50	028949	01	93	253686	01
8	074643	10	51	026985	15	94	227049	09
9	185906	01	52	034937	15	95	263501	05
10	086634	11	53	349652	07	96	163699	10
11	163509	05	54	010919	15	97	272163	01
12	002705	13	55	244754	06	98	253542	15
13	200081	12	56	225721	10	99	220747	06
14	091710	14	57	266996	12	100	337463	14
15	000592	16	58	252759	12	101	314757	10
16	165646	15	59	142377	11	102	257928	10
17	015349	12	60	271798	01	103	230717	01
18	212003	10	61	166416	01	104	101916	01
19	080019	10	62	297401	11	105	019132	12
20	076367	06	63	102994	01	106	321384	15
21	106848	14	64	292103	04	107	238061	01
22	123443	15	65	279128	13	108	177220	13
23	116194	14	66	020095	11	109	138872	11
24	133897	05	67	056464	12	110	070296	12
25	054285	10	68	194235	13	111	340256	14
26	142020	12	69	125254	14	112	268225	01
27	043762	12	70	078375	16	113	162275	01
28	027653	10	71	126245	15	114	134553	14
29	033910	10	72	037256	01	115	100506	13
30	123655	14	73	029853	05	116	099235	99
31	004744	10	74	136675	15	117	145740	01
32	001261	04	75	169236	01	118	181628	12
33	203494	16	76	030668	11	119	000000	06
34	288075	01	77	112939	15	120	075468	13
35	379487	03	78	100886	15	121	013820	13
36	283269	05	79	066950	16	122	050950	12
37	006895	05	80	129182	12	123	016292	12
38	024792	06	81	200386	10	124	169161	14
39	163703	02	82	149042	14	125	342542	15
40	185121	12	83	359253	14	126	078221	10
41	214520	14	84	018681	13	127	294794	14
42	247582	01	85	332257	14	128	091495	05
43	196310	14	86	235108	01			

01	02	03	04	05	06	07	08	09
破損	変形	擦傷	摩耗	固渋	汚損	弛み	狂い	変質
10	11	12	13	14	15	24	99	16
漏洩	開放	不安定	短絡	アーク	溶損	腐食	その他	不明

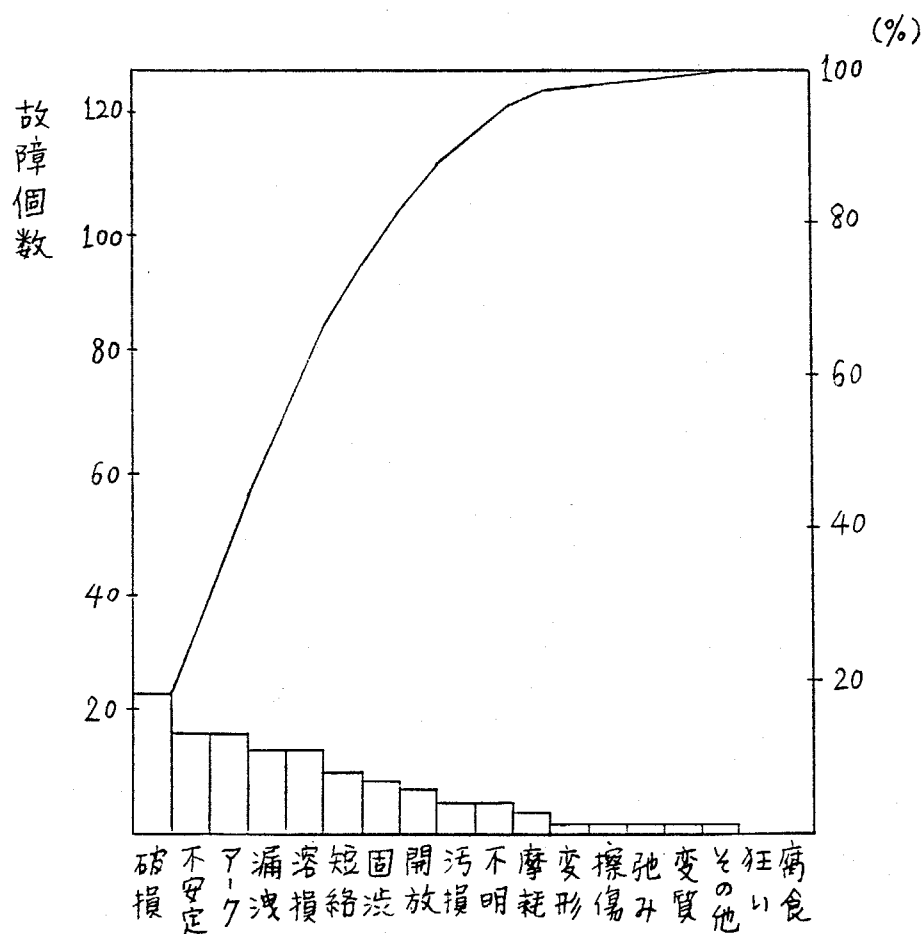


図 4.7 故障モードコードとパレート図

もしくは 24ヶ月経過したら行う) から故障までの走行距離と故障モードからなっている。故障モードコードの内容とその単純集計結果は図4.7に示す通りである。ここで故障原因不明データが5件あるが、これらは故障原因を調べたが

その原因がよくわからないものであり、本論文で対象とする故障原因不明データとはやや意味が異なる。すなわち本論文で対象とする故障原因不明データは、故障原因を調べていないもの、もしくは調べたがその結果が報告されていないものである。真の故障原因が何かによって原因不明になる可能性が異なる場合には、本論文での基本的な仮定である「故障原因が観測されるか否かはそのシステムの故障時間や真の故障原因に無関係である」ことが満たされないからである。そこで表4.4のデータについては、原因不明も1つのモードとみなすことにする。一方、鉄道車両の故障データにおいて、本論文で対象とする故障原因不明データが存在することも事実である。吉本^[51]は、新幹線のATCシステムの故障データ解析において、故障原因データの報告は走行距離データの報告よりも数ヶ月遅れるため、システム全体の故障率推定には故障原因不明の走行距離データも用いるが、各モードごとの故障率推定には故障原因不明の走行距離データを除いているとしている。

そこでここでは、表4.4のデータに対して、模擬的に無作為に選んだ64個すなわち50%のデータの故障原因を欠測値とし、(a) 128個のもともとの完全データに基づく推定結果、(b) 故障原因不明データを除いた64個の完全データに基づく推定結果、および(c) 4.3節で提案した方法(Ⅱを用いている)による故障原因不明を含んだ128個の不完全データに基づく推定結果を比較検討してみる。推定対象の故障モードは、図4.7の単純集計で最も多かった「破損」にした。破損故障は128個の中で23個あり、模擬的に半数の故障原因も欠測させた後は64個の中で13個になった。まず信頼度関数をノンパラメトリックに推定した結果を図4.8(a), (b), (c)に示す。(a), (b), (c)の間に大きな差は見られないが、(b)と(c)が(a)よりも小さな値になっているのは13個の破損故障が含まれているためであろう。(c)は故障原因不明デー

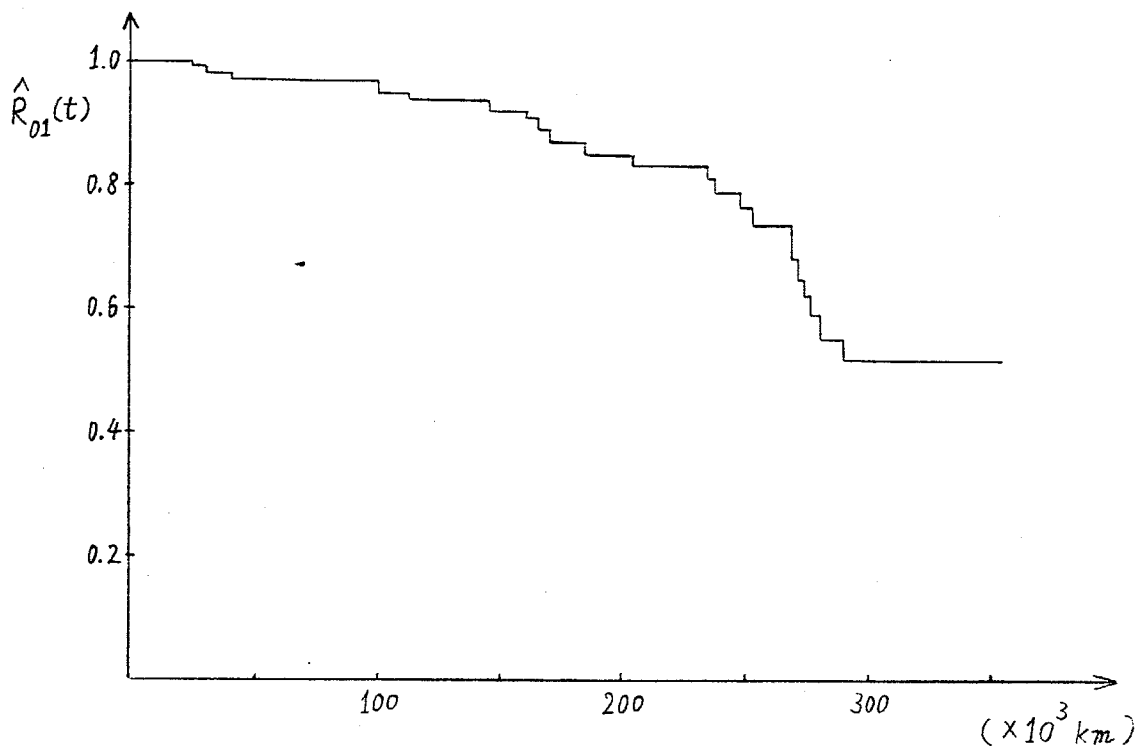


図 4.8 (a) 128 個の完全データに基づく信頼度関数の推定値

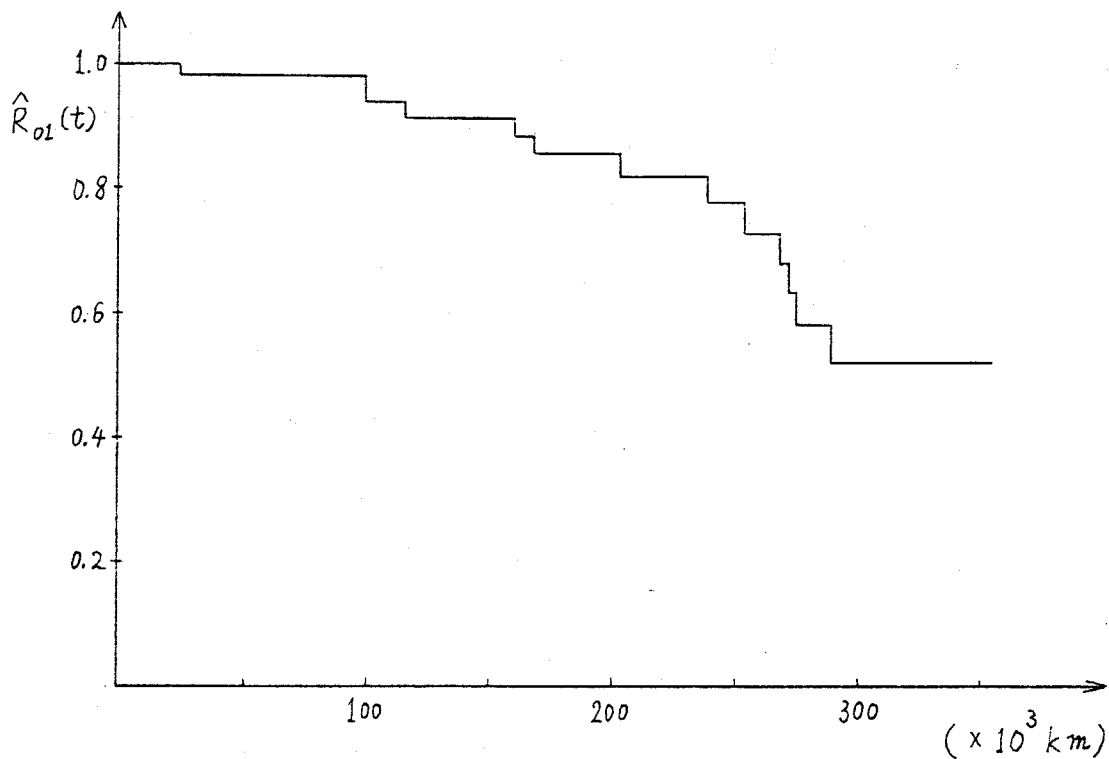


図 4.8 (b) 64 個の完全データに基づく信頼度関数の推定値



図 4.8 (c) 128個の不完全データに基づく信頼度関数の推定値

タの故障時点のほとんどが階段関数の不連続点になるのでなめらかになっている。さらに信頼度関数のノンパラメトリック推定値をもとに、ワイブル確率紙に打点した結果も図 4.9 (a), (b), (c) に示す。(a) では 25 万 km を過ぎたあたりから傾きが急になる凸性を顕著にみせているのに対し、(b) では打点数が少ないため、この凸性が母集団の性質なのかそれとも順序統計量のバラツキによるものなのかの判断は難しい。一方 (c) では (a) に近い傾向を示しており、本論文で提案した方法で故障原因不明データを併用することの効果表めれていると言える。たとえば、第 3 章で得られた知見を適用すれば、この凸性より「破損」モードの中にも異なるタイプの故障モードが存在しそれらが競合している可能性が示唆されよう。ただしこのような推測においては、とくに (c) ではワイブル確率紙上の点の重みが点ごとに大きく異なるので、それについても考慮した上で曲線をあてはめることが肝要であろう。たとえば 3 万 km から 9 万

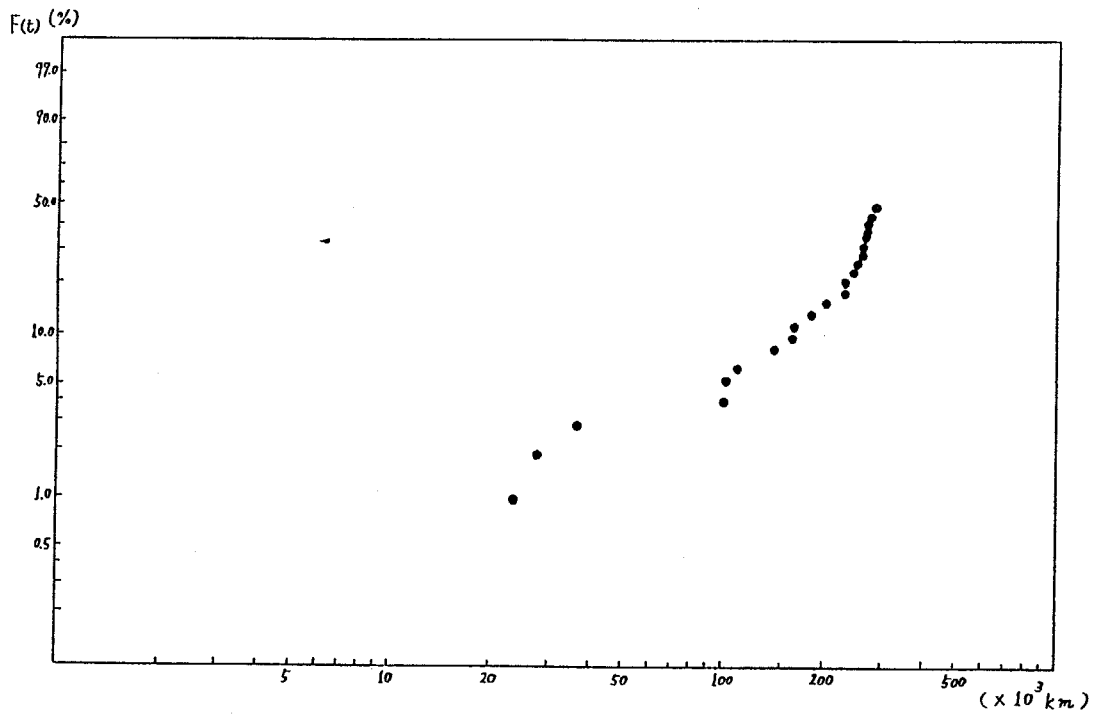


図 4.9 (a) 128個の完全データに基づくワイブル解析

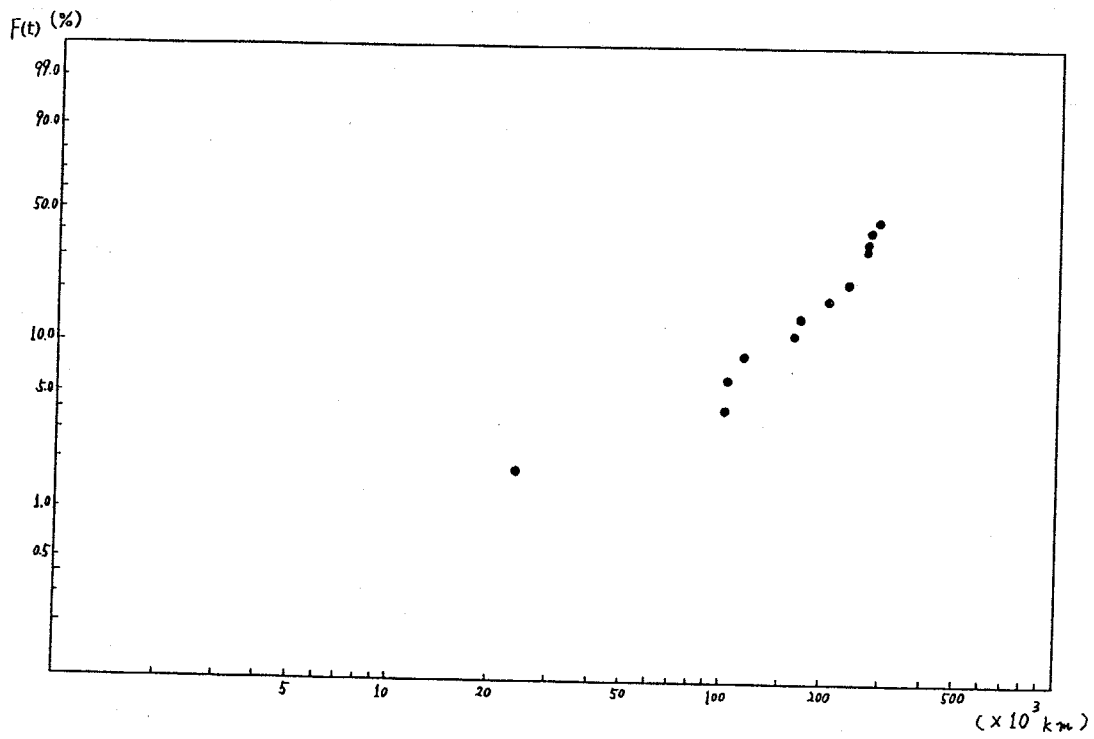


図 4.9 (b) 64個の完全データに基づくワイブル解析

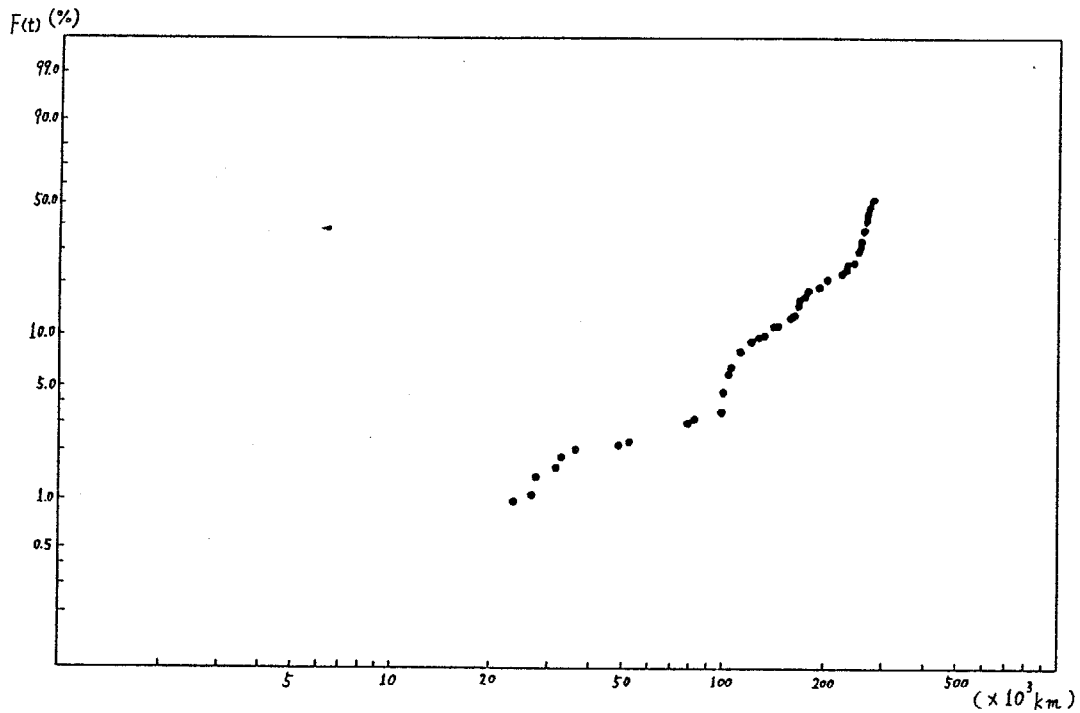


図 4.9(c) 128個の不完全データに基づくワイブル解析

km および 13 万 km から 15 万 km あたりの点はいずれも故障原因不明データであるので、これらの点の挙動は不安定であることに注意しなければならない。

4.5.2 大脳腫瘍患者の生存時間データの解析

Dinse^[15] は、本論文と同様に故障原因不明データの存在を指摘しその解析方法を論じている。ただし原因別ハザード分析の立場にあり、各モードごとの寿命特性の推定を目的としていない点が本研究と異なる。ここでは Dinse の挙げているデータより各モードごとの寿命特性の推定を行う。

Dinse のデータは、172 人の大脳腫瘍患者の腫瘍の発生から一定の大きさ

表4.5 大脳腫瘍患者の生存時間データ [15]

t_i	d_{1i}	d_{2i}	d_{0i}	打ち切り	t_i	d_{1i}	d_{2i}	d_{0i}	打ち切り
1	2	0	1	6	20	0	0	0	1
2	6	1	0	4	21	1	0	0	0
3	8	1	2	8	22	0	1	1	0
4	4	1	2	8	23	1	0	1	0
5	3	1	6	4	24	1	0	0	0
6	2	4	4	4	25	0	0	0	1
7	0	3	1	3	26	0	0	0	0
8	2	1	1	3	27	0	0	0	0
9	4	1	2	9	28	1	0	0	0
10	3	3	0	5	29	0	0	0	0
11	0	0	2	3	30	0	0	0	1
12	0	0	0	1	31	0	0	0	1
13	0	0	2	5	32	0	0	0	1
14	0	0	1	2	33	0	0	0	1
15	1	0	0	3	34	0	0	0	3
16	0	0	1	2	35	0	0	0	1
17	0	1	0	1	36	0	0	0	0
18	0	0	1	1	37	0	0	0	1
19	2	1	1	0					

(致死状態)になるまでの生存時間(単位は月)と,致死状態に至る直前の症状(歩行不可能もしくは歩行可能)からなっている. ここでは歩行不可能をモード1, 歩行可能をモード2とする. データにはランダムセンサリングがありセンサーデータについては必然的に症状は観測されていない. また症状不明データは29個含まれている. この場合, 観測単位が月でタイが非常に多いため区間分けされたデータとして処理することにし, 表4.5のように生命表データの形に整理した.

まず, 式(4.7)に基づき各モードのハザード関数をノンパラメトリックに推定した結果が図4.10である. ハザード関数のノンパラメトリック推定を行

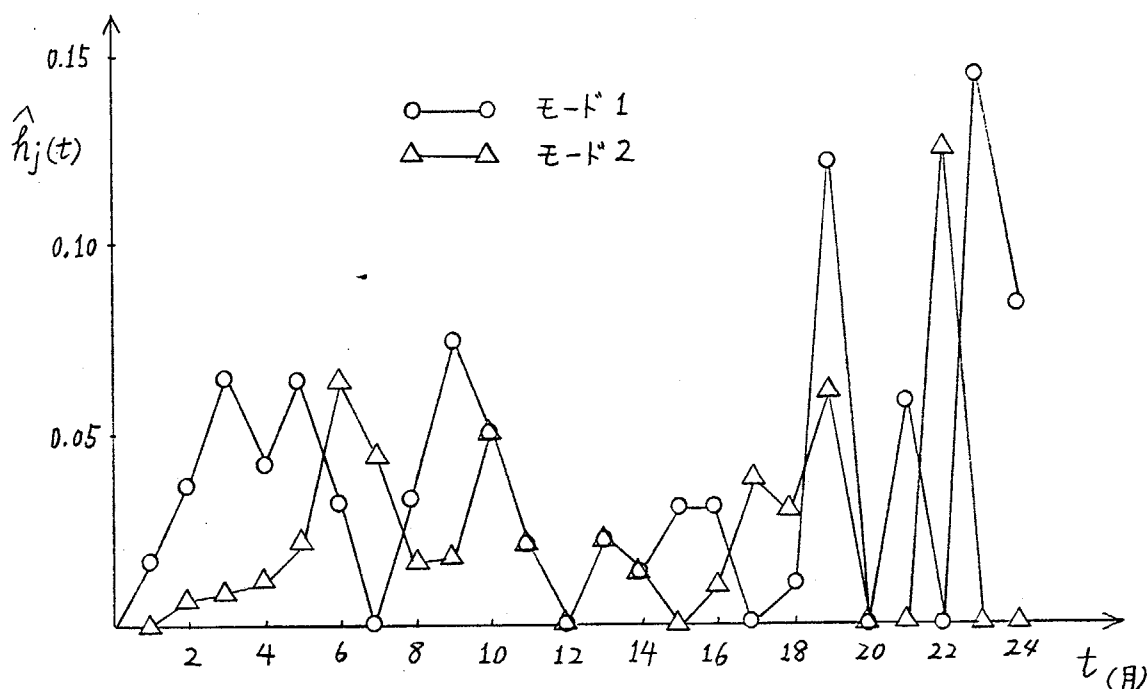


図 4.10 ハザード関数のノンパラメトリック推定値

うには各区間のデータ数が不十分とも思えるが、図 4.10 からそれぞれの故障率がほぼ一定であることや、2つのモード間の比例ハザード性などが読みとれる。また式(4.9)より各モードの信頼度関数を求めたのが図 4.11 である。さらにこの値をもとにワイブル確率紙にプロットした結果が図 4.12 である。図 4.12 より、モード1の寿命分布はワイブル分布によくあてはまることわかる。簡便的に直線をあてはめた結果、 $\hat{m}_1 = 0.95$, $\hat{\theta}_1 = 26.0$ と推定され、指数分布とみなせる結果となった。一方、モード2はやや傾きが増減する傾向にあるが、第7章で述べるようにワイブル確率紙においては、時間の小さいところでの点の位置は不安定であることからワイブル分布に従うとみてよいだろう。これについても簡便的に母数を推定したところ、 $\hat{m}_2 = 1.25$, $\hat{\theta}_2 = 35.0$ となりこちらもほぼ指数分布とみなせる。

以上の解析結果からこのデータに対して第2章で取り上げた指数分布を仮定した競合モデルを適用することが考えられるが、ここではランダムセンサリン

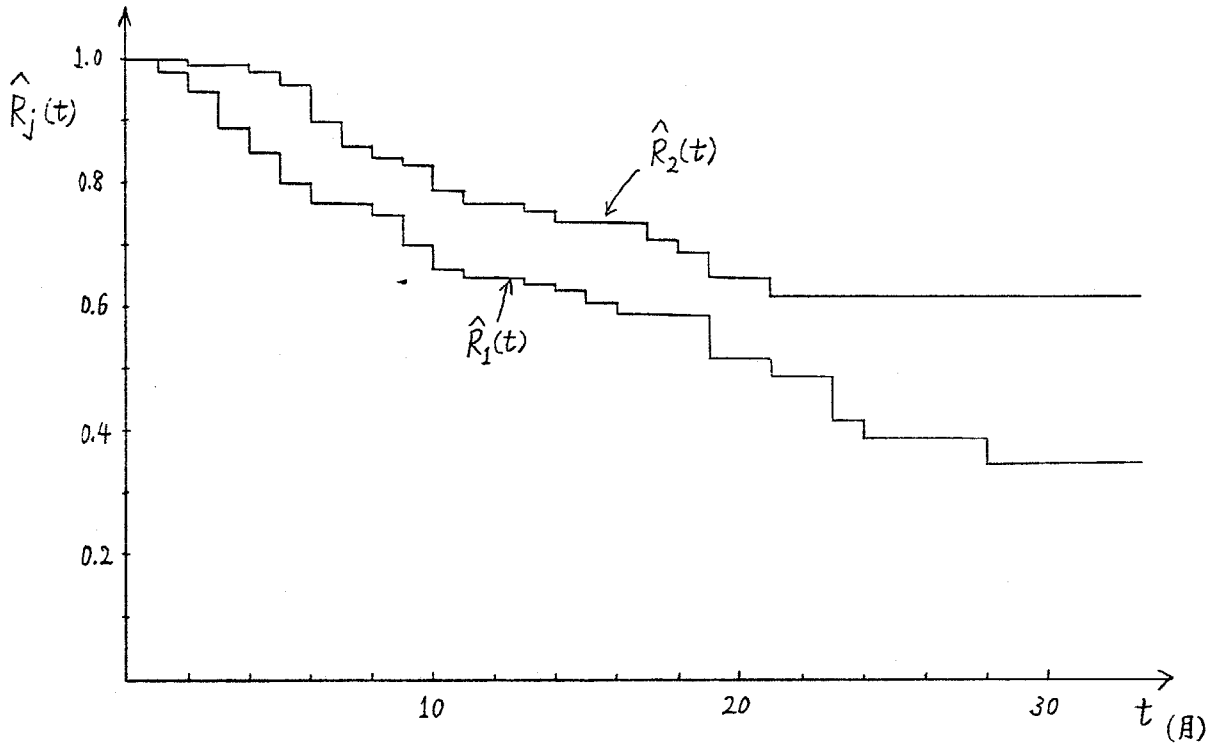


図 4.11 信頼度関数のノンパラメトリック推定値

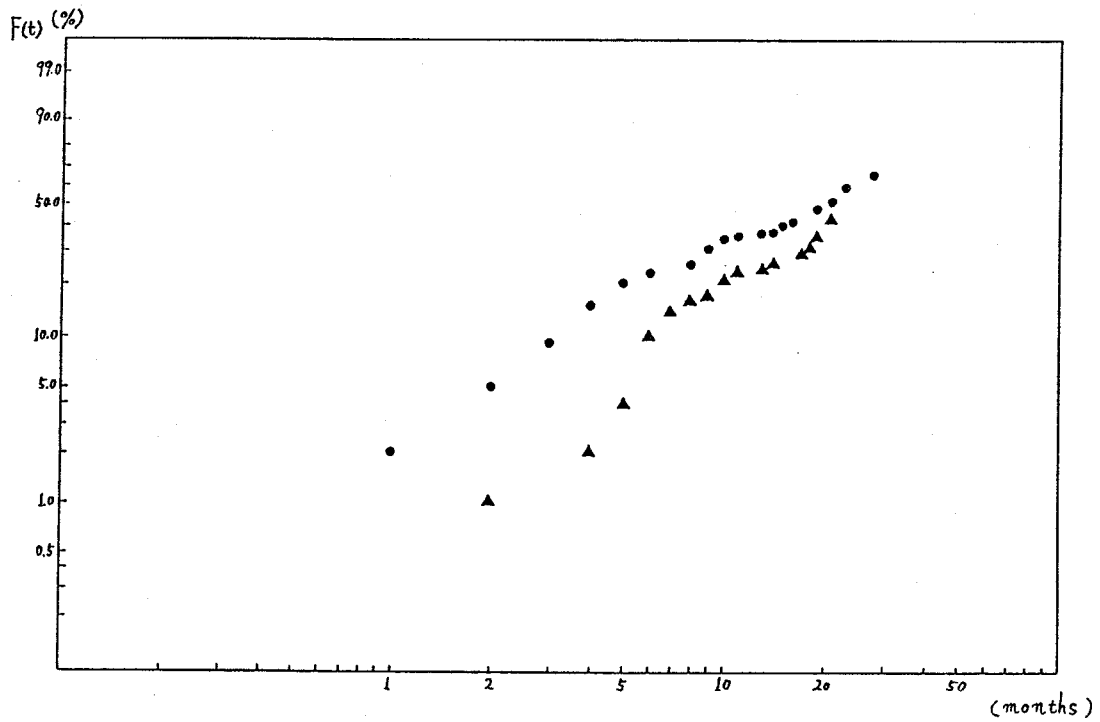


図 4.12 モード1, モード2のワイブル解析

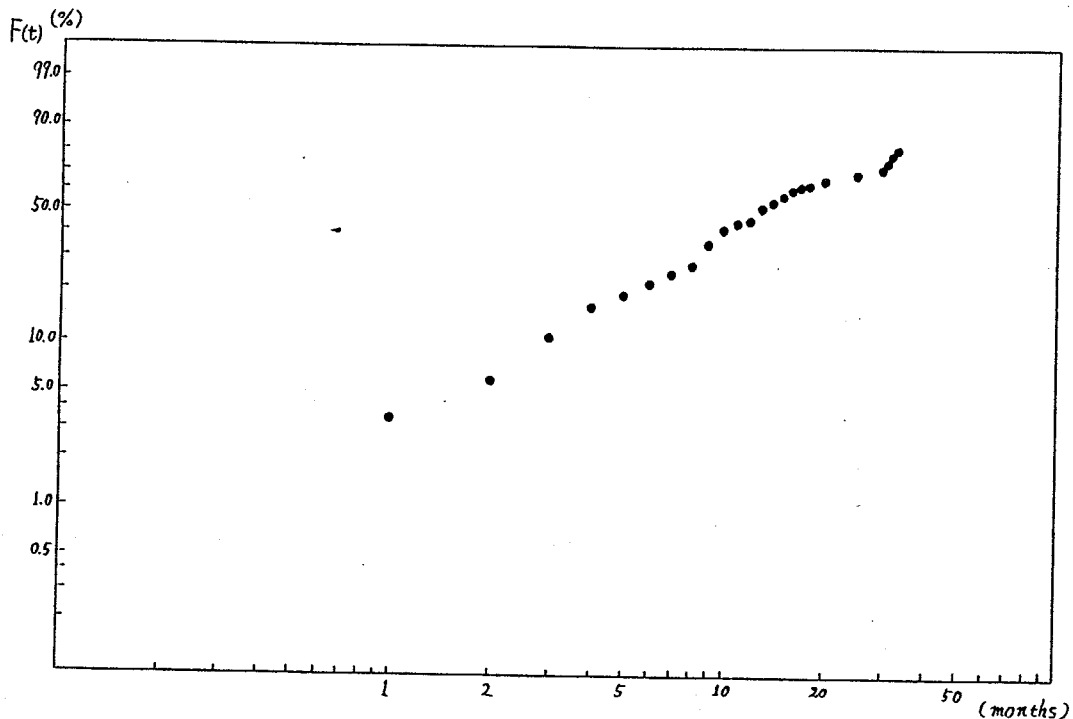


図 4.13 打切り時間データのワイブル解析

グが存在するため、打切り分布もまた指数分布でなければ適用できない。そこで「打切り」も1つのモード（これをモード3とする）とみなし、モード1、2と同様に打切り時間データに対してワイブル解析を行った結果が図4.13である。これより打切り時間はワイブル分布にあてはまり、しかも母数は $\hat{m}_3 = 1.10$, $\hat{\theta}_3 = 19.5$ と推定されるので指数分布とみなせる。よって $k=3$ の指数分布を仮定した競合モデルを仮定し、2.4節に示した方法が適用できる。このとき、原因不明データの真の原因はモード1、2のいずれかであり、モード3（打切り）ではないことに注意して λ_1, λ_2 の推定値を求めると

$$\lambda_1 = \frac{\sum_{i=1}^{37} d_{1i}}{\sum_{i=1}^{37} \sum_{j=1}^2 d_{ji}} \times \frac{\sum_{i=1}^{37} \sum_{j=0}^2 d_{ji}}{n} \times \frac{n-1}{\sum_{l=1}^{172} X_l} = 0.037$$

となり、同様に $\lambda_2 = 0.017$ を得る。

4.6 結 言

本章では、ノンパラメトリックな競合モデルについて考察し、その中で故障原因不明データを含む不完全な競合データに基づく寿命特性の推定問題を取り上げ、以下の結果を得た [34][35]。

- 1) 付加的仮定を用いない完全なノンパラメトリック推定において、十分統計量に基づく信頼度関数の強一致推定量を導くとともに、ノンパラメトリック推定での問題点を指摘した。
- 2) 1) で指摘した問題点を踏まえて、付加的仮定のもとで、区間分けされたデータと区間分けされていないデータに対して、それぞれEMアルゴリズムに基づき条件付故障原因確率を推定しながら故障原因不明データを利用する寿命特性の推定法を提示した。
- 3) 2) で与えた推定法により、故障原因不明データを無視した場合に比べ推定精度は大幅に改善され、また半教程度の故障原因が観測されていれば完全な競合データに比しかなりの推定効率が得られるというパラメトリック推定の場合とほぼ同様の結果を、シミュレーション実験で確認した。
- 4) 2) で与えた推定法より、ノンパラメトリック推定において何らかの事前情報があるとき、その情報をEMアルゴリズムのEステップの中で組み入れていくという統計解析における一般的な指針を得た。
- 5) 第2章から本章までに提唱した推定法を実際の寿命データに適用した解析結果を述べ、不完全な競合データに対する解析アプローチを例示した。

第5章 コンポーネントの寿命データに基づくシステムの寿命特性の推定

5.1 緒言

個別生産のシステムや製造コストの高い大規模システムの開発段階において、システムそのものの寿命試験で十分な寿命データを得ることは難しい。これに対し、コンポーネントのレベルでの寿命試験は、比較的容易に行える場合がある。また、汎用的コンポーネントで既存の多くのシステムに用いられているものについては、フィールドにおけるシステムの寿命データから前章までに論じた競合モデルを用いてコンポーネントの寿命特性に関する情報を得ることが可能である。このような場合、システムを構成する各コンポーネントの寿命情報を集め、これらに基づきシステム全体の寿命特性を推測する必要が生じてくる。コンポーネントの寿命データからシステムの寿命特性を推定する問題では、寿命特性として信頼度を取り上げたものとして、直列-並列システムについては Zacks and Even⁽⁵⁴⁾、Rutemiller⁽⁴⁴⁾、 k -out-of- m システムについては Basu and Mawazing⁽⁵⁾ の研究がある。本章では、5.2 節においてシステムの範囲をコヒーレントシステムまで拡張し、既存の結果をも含めた統一的な信頼度の推定量の表現を与える。一方、MIL-STD等では寿命特性として、非修理系の場合はMTTF (mean time to failure)、修理系の場合にはMTBF (mean time between failures)を採用することが多い。我が国でも、とく

にシステムの任務時間 (mission time) が明確でない消費財などではMTTFが用いられることが少なくない。そこで 5.3 節では 5.2 節の結果を基に、コヒーレントシステムのMTTFに対するパラメトリックおよびノンパラメトリック推定について論じる。さらに 5.4 節では前章までに述べた競合モデルによるシステムの寿命データの解析結果を上記の問題へ活用する方法について論じる。

5.2 コヒーレントシステムの信頼度の推定

5.2.1 異種コンポーネントの場合

m 個のコンポーネントからなるコヒーレントシステムを想定する。記号と定義は Barlow and Proschan⁽³⁾ に準ずる。コンポーネント i の状態を表わす 2 値変数

$$x_i = \begin{cases} 1 \cdots \text{コンポーネント } i \text{ が正常} \\ 0 \cdots \text{コンポーネント } i \text{ が故障} \end{cases} \quad (i=1, \dots, m)$$

とシステムの状態を表わす 2 値変数

$$\phi = \begin{cases} 1 \cdots \text{システムが正常} \\ 0 \cdots \text{システムが故障} \end{cases}$$

を用い、 $\phi = \phi(x_1, \dots, x_m)$ を構造関数とする。ここで、

- 1) $\phi(x_1, \dots, x_m)$ が各 x_i について単調非減少。
- 2) すべてのコンポーネントが関連性 (relevance) をもつ。

の2条件を満足するとき、そのシステムはコヒーレントシステムと呼ばれる。コヒーレントシステムは合理的な決定的構造をもつシステムとして最も広い範囲を包括し、直列-並列システム、 k -out-of- m システムなどはもちろんこの中に含まれる。また構造関数 $\phi(x_1, \dots, x_m)$ は最小バス集合、最小カット集合による表現定理 ((3), p.9) からわかるようにその表現は一意的でないが、 $x_i^2 = x_i$ を利用して展開し共通項をまとめると

$$\phi^c(x) = \sum_j a_j \prod_{i \in S_j} x_i \quad (5.1)$$

の形に書ける。ここで a_j は零でない整係数で、 S_j は $\{1, \dots, m\}$ のある部分集合である。式(5.1)は添字 j の順序を除いて一意であり、これを構造関数の標準形 (canonical form) と呼ぶ。これはコヒーレントシステムの信頼度関数 ((3), p.21) と同等である。すなわち各コンポーネントの寿命がたがいに独立なとき、時刻 t における各コンポーネントの信頼度 $R_1(t), \dots, R_m(t)$ とシステムの信頼度 $R_s(t)$ との間は

$$R_s(t) = \phi^c(R_1(t), \dots, R_m(t)) = \sum_j a_j \prod_{i \in S_j} R_i(t) \quad (5.2)$$

の関係が成り立つ。

さて、コンポーネント i の寿命 T_i がたがいに独立に2母数指数型分布族に属する分布 $f(t; \theta_i)$ にしたがって、各コンポーネントごとに独立に寿命試験を行い θ_i に対する完備十分統計量 Z_i が得られたとする ($i=1, \dots, m$)。このとき T_1, \dots, T_m の同時分布は独立性の仮定から多母数指数型分布族の分布となり、 (Z_1, \dots, Z_m) が十分統計量となる。さらに $(\theta_1, \dots, \theta_m)$ の可能な値の全体は m 次元の長

方形を含むので, Lehmannの定理 ([25], p.132) より (Z_1, \dots, Z_m) は完備である.

ところで式(5.2)は $R_i(t)$ ($i \in S_j$) の積の重み付け和であるから, たがいに独立なコンポーネントの信頼度の不偏推定量 $\hat{R}_i(t)$ ($i=1, \dots, m$) を式(5.2)に代入することでシステムの信頼度の不偏推定量が得られる. すなわち

$$\hat{R}_S(t) = \phi^c(\hat{R}_1(t), \dots, \hat{R}_m(t)) \quad (5.3)$$

が $R_S(t)$ の不偏推定量になる. さらに $\hat{R}_i(t)$ がそれぞれ $R_i(t)$ の一様最小分散不偏推定量 (Uniformly Minimum Variance Unbiased Estimator, 以下UMVUEと記す) であれば, $\hat{R}_S(t)$ は完備十分統計量の関数で不偏であるからこれもUMVUEとなる. $f(t; \theta_i)$ として指数分布を考えた場合には以下のように具体的な推定量の形を示すことができる.

コンポーネント i の寿命 T_i が指数分布

$$f(t; \theta_i) = \frac{1}{\theta_i} \exp\left(-\frac{t}{\theta_i}\right)$$

にしたがうとする. そしてコンポーネント i について n_i 個を定数打ち切り方式 (r_i 個で打ち切り) で観測し

$$T_{(i1)} \leq T_{(i2)} \leq \dots \leq T_{(ir_i)}$$

を得るものとする. このときコンポーネント i の信頼度 $R_i(t)$ のUMVUEは

$$\hat{R}_i(t) = \left[\left(1 - \frac{t}{Z_i} \right)^+ \right]^{r_i - 1} \quad (5.4)$$

で与えられる(4). ただしここで

$$Z_i = \sum_{j=1}^{r_i} T(t_j) + (n_i - r_i) T(t_{r_i})$$

$$(y)^+ = \max(0, y)$$

である. m 個のコンポーネントについて独立にこのような試験がなされたとすれば各コンポーネントの信頼度のUMVUE $\hat{R}_1(t), \dots, \hat{R}_m(t)$ が式(5.4)で与えられ, これらを式(5.3)に代入したものが, 構造関数 $\phi(x)$ をもつコヒーレントシステムの信頼度 $R_s(t)$ に対するUMVUEである. なお従来の直列・並列システムならびに k -out-of- m システムにおける $R_s(t)$ のUMVUEの形はすべて(5.3)の特殊なケースになっている.

5.2.2 同種コンポーネントの場合

m 個のコンポーネントの寿命がたがいに独立に同一分布にしたがう場合, コンポーネントの共通の信頼度を $R(t)$ で表わせば, 構造関数 $\phi(x)$ をもつコヒーレントシステムの信頼度は

$$R_s(t) = \phi^c(R(t), \dots, R(t)) = \sum_{l=1}^m b_l [R(t)]^l \quad (5.5)$$

と表わされる. ここで b_l は整係数であり, 式(5.5)は一意に定まる.

さて、コンポーネントの寿命分布を $f(t; \theta)$ とし、これに対する十分統計量 Z が得られたとき、 $l = 1, \dots, m$ に対して $[R(t)]^l$ の不偏推定量 $[\widehat{R}(t)]^l$ が求まれば、それらを式(5.5)に代入することにより $R_s(t)$ の不偏推定量が得られることが容易にわかる。すなわち

$$\widehat{R}(t) = \sum_{l=1}^m b_l [\widehat{R}(t)]^l \quad (5.6)$$

が不偏推定量である。さらにこの場合も Z が完備十分統計量で $[\widehat{R}(t)]^l$ が $[R(t)]^l$ の UMVUE であれば、式(5.6)が $R_s(t)$ の UMVUE となる。 $f(t; \theta)$ として指数分布を考えたときは以下のようになる。指数分布

$$f(t; \theta) = \frac{1}{\theta} \exp\left(-\frac{t}{\theta}\right)$$

にしたがう n 個のコンポーネントを定数打ち切り方式 (r 個で打ち切り) で観測し

$$T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(r)}$$

を得るものとする。このとき総試験時間

$$Z = \sum_{j=1}^r T_{(j)} + (n-r)T_{(r)}$$

が θ について完備十分統計量になる。また、指数分布においては信頼度について

$$[R(t)]^l = \exp\left(-\frac{lt}{\theta}\right) = R(lt)$$

が成り立つことから $[R(t)]^\ell$ の UMVUE は

$$\hat{R}(\ell t) = \left[\left(1 - \frac{\ell t}{Z} \right)^+ \right]^{r-1}$$

で与えられる ($\ell = 1, \dots, m$). これを (5.5) に代入した

$$\hat{R}_s(t) = \sum_{\ell=1}^m b_\ell \left[\left(1 - \frac{\ell t}{Z} \right)^+ \right]^{r-1} \quad (5.7)$$

が構造関数 $\phi(x)$ をもつコヒーレントシステムの信頼度に対する UMVUE である. なお, 母数が等しい指数分布の場合における従来の直列・並列システム, k -out-of- m システムでの $R_s(t)$ の UMVUE の形もすべて式 (5.7) の特殊なケースになっている.

5.2.3 推定量の比較

この節では寿命分布 $f(t; \theta_i)$ として指数分布を仮定した場合の $R_s(t)$ に対する UMVUE と MLE とを偏りと平均2乗誤差について比較する. まず母数がすべて異なる場合, 構造関数 $\phi(x)$ をもつコヒーレントシステムの信頼度 $R_s(t)$ に対する MLE は, 5.2.1 節と同様のデータが得られたとして

$$\tilde{R}_s(t) = \phi^c(\tilde{R}_1(t), \dots, \tilde{R}_m(t))$$

で与えられる。ただし $\tilde{R}_i(t)$ は $R_i(t)$ の MLE

$$\tilde{R}_i(t) = \exp\left(-\frac{r_i t}{Z_i}\right)$$

である ($i=1, \dots, m$). \hat{R}_s の期待値は標準形式 (5.2) に基づき

$$E[\hat{R}_s(t)] = \sum_j a_j \prod_{i \in S_j} E[\tilde{R}_i(t)]$$

と書け、 $E[\tilde{R}_i(t)]$ ($i=1, \dots, m$) で表わされることがわかる。また、平均2乗誤差は

$$MSE[\hat{R}_s(t)] = E[\hat{R}_s(t)]^2 - 2R_s(t)E[\hat{R}_s(t)] + [R_s(t)]^2$$

であるが、 $E[\hat{R}_s(t)]^2$ も $E[\hat{R}_s(t)]$ と同様に $E[\tilde{R}_i(t)]^2$, $E[\tilde{R}_i(t)]$ ($i=1, \dots, m$) で表わされる。

一方、UMVUE $\hat{R}_s(t)$ の分散は

$$\text{Var}[\hat{R}_s(t)] = E[\hat{R}_s(t)]^2 - [R_s(t)]^2$$

であるが、この場合も推定量の定義式より $E[\hat{R}_s(t)]^2$ が $E[\tilde{R}_i(t)]^2$ ($i=1, \dots, m$) で表わされることが容易に導ける。したがって構造関数 $\phi(x)$ が与えられれば、その都度システム全体の信頼度に対する推定量の偏りや平均2乗誤差は、コンポーネントの信頼度に対する推定量の1次、2次のモーメントを用いて評価される。

$E[\tilde{R}_i(t)]$, $E[\tilde{R}_i(t)]^2$, $E[\tilde{R}_i(t)]^2$ に対する正確な評価は Zacks and Even⁽⁵³⁾

によって与えられており、また近年、Chao⁽¹⁰⁾はこれらに対する近似式を提案している。

次に母数がすべて等しい場合であるが、このときの $R_s(t)$ に対するMLEは、5.2.2節と同様のデータが得られたとして

$$\hat{R}_s(t) = \sum_{l=1}^m b_l \left[\exp\left(-\frac{yt}{Z}\right) \right]^l$$

で与えられる。この場合には、 $MSE[\hat{R}_s(t)]$ 、 $Var[\hat{R}_s(t)]$ の評価はともに母数がすべて異なる場合に比べてやや面倒になるが、Chaoはk-out-of-mシステムについて $E[\hat{R}_s(t)]$ 、 $E[\hat{R}_s(t)]^2$ 、 $E[\tilde{R}_s(t)]^2$ に対する近似式を与えている。Chaoの与えた近似式はコヒーレントシステムの場合にも適用できる。それは $\hat{R}_s(t)$ が $\hat{R}_l(t)$ ($l=1, \dots, m$)の重み付け和で表われ、 $\tilde{R}_s(t)$ が $\tilde{R}(t)$ の多項式で表わされるという推定量の形から導き出せることである。以下に簡単な例を示し、具体的に $\hat{R}_s(t)$ と $\tilde{R}_s(t)$ の比較を行う。

(例) 図5.1に示すようなブリッジ構造(bridge structure)システムを考える。このシステムは直列・並列システムでもk-out-of-mシステムでもないがコヒーレントシステムである。構造関数は例えば最小パスによる表現を用いると、 $x = (x_1, \dots, x_5)$ として

$$\phi(x) = 1 - (1 - x_1 x_4) \times (1 - x_2 x_5) \times (1 - x_1 x_3 x_5) \times (1 - x_2 x_3 x_4)$$

と書ける。また、最小カットによる表現を用いると

$$\begin{aligned} \phi(x) = & [1 - (1 - x_1)(1 - x_2)] \times [1 - (1 - x_4)(1 - x_5)] \\ & \times [1 - (1 - x_1)(1 - x_3)(1 - x_5)] \times [1 - (1 - x_2) \\ & (1 - x_3)(1 - x_4)] \end{aligned}$$

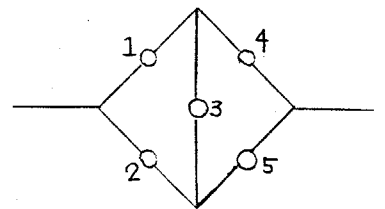


図 5.1

ブリッジ構造システム

と書ける。これらを標準形に直すと

$$\begin{aligned}\phi^c(x) &= x_1 x_4 + x_2 x_5 + x_1 x_3 x_5 + x_2 x_3 x_4 - x_1 x_2 x_3 x_4 \\ &\quad - x_1 x_2 x_3 x_5 - x_1 x_2 x_4 x_5 - x_1 x_3 x_4 x_5 - x_2 x_3 x_4 x_5 \\ &\quad + 2x_1 x_2 x_3 x_4 x_5\end{aligned}$$

となる。

ここで、すべてのパラメータが等しい場合を考えると、寿命時間 T におけるブリッジ構造システムの信頼度は

$$R_S(t) = 2[R(t)]^2 + 2[R(t)]^3 - 5[R(t)]^4 + 2[R(t)]^5$$

となる。ただし $R(t)$ はすべてのコンポーネントの共通の信頼度である。このとき $R_S(t)$ の UMVUE は

$$\begin{aligned}\tilde{R}_S(t) &= 2\left[\left(1 - \frac{2t}{Z}\right)^+\right]^{r-1} + 2\left[\left(1 - \frac{3t}{Z}\right)^+\right]^{r-1} - 5\left[\left(1 - \frac{4t}{Z}\right)^+\right]^{r-1} \\ &\quad + 2\left[\left(1 - \frac{5t}{Z}\right)^+\right]^{r-1}\end{aligned}\quad (5.8)$$

である。一方 MLE は

$$\begin{aligned}\hat{R}_S(t) &= 2\exp\left(-\frac{2rt}{Z}\right) + 2\exp\left(-\frac{3rt}{Z}\right) - 5\exp\left(-\frac{4rt}{Z}\right) \\ &\quad + 2\exp\left(-\frac{5rt}{Z}\right)\end{aligned}$$

である。

ここで (5.8) が 0 になる確率について触れておく。この確率は、打ち切り数

r と、使用時間 t と平均寿命 θ との比 $k = t/\theta$ によって定まるもので、 $2Z/\theta$ が自由度 $2r$ の χ^2 分布にしたがうことから

$$\Pr\{\hat{R}_s(t) = 0\} = \Pr\{Z < Z_t\} = \Pr\left\{\frac{2Z}{\theta} < \frac{4t}{\theta}\right\} \quad (5.9)$$

で求めることができる。

$\theta=1$ の場合の $\hat{R}_s(t)$ と $\tilde{R}_s(t)$ の偏りと平均2乗誤差をモンテカルロ・シミュレーションと Chao の近似式で調べた結果を表5.1、表5.2に示す。ここでは簡単のため、 $n=r$ とした。シミュレーションの繰り返し数は1000回である。表5.1が偏りで、上段に近似式による値、下段にシミュレーション結果を示した。ただし $\hat{R}_s(t)$ については不偏性が証明されているので近似値は控え入していない。表5.1よりMLE $\hat{R}_s(t)$ は負の偏りをもっていることがわかる。表5.2が平均2乗誤差で、やはり上段が近似値、下段がシミュレーション結果である。表2からわかることとして、 $k=t/\theta$ が小さいところではUMVUEの方がMLEより優れており、逆に k が大きいところではMLEの方が優れている。このような結果は Zacks and Even が単一システムについて厳密に評価した結果、さらに Basu and Mawaziny が k -out-of- m システムについてシミュレーションで調べた結果と一致している。このことからコヒーレントシステムにおいては、概してこのような傾向があるのではないかと類推される。また表5.1、5.2で与えた n と t の範囲では式(5.9)の値は $n=5$, $t=0.50$ で最大になるが、その場合にも式(5.9)の値は0.005以下であるから、 $\hat{R}_s(t)$ が0になる確率はほとんど無視できるであろう。

また表5.1、5.2よりChaoの近似式による値は一部を除いてシミュレーション結果と非常に良くあっている。 $n=5$ $t=0.10$ の平均2乗誤差では、UM

VUE、MLE と近似値がシミュレーション結果の約半分になっているが、これは Chao の近似式が推定量の1次、2次のモーメントをそれぞれ $1/n^2$ の項まで展開したものであるため、 $n=5$ では十分に近似されていなければいけないことによるものと思われる。

表 5.1 偏りの比較 ($\theta=1$)

τ	$R_S(\tau)$	n	5	10	20
0.10	.980559	UMVUE	.000227	.000539	.000382
		MLE	-.015431 -.016244	-.006607 -.006076	-.003025 -.002609
0.20	.927377	UMVUE	.003571	.000567	.001067
		MLE	-.046239 -.040671	-.020640 -.019625	-.009569 -.008440
0.50	.669513	UMVUE	-.007309	.010235	.000030
		MLE	-.077887 -.065592	-.041102 -.030782	-.021092 -.020890

*) 上段が近似式による値
下段がシミュレーションによる値

表 5.2 平均二乗誤差の比較 ($\theta=1$)

τ	$R_S(\tau)$	n	5	10	20
0.10	.980559	UMVUE	.000511 .001013	.000197 .000224	.000101 .000086
		MLE	.000933 .002445	.000326 .000412	.000101 .000119
0.20	.927377	UMVUE	.006130 .007130	.002420 .003057	.001051 .001050
		MLE	.010333 .012788	.003638 .004429	.001218 .001341
0.50	.669513	UMVUE	.087899 .071776	.032658 .022231	.013506 .011851
		MLE	.048257 .049067	.027695 .022002	.011743 .012141

*) 上段が近似式による値
下段がシミュレーションによる値

5.3 コヒーレントシステムのMTTFの推定

5.3.1 推定量の導出

まず、 m 個の独立なコンポーネントからなる直列システムのMTTFの推定問題を取り上げる。この直列システムにおける結果がコヒーレントシステムでの推定法の基礎になるからである。また、具体的には推定量の形を与えるために各コンポーネントの寿命は指数分布にしたがうものとする。すなわちコンポーネント i の寿命 T_i の確率密度関数を

$$f(t; \theta_i) = \frac{1}{\theta_i} \exp\left(-\frac{t}{\theta_i}\right)$$

とする ($i=1, \dots, m$)。ここで θ_i はすべて異なり関数的に独立とする。

いま、各コンポーネントごとに独立に寿命試験を行い、それぞれ n_i 個のたがいに独立な完全データ ($T_{11}, \dots, T_{1n_1}, \dots, T_{m1}, \dots, T_{mn_m}$) が得られるものとする。すると、 $(Z_1 = \sum_{j=1}^{n_1} T_{1j}, \dots, Z_m = \sum_{j=1}^{n_m} T_{mj})$ が $(\theta_1, \dots, \theta_m)$ に対する完備十分統計量となる。またこのとき、直列システムのMTTFを μ_s で表わすと

$$\mu_s = 1 / \left[\sum_{i=1}^m \frac{1}{\theta_i} \right] \quad (5.10)$$

であるから、 μ_s のMLEは各コンポーネントの θ_i に対するMLE $\tilde{\theta}_i = Z_i$

n_i を式(5.10)の θ_i にそれぞれ代入することで得られる。すなわち

$$\tilde{\mu}_s = 1 / \left[\sum_{i=1}^m \frac{n_i}{Z_i} \right] \quad (5.11)$$

である。一方、 μ_s の UMVUE は以下のように求められる。いま、各コンポーネントの第1サンプル $(T_{11}, T_{21}, \dots, T_{m1})$ から構成される統計量 $\min(T_{11}, \dots, T_{m1})$ は μ_s の不偏推定量になっている。また (Z_1, \dots, Z_m) が同時完備十分統計量であるから Lehmann-Scheffé の定理を適用すれば (Z_1, \dots, Z_m) を与えたときの $\min(T_{11}, \dots, T_{m1})$ の条件付期待値が μ_s の UMVUE となる。 Z_i を与えたときの T_{i1} の条件付分布は [4] にあるように

$$F(t|Z_i) = \begin{cases} 1 - \left(1 - \frac{t}{Z_i}\right)^{n_i-1}, & t \leq Z_i \\ 0, & t > Z_i \end{cases}$$

であり、 $\sum_{i=1}^m n_i$ 個のデータはすべて独立であるから (Z_1, \dots, Z_m) を与えたときの $\min(T_{11}, \dots, T_{m1})$ の条件付分布は

$$F(t|Z_1, \dots, Z_m) = \begin{cases} 1 - \prod_{i=1}^m \left(1 - \frac{t}{Z_i}\right)^{n_i-1}, & t \leq \min(Z_1, \dots, Z_m) \\ 0, & t > \min(Z_1, \dots, Z_m) \end{cases}$$

となる。これより μ_s の UMVUE $\hat{\mu}_s$ は

$$\hat{\mu}_s = \int_0^{\infty} \prod_{i=1}^m \left[\left(1 - \frac{t}{Z_i}\right)^+ \right]^{n_i-1} dt \quad (5.12)$$

で与えられることを知る. ここで $(y)^+ = \max(0, y)$ である. またここで注意すべきことは式(5.12)の被積分関数が時刻 t における直列システムの信頼度に対する $\overline{U}M\overline{V}\overline{U}E$ [52] になっていることである. 式(5.12)の被積分関数は t の多項式であるから, $\hat{\mu}_S$ は Z_1, \dots, Z_m の有理関数として表現される. たとえば, $m=2$ のときには

$$\hat{\mu}_S = \begin{cases} Z_1 \sum_{k=0}^{n_2-1} \frac{(n_1-1)!(n_2-1)!}{(n_2-k-1)!(n_1+k)!} \left(-\frac{Z_1}{Z_2}\right)^k, & Z_1 \leq Z_2 \\ Z_2 \sum_{k=0}^{n_1-1} \frac{(n_1-1)!(n_2-1)!}{(n_1-k-1)!(n_2+k)!} \left(-\frac{Z_2}{Z_1}\right)^k, & Z_1 > Z_2 \end{cases}$$

となる.

次にこの場合のノンパラメトリック推定法について考える. いまコンポーネント i における k_i 番目の寿命データを T_{ik_i} とすると $\min(T_{1k_1}, \dots, T_{mk_m})$ は μ_S の不偏推定量である(このとき, この不偏推定量をカーネルという). そこで, このカーネルに基づき各コンポーネントの n_i 個のサンプル内で対称となる \overline{U} -統計量

$$\mu_S^* = \frac{1}{n_1 \cdots n_m} \sum_{k_1=1}^{n_1} \cdots \sum_{k_m=1}^{n_m} \min(T_{1k_1}, \dots, T_{mk_m}) \quad (5.13)$$

を考えることができる. 式(5.13)はノンパラメトリックな状況における $\overline{U}M\overline{V}\overline{U}E$ である(たとえば [52] p. 153, 定理 3.7.1 を参照).

以上の結果をもとに, コヒーレントシステムへの拡張を考える. 各コンポーネントの信頼度 $R_1(t), \dots, R_m(t)$ とシステムの信頼度 $R_S(t)$ の間には式(5.2)の関係があるが, さらにシステムのMTTF μ_S の間には

$$M_S = \int_0^{\infty} R_S(t) dt = \int_0^{\infty} \sum_j a_j \prod_{i \in S_j} R_i(t) dt \quad (5.14)$$

の関係があるので、 M_S の推定問題は S_j に含まれるコンポーネントからなる直列システムのMTTFの推定問題に帰着することがわかる。

たとえば、 $\tilde{R}_i(t)$, $\hat{R}_i(t)$ をそれぞれ $R_i(t)$ のMLE, UMVUEとすると、 M_S のMLE, UMVUEは、それぞれ

$$\tilde{M}_S = \int_0^{\infty} \sum_j a_j \prod_{i \in S_j} \tilde{R}_i(t) dt \quad (5.15)$$

$$\hat{M}_S = \int_0^{\infty} \sum_j a_j \prod_{i \in S_j} \hat{R}_i(t) dt \quad (5.16)$$

となる。直列システムにおける式(5.11), (5.12) はそれぞれ式(5.15), (5.16)の特別な場合になっている。

次にノンパラメトリック推定についてであるが、いまコンポーネント i の寿命を t_i ($i=1, \dots, m$) とすると、コヒーレントシステムの寿命 t_S は、システムの最小パス集合を P_1, \dots, P_p としたとき

$$t_S = \max_{1 \leq j \leq p} \min_{i \in P_j} t_i$$

となることが知られている ([3], p. 12)。そこで、コンポーネント i における k_i 番目の寿命データを T_{ik_i} とすれば

$$\max_{1 \leq j \leq p} \min_{i \in P_j} T_{ik_i} \quad (5.17)$$

が M_S の不偏推定量になる。したがって式(5.17)をカーネルとする \bar{U} -統計量

$$M_S^* = \frac{1}{n_1 \cdots n_m} \sum_{k_1=1}^{n_1} \cdots \sum_{k_m=1}^{n_m} \max_{1 \leq j \leq p} \min_{i \in I_j} T_{i k_i} \quad (5.18)$$

を考えることができる。式(5.18)はノンパラメトリックにおける $\bar{U}M\bar{V}\bar{U}E$ であるが、これを別な見地から見直すことができる。 M_S と $R_1(t), \dots, R_m(t)$ の間には式(5.14)の関係があり、 M_S のMLE, $\bar{U}M\bar{V}\bar{U}E$ はそれぞれ $R_i(t)$ のMLE, $\bar{U}M\bar{V}\bar{U}E$ を式(5.14)の $R_i(t)$ に代入することで得られたわけであるが、まったく同様にして、 $R_i(t)$ に対する経験信頼度関数(階段関数)を $R_{E,i}(t)$ と表わしたとき、 $R_{E,i}(t)$ を式(5.14)の $R_i(t)$ に代入した

$$\int_0^{\infty} \sum_j a_j \prod_{i \in S_j} R_{E,i}(t) dt \quad (5.19)$$

を考えると、式(5.19)は式(5.18)に一致する。いずれも $\prod_{i=1}^m n_i$ 通りのシステムの擬似故障データの算術平均に他ならないからである。また、式(5.19)はノンパラメトリックにおけるMLEである。これは経験分布関数がノンパラメトリックにおける分布関数のMLEであることに基づくことである。

5.3.2 偏りと平均2乗誤差による比較

5.3.1 節で取り上げたコヒーレントシステムのMTTFに対する各種推定量の性質は、システムの構造、前提とするコンポーネントの寿命分布、データ数

n_i などに依存し, 一般論として論じることは困難である. そこでここでは, 5.2.3節のように各コンポーネントの寿命分布として指数分布を仮定した場合を中心に調べてみる. また, システムの構造は一般に多様であるが, 推定量の多からわかるように, 評価の上で基本となるのは直列システムである. 最も簡単な場合である $m=2$ の直列システムにおける MLE の偏り, MSE の近似式として

$$E(\tilde{\mu}_s - \mu_s) = -\frac{(a+1)\theta_1^2\theta_2^2}{an_1^2(\theta_1+\theta_2)} \left[a(n_1+1)\theta_1^2 + 2(a n_1 - a - 1)\theta_1\theta_2 + (a n_1 + 1)\theta_1^2 \right] + O\left(\frac{1}{n_1^3}\right) \quad (5.20)$$

$$E(\tilde{\mu}_s - \mu_s)^2 = \frac{\theta_1^2\theta_2^2}{an_1^2(\theta_1+\theta_2)} \left[a n_1(\theta_1^4 + a\theta_2^4) + 2(a n_1 - 2a - 2) \times (\theta_1^4 + a\theta_2^2)\theta_1\theta_2 + (a+1)(a n_1 + 5a + 5)\theta_1^2\theta_2^2 \right] + O\left(\frac{1}{n_1^3}\right) \quad (5.21)$$

を得る. ここで $a = n_2/n_1$ である. またとくに $\theta_1 = \theta_2 = \theta$, $n_1 = n_2 = n$ の場合には, 偏り, MSE に対する厳密な評価式

$$E(\tilde{\mu}_s - \mu_s) = -\frac{\theta}{2(2n+1)} \quad (5.22)$$

$$E(\tilde{\mu}_s - \mu_s)^2 = \frac{2n+5}{4(2n+3)(2n+1)} \theta^2 \quad (5.23)$$

を得る. 一方, このときの \bar{U} -統計量の分散に対しては厳密な評価式

$$\text{Var}(\mu_s^*) = \frac{\theta_1^2 \theta_2^2 [2n_1 \theta_1^2 + (n_1 + a n_1 + 3) \theta_1 \theta_2 + 2a n_1 \theta_2^2]}{a n_1^2 (\theta_1 + \theta_2)^2 (\theta_1 + 2\theta_2) (2\theta_1 + \theta_2)} \quad (5.24)$$

を得る。式(5.20), (5.21) ならびに(5.24)の導出については付録2も参照されたい。また \bar{U} M \bar{V} \bar{U} E の分散に対する精度の良い近似式は残念ながら得られなかった。

式(5.20), (5.21) ならびに(5.24)によるMLEの偏り, MSEの近似値と \bar{U} -統計量の分散値を表5.3に示す。表5.3で, 式(5.22), (5.23)による $n_1 = n_2$, $\theta_1 = \theta_2 = 1.0$ の場合のMLEの偏り, MSEの正確な値との比較から, 式(5.20), (5.21)は実用上十分な近似精度を与えていると言える。さ

表5.3 直列システムにおけるMLEの偏り, MSEの近似値
ならびに \bar{U} -統計量の分散 ($n_1 = n_2 = n$, $\theta_1 = 1.0$)

θ_2 / μ_s		$n = 5$	$n = 10$	$n = 15$	$n = 20$
0.2 / .166667	MLE	-.009568 .003549	-.004707 .001890	-.003121 .001286	-.002334 .000974
	\bar{U} -統計量	.004690	.002291	.001515	.001132
0.4 / .285714	MLE	-.022276 .009227	-.011400 .004723	-.007658 .003173	-.005765 .002388
	\bar{U} -統計量	.012698	.006122	.004031	.003005
0.6 / .375000	MLE	-.032300 .015518	-.016864 .007615	-.011401 .005645	-.008611 .003771
	\bar{U} -統計量	.021045	.010080	.006622	.004929
0.8 / .444444	MLE	-.039669 .021731	-.020891 .010432	-.014162 .006858	-.010710 .005108
	\bar{U} -統計量	.029087	.013892	.009117	.006783
1.0 / .500000	MLE	-.045000 (-.045455) .027500 (.026224)	-.023750 (-.023810) .013125 (.012940)	-.016111 (-.016129) .008611 (.008553)	-.012188 (-.012195) .006406 (.006381)
	\bar{U} -統計量	.036667	.017500	.011481	.008542

* MLEについては上段が偏り, 下段がMSE

()内は式(5.22), (5.23)による正確な値

らに直列システムが負の偏りもち、U-統計量の分散はMLEのMSEに比べ約3割程度大きくなっていることがわかる。

次にUMVUEも含めて比較するために、シミュレーション実験により指数分布のもとでの3つの推定量の偏り、MSEも求めた結果を示す。対象とするシステムの構造としては、2コンポーネント直列システムに加えて2コンポーネント並列システムを取り上げた。シミュレーションの繰り返し数は5,000

表5.4 直列システムにおける各推定量の偏り、
MSEのシミュレーション結果 ($n_1=n_2=n, \theta_1=1.0$)

θ_2/μ_3	n	5	10	15	20
0.2/ .166667	UMVUE	-.000914 .004267	-.000323 .002072	-.000169 .001389	-.000085 .001044
	MLE	-.010269 .003718	-.004988 .001928	-.003281 .001321	-.002418 .001005
	U-統計量	-.000867 .004706	-.000426 .002302	-.000210 .001552	-.000131 .001170
0.4/ .285714	UMVUE	-.000199 .010919	.000100 .005200	-.000382 .003454	-.000212 .002579
	MLE	-.022343 .009419	-.011283 .004819	-.008021 .003287	-.005967 .002483
	U-統計量	-.000207 .012860	-.000161 .006199	-.000805 .004087	-.000615 .003093
0.6/ .375000	UMVUE	-.001304 .017687	.000537 .008355	-.000408 .005406	-.000555 .004021
	MLE	-.033558 .015423	-.016381 .007785	-.011790 .005178	-.009145 .003898
	U-統計量	-.001387 .021465	-.000206 .010353	.000805 .00680-	-.001150 .005083
0.8/ .444444	UMVUE	-.001204 .023902	.000799 .011285	-.000780 .007252	-.000437 .005415
	MLE	-.041134 .020955	-.020186 .010563	-.014912 .006986	-.011137 .005265
	U-統計量	-.001444 .029595	.000007 .014276	-.001364 .009355	-.000990 .007037
1.0/ .500000	UMVUE	.000266 .029691	-.000551 .014029	-.000889 .009160	-.000502 .006775
	MLE	-.045212 .025962	-.024341 .013213	-.016972 .008833	-.012676 .006594
	U-統計量	.000056 .036675	-.000987 .018180	-.000989 .011913	-.001166 .008769

* 上段が偏り、下段がMSE

回である。その結果が表 5.4, 5.5 である。表 5.4 での MLE, \bar{U} -統計量の偏り, MSE のシミュレーション値を表 5.3 と比較すれば, このシミュレーション結果がある程度信頼できるものと言える。表 5.4 において MSE をみると, MLE, $\bar{U}MVUE$, \bar{U} -統計量の順であり, MLE と $\bar{U}MVUE$ に大差はないが, \bar{U} -統計量はこれらに比べ劣ることがわかる。また表 5.5 より並列システムでは MLE が正の偏りもち, MSE に関しては $\theta_2 = 0.2$ の場合を除いて

表 5.5 並列システムにおける各推定量の偏り,
MSE のシミュレーション結果 ($n_1 = n_2 = n, \theta_1 = 1.0$)

θ_2/μ_s	n	5	10	15	20
0.2/ 1.033333	UMVUE	.000500 .191801	.001183 .096248	.001545 .064768	-.000164 .048574
	MLE	.009855 .190846	.005849 .095984	.004658 .064631	.002169 .048492
	\bar{U} -統計量	.000452 .192333	.001286 .096462	.001587 .064884	-.000118 .048653
0.4/ 1.114286	UMVUE	-.002084 .178219	.000034 .088799	-.000097 .060332	.000740 .045093
	MLE	.020062 .180017	.011417 .089254	.007452 .060517	.006496 .045203
	\bar{U} -統計量	-.002076 .179508	.000295 .089588	.000326 .060972	.001144 .045521
0.6/ 1.225000	UMVUE	.000919 .178678	-.000229 .088414	.000711 .060019	.000604 .044942
	MLE	.033173 .185168	.016689 .090210	.012093 .060828	.009194 .045401
	\bar{U} -統計量	.001001 .181392	.000514 .089782	.001108 .061045	.001199 .045696
0.8/ 1.355556	UMVUE	-.003112 .193589	.001002 .097902	.000177 .064755	-.000350 .048893
	MLE	.036818 .204214	.021987 .101039	.014310 .066147	.010351 .049685
	\bar{U} -統計量	-.002873 .197265	.001794 .100449	.000762 .066129	.000204 .050024
1.0/ 1.500000	UMVUE	.001496 .227237	-.002419 .114540	-.000350 .077026	-.000079 .057904
	MLE	.047174 .241593	.021371 .118465	.015733 .078872	.012096 .058967
	\bar{U} -統計量	.001906 .232640	-.001983 .116946	-.000250 .079077	-.000436 .059604

*) 上段が偏り, 下段が MSE

はUMVUEが最も優れているが、概して3つの推定量に大きな優劣はないと言える。

5.3.3 頑健性に関する考察

ここでは、仮定しているモデルからのズレに対する頑健性について考察する。まず、5.2.2節のように同種コンポーネントを仮定したモデルを考える。指数分布の場合、共通の母数を θ としたとき、コヒーレントシステムのMTTFは式(5.14)からわかるように、すべて $A\theta$ の形をとる。ここで A は正の係数である。したがって全データをまとめ、 $\frac{\sum_{i=1}^m Z_i}{\sum_{i=1}^m n_i}$ により θ を推定し、これに A を乗じたものがMTTFのMLEでありUMVUEである。最も簡単な場合として2コンポーネント直列システムについて考えると、 $\theta_1 = \theta_2 = \theta$ とすると $\mu_s = \theta/2$ であるが、このとき

$$\hat{\mu}_s^* = \frac{Z_1 + Z_2}{2(n_1 + n_2)} \quad (5.25)$$

は不偏で、その分散は $\theta^2/[4(n_1 + n_2)]$ である。たとえば、 $n_1 = n_2 = 10$ 、 $\theta = 1.0$ のとき分散値は0.0125であるが、 θ_1 と θ_2 が異なるとして求めた場合にも式(5.23)よりMLEのMSEは0.01294であり、UMVUEのMSEは表5.4より読みとると0.0140であるからさほど劣らないと言える。これに対し、 $\theta_1 \neq \theta_2$ のとき $\theta_1 = \theta_2$ として μ_s を推定すると、たとえば $n_1 = n_2 = 10$ 、 $\theta_1 = 1.0$ 、 $\theta_2 = 0.6$ のとき式(5.25)は0.025の偏りをもつ上、MSEは0.0091となり、対応する表5.4のMLE、U

MVUEのMLEに比べ約10~17%増加している。したがって $\theta_1 = \theta_2$ という情報によほどの信頼がおける場合以外は異種コンポーネントとして解析した方が安全と言えよう。

次に、想定する分布形からのズレに対する頑健性を調べてみる。これについては指数分布を仮定している状況を考え、各コンポーネントの寿命が実際にはワイブル分布に従うときに、指数分布を前提として構成した各推定量のMSEをシミュレーション実験で求めた。図5.2は、2つのコンポーネントが同一のワイブル分布

$$F(t) = 1 - \exp(-t^\beta)$$

に従う場合に、これをそれぞれ異なる指数分布に従うものとして構成した直列システムの μ_s に対する各推定量の \sqrt{MSE} / μ_s を示している。これより形状母数 β が1から離れた場合(とくに大きいとき)の \bar{U} -統計量の頑健性が顕著に表われており、従来から指摘されてきたように(たとえば[53]を参照)、指数分布を仮定したパラメトリックな推定法が頑健性に乏しいことが、この種の問題でも確認された。また、図5.3は並列システムの場合であるが、この場合には β が1付近でも、 \bar{U} -統計量はUMVUE, MLEに比べてほとんど差がなく、1より大きいところでは非常に優れていることが明らかになり、 \bar{U} -統計量の有用性が示されたものと言える。

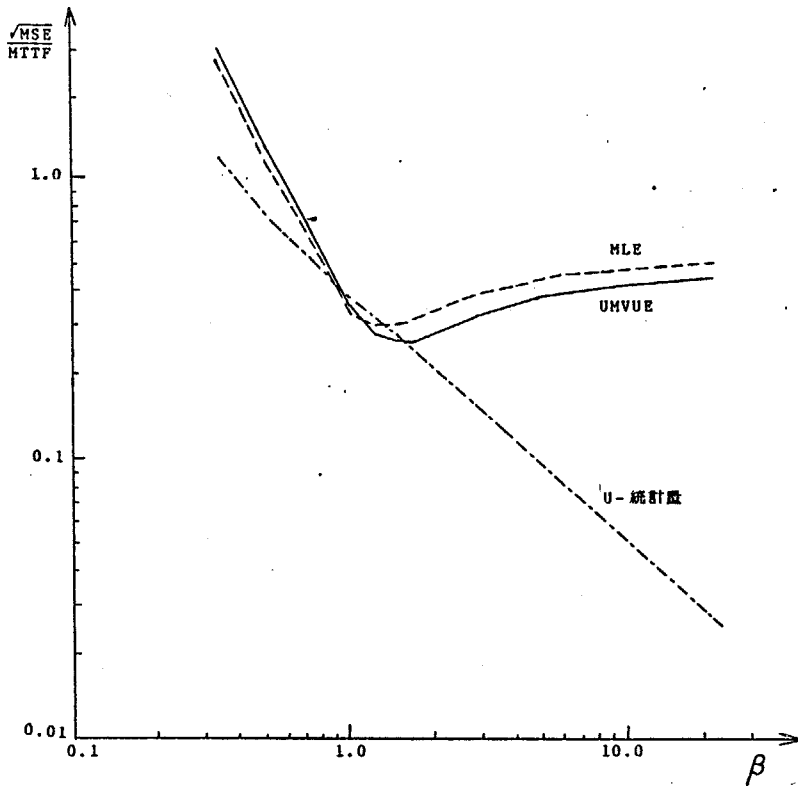


図 5.2 直列システムにおける頑健性の比較
($n_1 = n_2 = 5$)

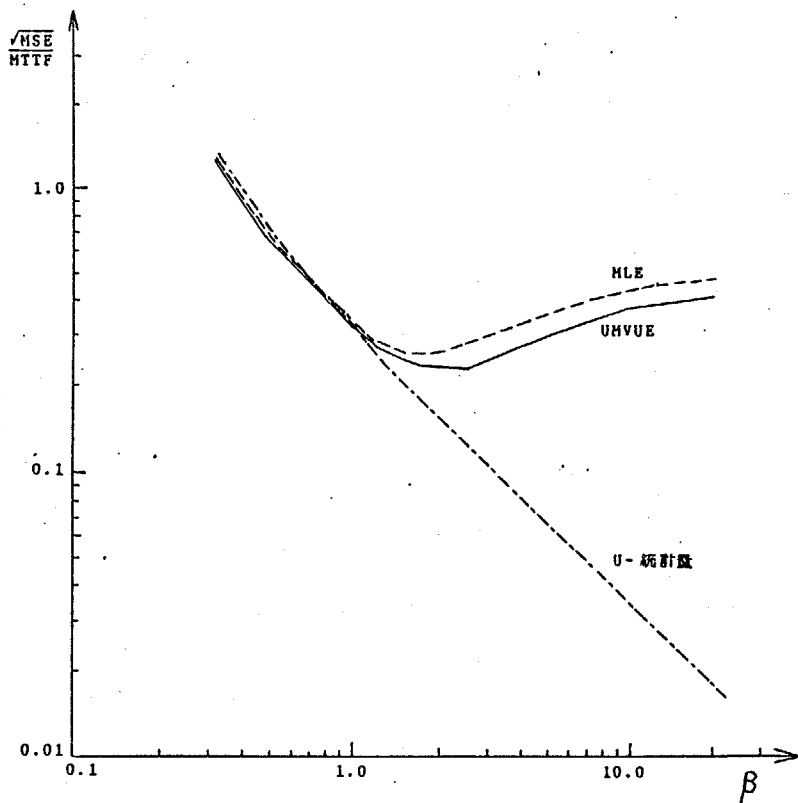


図 5.3 並列システムにおける頑健性の比較
($n_1 = n_2 = 5$)

5.4 競合データとコンポーネントデータの併合

新製品開発あるいは一品生産のシステムの開発にあたっては、この段階で製品、システムの信頼性特性を的確に把握しておく必要がある。一般に、新製品においても製品を構成するすべての部品が新しいとは限らず、一品生産のシステムにおいてもすべてのコンポーネントがそのシステム固有のものではなく、標準的、汎用的なものも含まれていることが多い。このようなとき、既にフィールドで用いられている部品、コンポーネントについては、それらが用いられているシステムの寿命データから競合モデルを用いることによりそれらの寿命情報も得ることが考えられる。また、そのコンポーネントの寿命試験データが過去にあれば、寿命試験とフィールドとの対応をみることもできる。一方、新しい部品、固有なコンポーネントについては、コンポーネントレベルの寿命試験により新たに寿命情報を採取する必要があり、フィールドにおける実際の寿命特性への外挿も考慮されねばならない。そこで、既存システムの寿命データと寿命試験でのコンポーネントの寿命データを併合して、新製品あるいは一品生産システムの寿命特性を統計的に推定することが当然考えられる。新製品開発においては、このような統計解析の上で、少数の試作品の寿命データとの対応をみるのがより効果的である。

いま、 m 個のコンポーネントからなるシステムを考えたとき、最初の m_0 個のコンポーネントは同種のものが既存のシステムにも用いられているものとし、残りの $m - m_0$ 個のコンポーネントは新しく開発されたコンポーネントとする。コンポーネント i ($i = 1, \dots, m_0$) は、システム \mathcal{S}_i に用いられており、 n_i 個のシステム \mathcal{S}_i の寿命データを得るとする。またコンポーネント i (i

$= m_0 + 1, \dots, m$) については, n_i 個のコンポーネント i の寿命データを得るものとする.

まず, 分布形を仮定したパラメトリックな推定について考える. このとき, システム J_i の寿命データからは, 第2章, 第3章で述べた方法により, 完全な競合データはもとより, 不完全な競合データに対しても, コンポーネント i の寿命分布 $f(t; \theta_i)$ のパラメータベクトル θ_i のMLEを得ることができる ($i = 1, \dots, m_0$). 一方, コンポーネント i の寿命データから, そのパラメータベクトル θ_i のMLEを求めることは, 既存の方法により容易になされる ($i = m_0 + 1, \dots, m$). 各コンポーネントの信頼度 $R_i(t; \theta_i)$ は, パラメータベクトル θ_i の関数であるから, 信頼度のMLEを求めることも容易である ($i = 1, \dots, m_0, m_0 + 1, \dots, m$). そしてこれらを式(5.2)の右辺に代入すれば, 新たに対象とするシステムの信頼度に対するMLEも導くことができる. また, 5.3節にみたように信頼度以外の寿命特性についても, 理論的にはすべて信頼度関数により表現されるものであるから, 少なくとも解析的にはMLEが求まることになる.

ノンパラメトリック推定においても, 信頼度関数を中心とした推定が有効な手段となる. このとき, たとえばMTTFの推定で言えば, 式(5.18)でなく式(5.19)による表現が重要な意味をもってくる. 完全データ(打ち切りなし)においては, 2つの式は全く同等であるが, 式(5.18)は完全データにおいてのみ成立するものであるのに対し, 式(5.19)はセンサードデータ(ランダムセンサードデータも含む)へ拡張されるものである. このことは, システム J_i の寿命データがコンポーネント i にとってはランダムセンサードデータになっているという事実をみれば, その意義が明らかである. すなわち, m_0 個のコンポーネントについては, 得られている競合データより, 第4章で述べた

K-M推定量, もしくは不完全な競合データのために修正されたK-M推定量を求め, それを式(5.19)の経験信頼度関数 $R_{E_i}(t)$ の代わりに用いればよい. このようにノンパラメトリック推定の場合も, 式(5.2)に基づいて競合データとコンポーネントデータを併合した推定が可能となる.

5.5 結言

本章では, 一品生産のシステムや大規模かつ複雑なシステムで製造コストの高いものを対象として, コンポーネントの寿命データからシステム全体の寿命特性を効率的に推定する方法について考察し, 以下の結果を得た[32][36].

- 1) 寿命特性として信頼度を取り上げた場合については, 従来の直列-並列システム, k -out-of- m システムでの結果を拡張してコヒーレントシステムにおける推定法を示し, かつ従来の結果を包含する統一的表現を与えた. また, 最尤推定量と一様最小分散不偏推定量の比較において, とくに高信頼度のところでは, ここで与えた一様最小分散不偏推定量が優れた性質をもっていることが明らかになった.
- 2) 寿命特性としてMTTFを取り上げた場合については, 最尤推定量, 一様最小分散不偏推定量ならびにノンパラメトリック推定量(U -統計量)が, 1)の結果に基づき, コヒーレントシステムに対して統一的な表現で与えられることを示した. また, 推定精度, 計算の手間, 頑健性なども総合的に考えたとき, ここで提示したノンパラメトリック推定量である U -統計量が優れた性質をもっていることが明らかになった.

- 3) フィールドにおけるシステムの寿命データから競合モデルを用いた解析により得られたコンポーネントの寿命情報を 1), 2) の推定問題へ適用し, コンポーネントレベルでの寿命試験データと併合する方法を, パラメトリックおよびノンパラメトリック推定について明らかにした. これにより, 新製品ならびに一品生産システムの開発段階における信頼性特性の効率的な推定が可能となる.

第6章 故障メカニズムモデルと構造関数の推定

6.1 緒言

システムの信頼性を定量的に把握し、信頼性の向上のために様々な対策をとる上で、システムの故障メカニズムを明らかにすることは必要不可欠である。システムの故障には、一般に多くの故障モード（症状）が存在し、各々の故障モードに対してその原因となる故障コンポーネントの集合が対応している。システムの企画、設計段階では、FMEAにより、まずシステムをその機能からいくつかのブロック（サブシステム）に分け、各ブロックごとに故障モードを列挙し、さらに各故障モードに対応する故障原因を網羅的に調べていく。またFTAにより、故障原因となる各コンポーネントがどのように関連しているかを定性的に把握する。FTAにおいて基になる構造は言うまでもなく第5章で取り扱ったツリー構造である。FTAにおいては、トップ事象としてシステムの故障を取り上げた総括的な解析から、各故障モードもトップ事象にする詳細な解析まで目的に応じて使い分けることが肝要である。

このように企画、設計の段階でシステムの故障メカニズムを調べた後、信頼性試験へと移行するわけだが、信頼性試験において予期せぬ故障モードが現れたり、また1つの故障モードが予期せぬ故障コンポーネントの集合から生まれることがある。このような場合には、故障物理など固有技術を中心とした方法で、故障の中身を詳細に検討し、そのメカニズムを解明しなければならない。

また、ある程度明らかになれば再現試験によって確認する必要がある。

一方、フィールドにおける使用段階においても、様々な使用環境や使用仕方から予期せぬ故障モードや故障コンポーネントの集合が得られることがある。従来はこの場合においても、1つ1つの故障の中身を固有技術を中心に検討するという方法がとられており、いわゆる統計的な見方がなかったとよい。もちろん固有技術なしに故障メカニズムの解明は不可能であるが、統計的方法は真の原因、真の構造を明らかにするプロセスを効率的に進めるための有効な手段になり得ると考えられる。故障時間データと同様、故障モードと故障原因に関するデータも、フィールドにおいては統計的方法を用いるだけのデータ数がそろわない場合も少なくない。

このような観点から本章では、故障モードや故障原因といった質的データから故障メカニズムを同定するための統計的方法について考察する。3.2節では、様々な故障メカニズムモデルと信頼性データを整理分類する。3.3節では、推定のためのモデルとして確率的コヒーレントシステムを導入し、その性質を明らかにする。さらに3.4節では、構造関数の推定法を提示する。

6.2 故障メカニズムモデルと信頼性データ

故障メカニズムモデルは、決定的モデルと確率的モデルとに大別される。コヒーレント構造は1つの合理的な決定的故障メカニズムモデルと考えられる。このとき、故障コンポーネントの集合であるカット集合（あるいは最小カット集合）も1つの故障モードと対応させて考えることもできるし、1つの故障モードに対してコヒーレント構造を仮定し複数の故障コンポーネントの集合との

関係を表わすこともできる。コヒーレント構造は、システムおよび各コンポーネントに対して2状態を考えたものから、全順序集合を形成する多状態を考えた多状態コヒーレントシステムへと研究が進められている。

コヒーレントシステムモデルの特徴はオ1に、コンポーネントの状態とシステムの状態との間に決定的な関数関係があることである。オ2は、単調性すなわち負の作用をするコンポーネントがないことである。オ3には、すべてのコンポーネントが関連性(relevance)をもつこと、すなわちシステムの状態に全く関与しないコンポーネントがないことである。近年、ノンコヒーレントシステムの研究がなされているが(たとえば[19]を参照)、それらは主にオ2の特徴を否定したものである。すなわち複数個のコンポーネントの故障状態の組合せがシステムの状態に対して正の作用を施す場合を表現している。コンポーネントの状態の定義の仕方によっては、確かにそのような場合もありうるが、コヒーレントシステムモデルを現実のシステムにあてはめるとき、最も大きな障害になるのはむしろオ1の特徴、すなわち決定的な関数関係にあると思われる。コンポーネントの状態ならびにシステムの状態も連続的に把握でき、かつシステムを取り巻く環境条件が完全に把握できれば、両者の間には物理的に決定的な関数関係が存在するはずであるが、現実には状態は離散的な有限状態でしか把握できないことが多く、状態の識別にも測定誤差を含めて不確実性が残る。そのため、コンポーネントの状態とシステムの状態との間に再現性がない場合が生じてくる。

一方、確率的故障メカニズムモデルとしては、システムもしくはユニットに対していくつの離散的な劣化状態を考え、各状態間の推移確率ならびに各状態から最終状態である故障状態への推移確率を与える確率的劣化モデルがある。推移確率を一定で既知としたマルコフ劣化モデルにおける点検取替ルールにつ

いては数多くの研究がなされているが、これら劣化モデルの適合性や推移確率の推定などの問題については、ほとんど手がつけられていないのが現状である。

本章では、これら一連の故障メカニズムモデルを広範囲にわたって統一的に表現するものとして、確率的コヒーレントシステムモデルを導入し、このモデルのもとで故障メカニズムをデータより探索していく方法について考察する。

そこでまず、得られるデータについて整理してみる。信頼性データは図6.1のように原因系と結果系に大別されよう。

信頼性における統計的方法は、初期においては故障時間だけを扱い、しかも寿命分布の母数推定・検定に主眼を置いた単一母集団分析が中心であった。一方、故障物理、再現試験では、1つの故障事象に対して、図6.1にあるすべてのデータから（統計的ではなく）物性面からの解明を試んでいる。第2章から第4章で論じた競合モデルによる統計的方法や、比例ハザードモデルなどはこれらも結びつけることも意図して発達したものと見えるが、依然として故障時間を中心とした解析方法である。

ところで、故障モードとは本来「症状」を指す言葉であるが、この症状と原因（故障しているコンポーネント）が完全に対応づけられていれば、これらを

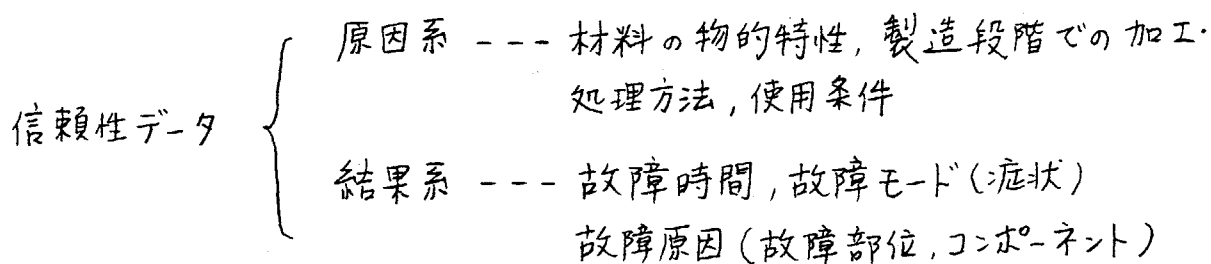


図 6.1 信頼性データの分類

同一視することができ、第2章から第4章では主にこの立場をとってきた。これに対し、1つの故障モードに対して複数の故障原因が存在する場合があるが、そのモードに限定すれば、この関係はシステムの故障に対し複数の故障モード（もしくはこれに対応する故障原因）が存在する場合に置き換えて考えることができ、同一の枠組みの中で議論ができる。

故障原因データは、故障コンポーネントの集合という型で与えられることが多いが、故障時間が伴うことにより、その情報の程度から次の3つに分類できる。

- 1) システムが故障するまでに故障したコンポーネントとその故障時間がすべて観測されているデータ。
- 2) システムの故障時点で故障したコンポーネントと、それ以前に故障したコンポーネントとが区別して観測されるが、後者についてはその間の故障順序や故障時間は不明なデータ。ただし、システムの故障時間は観測されている。
- 3) システムの故障時点で故障したコンポーネントと、それ以前に故障したコンポーネントとが区別されずに観測され、その間の故障順序や故障時間は不明なデータ。ただし、システムの故障時間は観測される。

ところで、第2章から第4章までに論じた競合モデルは直列システムモデルと同等であり、コヒーレントシステムモデルの特別な場合として位置付けられる。一般に対象とするシステムがコヒーレントシステムである場合に対する競合理論の適用可能性について考えてみる。まず、1)のデータが観測されたならば、コンポーネントレベルにおいては故障時間もしくは打ち切り時間が観測されることになり完全な競合データとなる。したがって既存の競合理論の枠組みの

中で解析が可能である。次に、2) であるがこの場合は、コンポーネントレベルでみれば、故障時間（システム故障時点まで故障したとき）と打ち切り時間（システム故障時点で未故障のとき）に加えて、見逃し時間（システム故障時点以前に故障したとき）が観測されることになる。さらに 3) の場合には、打ち切り時間と見逃し時間となる。2), 3) に対するパラメトリックな尤度構成は容易であり、最尤推定値を求める上での数値計算上の困難は伴うにしても理論上は推定可能である。ノンパラメトリックの場合にも、Turnbull [50] が与えた部分集合データ（区間データの一般化）に基づく分布関数の推定法（これはEMアルゴリズムの応用である）により、各コンポーネントの寿命特性の推定が可能となる。したがって、競合モデルに基づく方法は一般のコヒーレントシステムにおいてもかなりの範囲で適用されることがわかる。

6.3 確率的コヒーレントシステムの定義とその性質

6.3.1 定義

m 個のコンポーネントからなるシステムを考える。コンポーネント i の状態を示す 2 値変数を

$$x_i = \begin{cases} 1 & \text{--- コンポーネント } i \text{ が正常のとき} \\ 0 & \text{--- コンポーネント } i \text{ が故障のとき} \end{cases}$$

と定義する ($i=1, \dots, m$)。同様にシステムの状態を表わす 2 値変数を

$$\phi = \begin{cases} 1 & \text{--- システムが正常のとき} \\ 0 & \text{--- システムが故障のとき} \end{cases}$$

と定義する。コンポーネントの状態を表わすベクトル $\mathbf{x} = (x_1, \dots, x_m)$ に対して $C_0(\mathbf{x}) = \{i \mid x_i = 0\}$, $C_1(\mathbf{x}) = \{i \mid x_i = 1\}$ と定義する。

そこで、システムならびにコンポーネントの確率的挙動について考えてみる。 $\{X_i(t), t \geq 0\}$ を $\{0, 1\}$ 上の値をとる右連続、非増加の確率過程とする。すなわち $\{X_i(t), t \geq 0\}$ はコンポーネント i の挙動を表わしている。さらに、 m 個のコンポーネントの挙動を示す多変量確率過程を $\{X(t), t \geq 0\} = \{(X_1(t), \dots, X_m(t)), t \geq 0\}$ で表わす。同様に、 $\{\Phi(t), t \geq 0\}$ をシステムの挙動を表わす $\{0, 1\}$ 上の値をとる右連続、非増加の確率過程とする。

このとき $\{\Phi(t), t \geq 0\}$ は $\{X(t), t \geq 0\}$ に以下のように依存する二重確率過程 (Doubly stochastic process) とする。オ1に、 $\{\Phi(t), t \geq 0\}$ で推移が生じうる時点は $\{X(t), t \geq 0\}$ の推移時点のみである。オ2に、 $\{X(t), t \geq 0\}$ の推移時点で $\{\Phi(t), t \geq 0\}$ の推移が生ずる確率は、その時点での $X(t)$ の状態 \mathbf{x} にもみ依存し、それまでの $\{X(t), t \geq 0\}$ の履歴や経過時間には依存しないものとする。すなわち $\{X(t) \geq 0\}$ のある推移時点も t^* と記せば、この確率は

$$P_c [C_0(\mathbf{x})] \triangleq P_r [\Phi(t^*) = 0 \mid X(t^*) = \mathbf{x}, \Phi(t^{*-0}) = 1] \quad (6.1)$$

と表わされるものとする。ここで $\Phi(t^{*-0}) = \lim_{\varepsilon \rightarrow 0} \Phi(t^* - \varepsilon)$ である。これを状態 \mathbf{x} におけるカット集合確率と呼ぶことにする。すなわちこれは、 $C_0(\mathbf{x})$ がカット集合になる確率と考えられる。また、便宜上

$$P_p [C_1(\mathbf{x})] \triangleq 1 - P_c [C_0(\mathbf{x})] \quad (6.2)$$

を定義し、これをパス集合確率と呼ぶことにする。

[定義 6.1]

コンポーネント i が確率的に関連性をもたない (stochastically irrelevant) とは、任意の (\cdot, \mathbf{x}) に対して

$$P_c [C_o(1_i, \mathbf{x})] = P_c [C_o(0_i, \mathbf{x})] \quad (6.3)$$

が成り立つことをいう。ここで

$$(\cdot, \mathbf{x}) = (x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_m)$$

である。そうでないときには、確率的に関連性をもつ (stochastically relevant) という。

[定義 6.2]

m 個のコンポーネントからなるシステムが次の 2 条件を満たすとき、確率的コヒーレントシステム (stochastic coherent system) であるという。

- 1) $P_c [C_o(\mathbf{x})]$ は \mathbf{x} の各要素について非増加である。
- 2) すべてのコンポーネントは確率的に関連性をもつ。

[定義 6.3]

確率的コヒーレントシステムにおけるコンポーネント i の相対重要度 (relative importance) を次式で定義する。

$$I_m(i) \triangleq \frac{1}{2^{m-1}} \sum_{(\cdot, \mathbf{x})} \{ P_p [C_1(1_i, \mathbf{x})] - P_p [C_1(0_i, \mathbf{x})] \}$$

このように定義した確率的コヒーレントシステムの特性について、以下考察する。

6.3.2 故障過程と状態の分類

まず、確率的コヒーレントシステムの故障過程について考えてみる。コンポーネントの初期状態を $x_0 = (1, \dots, 1)$ とする。便宜上、 $P_c[C_0(x_0)] = 0$ とするが、この仮定は自然であろう。この状態からあるコンポーネントの故障（もしくは複数個のコンポーネントの同時故障）により状態 x_1 へ推移したとき、 $P_c[C_0(x_1)]$ の確率でシステムの故障状態へ推移し、 $1 - P_c[C_0(x_1)]$ の確率でシステムの正常状態にとどまる。後者の場合には、次のコンポーネントの故障により x_2 へ推移し、このとき $P_c[C_0(x_2)]$ ($\geq P_c[C_0(x_1)]$) の確率でシステムは故障状態へ推移し、 $1 - P_c[C_0(x_2)]$ の確率で正常状態にとどまる。以下同様のプロセスを続け、各コンポーネントの信頼度関数 $R_1(t), \dots, R_m(t)$ がいずれも $t \rightarrow \infty$ で 0 に収束するならば、システムはいずれは故障状態へ推移する。

ところで、ベクトル $x = (x_1, \dots, x_m)$ はシステムを構成するコンポーネントの状態を表わすものだが、確率的コヒーレントシステムにおけるコンポーネントの状態は次の3つに分類できる。

- 1) $P_c[C_0(x)] = 1$ となるコンポーネント状態 x を決定的致命状態 (deterministic fatal state) と呼ぶ。この x の集合を Ω_d と記す。
- 2) $0 < P_c[C_0(x)] < 1$ となるコンポーネント状態 x を確率的致命状

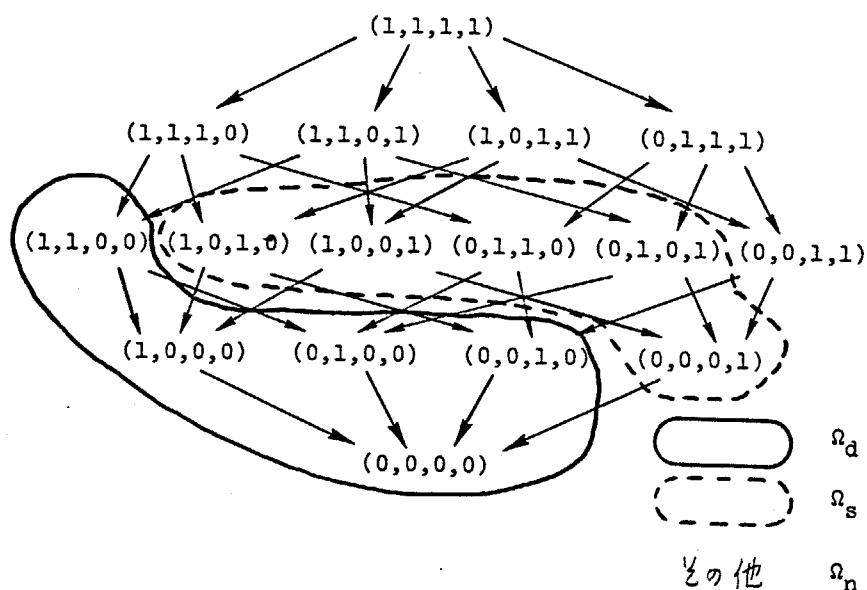


図 6.2 状態の分類と推移の例

態 (stochastic fatal state) と呼ぶ。この X の集合を Ω_s と記す。

3) $P_c[C_o(X)] = 0$ となるコンポーネント状態 X を非致命状態 (non-fatal state) と呼ぶ。この X の集合を Ω_n と記す。

図 6.2 は、 $m=4$ の確率的コヒーレントシステムにおける状態の例を示している。矢印は、複数コンポーネントの同時故障がないときの可能な推移を表わしている。

6.3.2 信頼度表現と特別な場合

確率的コヒーレントシステムの時刻 t での信頼度は、初期状態 $X_0 = (1, \dots, 1)$ から時刻 t までにシステムの故障状態に至る経路 (path) の確率を評価することで求まる。コンポーネントの寿命が独立で、それぞれ一定の故障

率 $\lambda_1, \dots, \lambda_m$ をもつ場合には, 以下に示すような推移率行列を用いてその信頼度が表現される.

いま, Ω_n, Ω_s に含まれる状態とシステムの故障状態 (これを f と記す) からなる状態ベクトルを考え, 時刻 t におけるこの状態確率ベクトルを $\pi(t)$ で表わす. ただしここで, Ω_s に属する状態に対する確率は, その状態にあってしかもシステムが正常状態にある同時確率を表わすものとする.

$\pi(t)$ に対応する推移率行列を A と記せば, A の要素は以下のように定まる.

いま, 1つの状態 x より1つのコンポーネントの故障により推移可能な決定的致命状態の集合, 確率的致命状態の集合, および非致命状態の集合を, それぞれ $\Omega_d(x), \Omega_s(x)$, および $\Omega_n(x)$ と記せば

a) $x \in \Omega_n$ から $x^* \in \Omega_n$ への推移率は, $x^* \in \Omega_n(x)$ であれば λ_j (ただし j は $C_1(x) - C_1(x^*)$ の要素), それなければ 0 である.

b) $x \in \Omega_n \cup \Omega_s$ から $x^* \in \Omega_s$ への推移率は, $x^* \in \Omega_s(x)$ であれば $\{1 - P_c[C_0(x^*)]\} \cdot \lambda_j$ (ただし j は $C_1(x) - C_1(x^*)$ の要素), それなければ 0 である.

c) $x \in \Omega_n \cup \Omega_s$ から f への推移率は

$$\sum_{x^* \in \Omega_s(x)} P_c[C_0(x^*)] \lambda_j + \sum_{x^0 \in \Omega_d(x)} \lambda_k$$

である. ただし j, k は, それぞれ対応する $C_1(x) - C_1(x^*)$, $C_1(x) - C_1(x^0)$ の要素である.

d) $x \in \Omega_n \cup \Omega_s$ に対する A の対角要素は

$$- \left\{ \sum_{x^* \in \Omega_n(x) \cup \Omega_s(x)} \lambda_j \right\}$$

で与えられる。ただし j は対応する $C_1(x) - C_1(x^*)$ の要素である。

e) 上記以外の要素はすべて 0 である。

この行列 A を用いれば、状態確率ベクトル $\pi(t)$ に対して方程式

$$\frac{d}{dt} \pi(t) = \pi(t) A \quad (6.4)$$

が成り立つ。式(6.4)の解の一般的表現として

$$\pi(t) = \pi(0) e^{At} \quad (6.5)$$

を得るが、 $\pi(t)$ において Ω_n , Ω_s に属する状態確率の和が、時刻 t における信頼度となる。式(6.5)から実際に信頼度の値を求めることは一般に難しいわけだが、信頼度計算はここでの主目的でないので、一般的表現を示すにとどめる。

次に、確率的コヒーレントシステムが表現するいくつかの典型的故障メカニズムモデルについて考察する。当然のことながら、通常のコヒーレント構造は確率的コヒーレント構造において確率的致命状態がない場合である。この他に興味深いものとして、次の2つを挙げておく。

a) 確率的劣化モデル

$N+1$ 個の離散的な劣化状態 $\pi_0, \pi_1, \dots, \pi_N$ を想定し、 π_0 を新品状態、

π_N を故障状態とし, 各状態間の推移確率 λ_{ij} が与えられている確率的劣化モデルが取替問題ではよく論ぜられる. ここで $j < i$ に対して $\lambda_{ij} = 0$, かつ λ_{iN} が i について非減少という仮定が通常なされるが, このモデルにおいて, π_i から π_N への推移を, π_i からいったん π_{i+1} へ推移し同時に π_{i+1} から π_N へ推移するとみなせば, この確率的劣化モデルはまさに確率的コヒーレントシステムである. すなわち, 劣化状態を故障コンポーネント数に対応させ, $P_c[C_0(x)]$ が $C_0(x)$ の要素数のみで決まる対称的な確率的コヒーレントシステムと同等である. また, このモデルはコヒーレント構造における *k-out-of-m* システムの確率的な拡張とみなせる.

b) 混合モデル

異なる故障モードをもつ2種類のシステムが p , $1-p$ の割合で混合モデルは競合モデルと対比してよく用いられる. 競合モデルでは, 個々のシステムが2種類の故障モードを潜在的に所有するものであり, これは直列システムモデルとして表わされるわけだが, これに対し混合モデルでは, それぞれ固有の1つのモードでのみ故障するシステムが2種類混合している. 確率的コヒーレントシステムは, 混合モデルにおける母集団の確率的挙動を表現できるという意味で混合モデルを含んでいる.

いま, 2つの独立なコンポーネントからなる確率的コヒーレントシステムを考へ, カット集合確率をそれぞれ $P_c[C_0(1,1)] = 0$, $P_c[C_0(1,0)] = 1-p$, $P_c[C_0(0,1)] = p$, $P_c[C_0(0,0)] = 1$ とすれば, コンポーネントの信頼度 $R_1(t)$, $R_2(t)$ に対して, システムの信頼度は

$$R(t) = R_1(t) R_2(t) + p R_1(t) [1 - R_2(t)] \\ + (1-p) R_2(t) [1 - R_1(t)]$$

$$= p R_1(t) + (1-p) R_2(t)$$

となり，混合モデルの信頼度関数に一致する。

6.4 構造関数の推定

6.4.1 コヒーレントシステムの構造関数の推定

この節では，6.2節に述べた故障原因データより故障メカニズムを探索，推定する方法について述べる。まずコヒーレント構造を仮定した上で，構造関数を推定する方法を述べるが，この方法は確率的コヒーレントシステムに対しても基礎的役割を果たす。

いま， n 個の同種システムを観測し，6.2節で分類した故障原因データのうち，1)もしくは2)のデータを得たものとする。ただしここでは故障順序が十分な情報であり，故障時間そのものは不要である。 j 番目のシステムに対して，システム故障時のコンポーネントベクトルを x_{0j} ，システム故障直前のコンポーネントベクトルを x_{1j} とそれぞれ記せば， $C_0(x_{0j})$ ， $C_1(x_{1j})$ がそれぞれ観測されたカット集合，パス集合となる（必ずしも最小カット集合，最小パス集合ではない）。このとき，次の命題が成り立つ。

[命題 6.1]

n 個のシステムについて， $C_0(x_{0j})$ ， $C_1(x_{1j})$ ($j=1, \dots, n$) が得られたとき，構造関数 $\phi(x)$ に対して次の不等式

$$\prod_{j=1}^n \prod_{i \in C_1(x_{1j})} x_i \leq \phi(x) \leq \prod_{j=1}^n \prod_{i \in C_0(x_{0j})} x_i \quad (6.6)$$

が成り立ち, 下限

$$\hat{\phi}_{L,n}(x) = \prod_{j=1}^n \prod_{i \in C_1(x_{1j})} x_i$$

は n に関して非減少であり, 上限

$$\hat{\phi}_{U,n}(x) = \prod_{j=1}^n \prod_{i \in C_0(x_{0j})} x_i$$

は n に関して非増加である. ここで $\prod_{i=1}^n x_i = 1 - \prod_{i=1}^n (1-x_i)$, $x_i^2 = x_i$ である.

証明は自明であるから省略するが, これより式(6.6)の下限 $\hat{\phi}_{L,n}(x)$ および上限 $\hat{\phi}_{U,n}(x)$ をそれぞれ $\phi(x)$ の推定値として用いることが考えられる. なお, 当然のことながら上限, 下限が一致したときには, それは $\phi(x)$ に等しい. $\hat{\phi}_{L,n}(x)$, $\hat{\phi}_{U,n}(x)$ の性質については次の命題が成り立つ.

[命題 6.2]

コンポーネントの故障率関数 $\lambda_1(t), \dots, \lambda_m(t)$ がそれぞれある有限区間 $(0, T]$ で正の値をとるならば, $\hat{\phi}_{L,n}(x)$, $\hat{\phi}_{U,n}(x)$ はともに, $n \rightarrow \infty$ では, 確率 1 で真の構造関数に一致する.

一方, 6.2 節に述べた 3) のデータが得られたときには, システム故障時に故障したコンポーネントがそれ以前に故障したものと区別されていないので, x_{0j} は観測されるが, x_{1j} は得られない. そのためこの場合には式(6.6)の上限 $\hat{\phi}_{U,n}(x)$ だけが与えられることになる.

6.4.2 カット集合確率の推定

ここでは, 確率的コヒーレント構造を仮定した故障原因データの解析法について考察する. このとき中心となるのは状態の分類であり, 確率的致命状態が現われない限り, コヒーレント構造のもとでの構造関数の推定法がそのまま使えることになる.

いま, 6.2 節に述べた 1) のデータが観測されるものとする. すなわち, コンポーネントの状態の推移がシステムの故障まで絶えず観測されるものとする. このとき j 番目のシステムに対し, p_j 個のシステム正常時のコンポーネント状態 $x_{1j1}, \dots, x_{1jp_j}$ とシステム故障時のコンポーネント状態 x_{0j} を得る ($j=1, \dots, n$).

そこで, システム正常時のコンポーネント状態の集合を S_p とする. すなわち

$$S_p = \{ x_{1jk}, k=1, \dots, p_j, j=1, \dots, n \}$$

である. また S_c をシステム故障時のコンポーネント状態の集合とする. すなわち

$$S_c = \{ x_{0j}, j=1, \dots, n \}$$

である。 S_p と S_c よりコンポーネント状態の分類がなされる。 $S_c - S_p$ は観測された決定的致命状態の集合であり、これに $y < x$, $x \in S_c - S_p$ なる y を加えたものを $\hat{\Omega}_d$ とする。そして $\hat{P}_c [C_0(x)] = 1$; $x \in \hat{\Omega}_d$ とする。また $S_p - S_c$ は観測された非致命状態の集合であり、これを $\hat{\Omega}_n$ とする。そして $\hat{P}_c [C_0(x)] = 0$; $x \in \hat{\Omega}_n$ とする。さらに $S_p \cap S_c$ が観測された確率的致命状態の集合であり、これを $\hat{\Omega}_s$ とする。 $x \in \hat{\Omega}_s$ なる x については

$$\tilde{P}_c [C_0(x)] = \frac{\text{システム故障時に } x \text{ が観測された回数}}{\text{コンポーネント状態として } x \text{ が観測された回数}}$$

を基本推定量 (basic estimator) として、単調性 (定義 6.2 の 1)) が成り立つようにフーリング修正 (たとえば [2] の第 2 章参照) を施した $\hat{P}_c [C_0(x)]$ が順序制約下での 2 項パラメータの最尤推定量として導かれる。この $\hat{\Omega}_s$ に含まれるコンポーネント状態は、システムの状態に再現性のないものであるので、とくにこの状態については故障の中身、使用環境や使い方の差異などを固有技術的に究明することが肝要であろう。

$\hat{\Omega}_s$, $\hat{\Omega}_d$, $\hat{\Omega}_n$ のいずれにも含まれていない x については $\hat{P}_c [C_0(x)]$ の推定はできないが、信頼度の評価においては、データ数 n が十分大きければ $\hat{P}_c [C_0(x)] = 0$ として差し障りないと思われる。

6.5 結 言

本章では、第 2 章から第 5 章までに述べた一連の統計的方法の適用範囲を拡張することを目的として、様々な故障メカニズムモデルと信頼性データを整理

分類し、新たな故障メカニズムモデルを導入するとともに、故障原因データから故障メカニズムを探索・推定する方法について考察した。ここで得られた結果は次の通りである[37]。

- 1) 決定的故障メカニズムモデルとしてコヒーレント構造を取り上げたとき、第2章から第4章で論じた競合モデルを適用するために必要となる信頼性データの型を明らかにし、その適用の可能性が十分大きいことを示した。
- 2) コヒーレント構造を仮定したときに、故障原因データから構造関数を推定する方法を与えた。これにより、第5章で論じたコンポーネントの寿命データからコヒーレントシステムの寿命特性を推定する問題を、構造関数未知の場合にも拡張するための指針を得た。
- 3) 様々な決定的および確率的故障メカニズムモデルをなるべく統一的に表現するものとして、確率的コヒーレントシステムを導入し、その性質を明らかにした。さらにコンポーネント状態のデータより、このシステムの構造を探索・推定する方法を示し、効率的な故障メカニズム究明のための緒を与えた。

第7章 ワイブル確率紙による母数の簡易推定法

7.1 緒言

システムの信頼性特性の評価においては、各故障モードもしくは各コンポーネントの寿命分布としてワイブル分布を仮定して統計的に解析することが多い。これはワイブル分布自体が様々な寿命データによくあてはまるという経験的事実にも依るが、ワイブル分布の形状母数、尺度母数という2つの母数の値を知ることが工学的に大至な意味をもっているからである。

一般に、システムにはいくつもの故障モードが存在し、その寿命分布も複雑なことが多いが、故障モードやコンポーネントのレベルで解析も行ない、それらが個々には初期故障型、偶発故障型、摩耗故障型のいずれのタイプに属するのかを知ることが、システムの運用上はもちろん改善の上でも極めて有用である。この際、ワイブル分布を個々の故障モードやコンポーネントの寿命分布にあてはめ、とくにその形状母数の値を知ることが有効な手段となる。このときシステムの寿命データを観測する場合については第3章で論じた。本章では第5章のようにコンポーネントの寿命データ、すなわち直接ワイブル分布に従う寿命データの解析を論ずるものである。

2母数ワイブル分布はいわゆる指数型分布族に属していないことと自明な十分統計量しか存在しないことなどから、母数推定は正規分布や指数分布の場合に比べ面倒である。たとえば最尤法を用いる場合にも第3章に述べたように連

立関数方程式を計算機による繰り返し計算で解かなければならない。そのため、母数の簡易推定法としてモーメント法やワイブル確率紙による推定法などが提案されてきた。また、データを対数変換したものの線形結合を推定量とする方法である最良線形不偏推定量、最良線形不変推定量は、2母数ワイブル分布の母数そのものの推定量にはあてはまらないことに注意すべきである。このように様々な推定量が提案されているが、これらの性質が明らかなのは漸近的な性質だけであり、小標本での正確な評価はほとんどなされていない。これに対し近年、柴田⁽⁴⁵⁾や信頼性研究会⁽⁴⁶⁾は各種推定量の小標本での性質をシミュレーションで調べ報告している。

本章では、ワイブル分布の母数の簡易推定法としてワイブル確率紙による推定法を取り上げ、その性質を計算機による数値積分で評価できる範囲において調べ、その中で推定精度と実用面から最も望ましいと思われる直線の引き方、すなわち推定法を与えることを目的とする。ワイブル確率紙は第3章で示したように得られたデータがワイブルにあてはまるかを確認、競合型あるいは混合型として扱うべきかなどを検討する上でも極めて有用であるが、ここでは母数推定に限ってワイブル確率紙をみたときの小標本での偏り、分散、平均2乗誤差について調べ、打点の方法や直接の引き方による推定精度の比較を試みる。

7.2 ワイブル確率紙の数理的構造

ワイブル分布の分布関数を

$$F(t) = 1 - \exp\left\{-\left(\frac{t}{\theta}\right)^m\right\} \quad (7.1)$$

と表わす. ここで, m は形状母数, θ は尺度母数と呼ばれるものである. ワイブル確率紙は式(7.1)が

$$\ln \ln \frac{1}{1-F(t)} = m \ln t - m \ln \theta$$

の関係を満たすことから, たとえば式(7.1)の分布にしたがう母集団からの完全データ

$$t_1 \leq t_2 \leq \dots \leq t_n$$

が得られるならば, n 個の点,

$$\left(\ln t_i, \ln \ln \frac{1}{1-\hat{F}(t_i)} \right) \quad (i=1, \dots, n)$$

を打ち, n 点に直線をあてはめ, その傾きと切片から m と θ を推定するものである. ここで $\hat{F}(t_i)$ は n パラメトリックに推定され, 通常平均ランク

$$\hat{F}(t_i) = \frac{i}{n+1}$$

がよく用いられる. この他にも X ジアンランク, 対称サンプル累積ランクなどがある.

また T が式(7.1)にしたがうとき, $X = \ln T$ は極値分布

$$F(x) = 1 - \exp \left\{ -\exp \left(\frac{x-u}{b} \right) \right\} \quad (7.2)$$

にしたがう。ここで $b = 1/m$, $u = \ln \theta$ の関係がある。式(7.2)は位置母数と尺度母数をもつ分布族に属しているので、 u と b の最良線形不偏推定量や最良線形不変推定量が求められ Lieblein and Zelen⁽²⁶⁾, Mann⁽²⁷⁾ らによってその係数が与えられている。また Bain⁽¹⁾ によっていくつかの簡易線形推定量とそのための係数が与えられている。

いま式(7.2)をもとにワイブル確率紙を見直すと、次のように考えられる順序統計量

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

に対して

$$U_{(i)} = \frac{X_{(i)} - u}{b} \quad (i=1, \dots, n)$$

と置くと、 $U_{(i)}$ の分布は母数によらず

$$E(U_{(i)}) = \alpha_i \quad (i=1, \dots, n)$$

$$\text{Cov}(U_{(i)}, U_{(j)}) = v_{ij} \quad (1 \leq i \leq j \leq n)$$

は母数と無関係な定数となる。そこで $X_{(i)} = u + bU_{(i)}$ の関係から

$$\begin{pmatrix} X_{(1)} \\ \vdots \\ X_{(n)} \end{pmatrix} = \begin{pmatrix} 1 & \alpha_1 \\ \vdots & \vdots \\ 1 & \alpha_n \end{pmatrix} \begin{pmatrix} u \\ b \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (7.3)$$

と表わせば

$$E(\varepsilon) = 0$$

$$E(\varepsilon\varepsilon') = b^2V$$

を満たす。ただし $\varepsilon' = (\varepsilon_1, \dots, \varepsilon_n)$ であり、 $V = (v_{ij})$ は $(U_{(1)}, \dots, U_{(n)})'$ の分散共分散行列である。ワイブル確率紙は式(7.3)の関係をもとに、縦軸に説明変数を取り、横軸に目的変数として変量 $x_i = \ln t_i$ をとったものと考えることができる。この考え方からすれば、少なくとも μ と b の推定については縦軸に $\alpha_1, \dots, \alpha_n$ をとることが自然であるが、通常は平均ランクに基づいて

$$\begin{aligned} \beta_i &= \ln \ln = \frac{1}{1 - \hat{F}(t_i)} \\ &= \ln \ln \frac{n+1}{n+1-i} \quad (i=1, \dots, n) \end{aligned}$$

をとる。 α_i の値は Lieblein and Zelen⁽²⁶⁾ が $n=2 \sim 6$ について与えているが、もとのワイブル分布の母数 m, θ の推定が目的であれば α_i をとることが良いとは限らないわけである。本章では、以下ワイブル確率紙による推定に関して、7.3節において極値分布の母数 μ, b の推定について述べ、7.4節ではワイブル分布の形状母数 m の推定について述べる。

7.3 極値分布の位置母数と尺度母数の推定

確率紙による推定の精度についてはこれまでには, Chernoff and Lieberman [12], Kimball [24]らが調べているが, これらの結果はいずれも「直線は最小2乗法で引かれたもの」という仮定で議論されたものである. すなわち, すべての点に同じ重みを与えたときに残差平方和を最小にする直線から定まる推定量の性質を調べているわけである. このときの残差は, 変量の軸(横軸)に平行な距離で測ったものである.

ワイブル確率紙において, 縦軸に $\alpha_1, \dots, \alpha_n$ をとり, 横軸方向の距離について最小2乗法で直線を引いたならば, 得られる極値分布の母数 u, b の推定量は

$$\begin{pmatrix} \tilde{u} \\ \tilde{b} \end{pmatrix}_A = (A'A)^{-1} A' X \quad (7.4)$$

となる. ここで

$$A = \begin{pmatrix} 1 & \alpha_1 \\ \vdots & \vdots \\ 1 & \alpha_n \end{pmatrix}, \quad X = \begin{pmatrix} X_{(1)} \\ \vdots \\ X_{(n)} \end{pmatrix}$$

である. 式(7.4)は不偏でありその分散共分散行列は

$$\text{Var} \begin{pmatrix} \tilde{u} \\ \tilde{b} \end{pmatrix}_A = (A'A)^{-1} A' \nabla A (A'A)^{-1} b^2$$

となる。これに対し最良線形不偏推定量は

$$\begin{pmatrix} \tilde{u} \\ b \end{pmatrix}_A^* = (A'V^{-1}A)'A'V^{-1}X \quad (7.5)$$

で与えられ、その分散共分散行列は

$$\text{Var} \begin{pmatrix} \tilde{u} \\ b \end{pmatrix}_A^* = (A'V^{-1}A)^{-1} b^2$$

となる。式(7.4)に対応する直線を引くことは計算の上では可能であるが、式(7.5)に対応する直線を引くことは困難であろう。

また最小二乗法とは別に、2つの点を結ぶ方法も考えられる。市田[18]は対数正規確率紙による推定において、この2点推定法について調べている。縦軸に $\alpha_1, \dots, \alpha_n$ をとったときの2点推定法も2点法Aと呼ぶことにする。順序統計量 $X_{(1)} \leq \dots \leq X_{(n)}$ に対し、2点法Aにより i 番目と j 番目の点を結んだときの u と b の推定量は

$$\hat{u}_A = \frac{\alpha_j X_{(i)} - \alpha_i X_{(j)}}{\alpha_j - \alpha_i} \quad (7.6)$$

$$\hat{b}_A = \frac{X_{(j)} - X_{(i)}}{\alpha_j - \alpha_i} \quad (7.7)$$

で与えられる($i < j$)。これらはともに不偏で、分散はそれぞれ

$$\text{Var}(\hat{u}_A) = \frac{b^2}{(\alpha_j - \alpha_i)^2} (\alpha_j^2 v_{ii} - 2\alpha_i \alpha_j v_{ij} + \alpha_i^2 v_{jj}) \quad (7.8)$$

$$\text{Var}(\hat{b}_A) = \frac{b^2}{(d_j - d_i)^2} (v_{ii} - 2v_{ij} + v_{jj}) \quad (7.9)$$

となる。2点推定法では、標本数 n に応じて式(7.8), (7.9) をそれぞれ最小にする i と j の組合せを求める必要がある。極値分布における2つの順序統計量の最適組合せの問題は Chan and Kabir [9] によってなされているが、これは順序統計量の漸近分布をもとに論じており、その上推定量は $a_i X_{(i)} + b_j X_{(j)}$ という形で、最適組合せと同時に係数 a_i, b_j も決めるものである。したがって、確率紙上での2点推定法とは直接対応していない。

そこで、ここでは小標本での確率紙上の2点法Aの最適組合せを与えるため、直接数値積分で行列 A, V の要素を計算し、それにより $n=5 \sim 20$ について \hat{u}_A, \hat{b}_A それぞれに対する最適組合せと係数 α_i, α_j を求めた。結果を表7.1に示す。なお、これらの計算法については付録3を参照されたい。表7.1より、 \hat{b}_A については $n=5 \sim 20$ の範囲では $X_{(j)}$ としてすべて最大値を用い、 i は標本数とともに大きくなるのがわかる。また \hat{u}_A については、 \hat{b}_A に比べなかほどのデータを使うべきであるということがわかる。ただしこれらの最適組合せは、すべて異常値を含まないという仮定のもとでのものである。

また、縦軸に平均ランクに基づく β_1, \dots, β_n をとったときの最小2乗法や2点法も考えられる(推定量の形は省略する)。これらを最小2乗法B, 2点法Bと呼ぶことにする。これに対し式(7.4)を最小2乗法A, 式(7.6)を2点法Aと呼ぶ。 u と b の推定について、これら4つの推定量の偏り, 分散, 平均2乗誤差を計算し比較したのが表7.2, 表7.3である。これらは $u=0, b=1$ のもとでの値であり、一般の場合には、偏りには b を、分散と平均2乗誤差には b^2 を乗じればよい。また、2点法については最適組合せで

表 7.1 u, b に対する 2 点法 A の最適組合せと係数

n	\hat{u}_A				\hat{b}_A			
	i	α_i	j	α_j	i	α_i	j	α_j
5	2	-1.0709	5	0.6902	1	-2.1867	5	0.6902
6	3	-0.6627	6	0.7773	2	-1.2750	6	0.7773
7	3	-0.8525	6	0.3653	2	-1.4441	7	0.8460
8	3	-1.0111	7	0.4528	2	-1.5884	8	0.9021
9	3	-1.1475	8	0.5244	2	-1.7144	9	0.9493
10	3	-1.3403	9	0.5846	3	-1.2672	10	0.9899
11	4	-0.9825	10	0.6362	3	-1.3739	11	1.0252
12	4	-1.0849	10	0.4112	3	-1.4703	12	1.0565
13	4	-1.1776	11	0.4626	3	-1.5581	13	1.0845
14	4	-1.2623	12	0.5080	3	-1.6387	14	1.1097
15	4	-1.3403	13	0.5485	3	-1.7133	15	1.1327
16	4	-1.4126	14	0.5851	4	-1.4126	16	1.1537
17	5	-1.1937	14	0.4302	4	-1.4800	17	1.1731
18	5	-1.2592	15	0.4666	4	-1.5431	18	1.1910
19	5	-1.3207	16	0.4999	4	-1.6024	19	1.2076
20	5	-1.3786	17	0.5304	4	-1.6584	20	1.2232

表 7.2 u に対する推定法の比較

n	推定法	偏り	分散	MSE
5	2点法 A	0	0.2634	0.2634
	2点法 B	-0.0010	0.2639	0.2639
	最小2乗法 A	0	0.2348	0.2348
	最小2乗法 B	0	0.2341	0.2341
10	2点法 A	0	0.1226	0.1226
	2点法 B	-0.0042	0.1362	0.1362
	最小2乗法 A	0	0.1180	0.1180
	最小2乗法 B	0	0.1174	0.1174
15	2点法 A	0	0.0849	0.0849
	2点法 B	-0.0041	0.0851	0.0851
	最小2乗法 A	0	0.0788	0.0788
	最小2乗法 B	0	0.0783	0.0783
20	2点法 A	0	0.0618	0.0618
	2点法 B	-0.0036	0.0618	0.0618
	最小2乗法 A	0	0.0594	0.0594
	最小2乗法 B	0	0.0587	0.0587

の値である。

表 7.2 より μ の推定については、次のようなことがわかる。

- 1) 2点法 B はわずかに負の偏りをもつが、それ以外はすべて不偏である。最小 2 乗法の不偏性も容易に証明できることである。
- 2) 2点法については、2点法 A のほうが 2点法 B よりも平均 2 乗誤差は小さい。逆に最小 2 乗法では、最小 2 乗法 B のほうが最小 2 乗法 A よりも平均 2 乗誤差はわずかに小さい。
- 3) 2点法 A と最小 2 乗法 B を比べると、最小 2 乗法 B のほうが平均 2 乗誤差は小さいが、その差はさほど大きくない。

一方、表 7.3 より b の推定については、次のようなことがわかる。

- 1) 2点法 B, 最小 2 乗法 B はともに正の偏りもち、それと 2点法 A, 最小 2 乗法 A に比べ平均 2 乗誤差はかなり大きい。
- 2) 2点法 A と最小 2 乗法を比べると、 $n \geq 10$ では 2点法 A のほうが平均 2 乗誤差は小さい。

b の推定量の性質で 1) の理由としては、 β_1, \dots, β_n の存在する範囲が $\alpha_1, \dots, \alpha_n$ の範囲より狭いことがあげられる。 $n=10$ での両者の値を比べたのが表 7.4 である。また 2) の理由としては、極値分布が左に大きくすそを引いた分布であるため、最小 2 乗法ではすべての点に同じ重みを与えることにより、とくに最小値のバラツキの影響が大きくなることが考えられる。

以上をまとめると、 μ の推定については 4 つの推定法に大差はないが、 b の推定では 2点法 A が最も良いということになる。

表 7.3 b に対する推定法の比較

n	推定法	偏り	分散	MSE
5	2点法 A	0	0.2068	0.2068
	2点法 B	0.2589	0.3277	0.3948
	最小2乗法 A	0	0.2063	0.2063
	最小2乗法 B	0.2515	0.3164	0.3797
10	2点法 A	0	0.0912	0.0912
	2点法 B	0.1190	0.1140	0.1282
	最小2乗法 A	0	0.0963	0.0963
	最小2乗法 B	0.1720	0.1280	0.1576
15	2点法 A	0	0.0546	0.0546
	2点法 B	0.0981	0.0658	0.0754
	最小2乗法 A	0	0.0600	0.0600
	最小2乗法 B	0.1359	0.0738	0.0923
20	2点法 A	0	0.0384	0.0384
	2点法 B	0.0801	0.0448	0.0512
	最小2乗法 A	0	0.0421	0.0412
	最小2乗法 B	0.1145	0.0494	0.0625

表 7.4 α_i と β_i の比較 ($n=10$)

i	1	2	3	4	5
α_i	-2.8798	-1.8262	-1.2672	-0.8681	-0.5436
β_i	-2.3506	-1.6061	-1.1443	-0.7941	-0.5007
i	6	7	8	9	10
α_i	-0.2575	0.0120	0.2837	0.5846	0.9899
β_i	-0.2377	0.0115	0.2618	0.5334	0.8746

7.4 ワイブル分布の形状母数の推定

一般に、ある推定量 \hat{b} が b にとって良い推定量であっても、 $1/\hat{b}$ が $m = 1/b$ の良い推定量になっているとは限らない。この点がワイブル分布の形状母数の推定も非常に難しいものになっている。そこでこの節では、このワイブル分布の形状母数の推定について調べることにする。 m の推定についても最小2乗法と2点推定法が考えられるが、最小2乗法を数値積分で評価しようとするとは n 重積分をすることになる。したがって最小2乗法については、シミュレーションによる評価に任せることにし、ここでは直接には取り扱わない。 m に対する最小2乗法というのは、 b の最小2乗推定量の逆数もとったものと、 $X_{(1)} \leq \dots \leq X_{(n)}$ を与えられたものとして縦軸方向の距離について残差平方和を最小にするものの2つが考えられるが、後者については日本品質管理学会信頼性研究部会 [46] のシミュレーション結果が報告されている。ここでは2重積分の範囲で調べられるものとして、2点推定法の評価を行なう。

形状母数 m に対する2点推定法とは、確率紙上にプロットされた点の中の2点を結び、その傾きから m を推定するものだが、この2点推定法は Dubey [16] の "percentile estimator" に類似している。Dubey は順序統計量の漸近分布をもとに $p_1 = 0.1673$, $p_2 = 0.9737$ として

$$\hat{m} = \frac{\ln \ln [1/(1-p_2)] - \ln \ln [1/(1-p_1)]}{X_{(np_2)} - X_{(np_1)}} \quad (7.10)$$

で推定すると、漸近効率 67% が得られるといっている。ここでは、小標本における順序統計量の最適組合せと縦軸の値のとり方について考察したい。

m に対する 2 点推定法についても、縦軸に $\alpha_1, \dots, \alpha_n$ をとったときのものを 2 点法 A と呼ぶことにすると、このときの推定量は

$$\hat{m}_A = \frac{\alpha_j - \alpha_i}{X_{(j)} - X_{(i)}} \quad (7.11)$$

となり、一方縦軸に β_1, \dots, β_n をとったときの 2 点法 B による推定量は

$$\hat{m}_B = \frac{\beta_j - \beta_i}{X_{(j)} - X_{(i)}} \quad (7.12)$$

となる。またワイブル確率紙での推定とは直接関係しないが、式(7.11)、(7.12)に類する形で m の不偏推定量

$$\hat{m}^* = \frac{K_{ij}}{X_{(j)} - X_{(i)}} \quad (7.13)$$

を考えることができる。ここで

$$K_{ij} = \frac{1}{E[1/(U_{(j)} - U_{(i)})]}$$

である。式(7.13)を不偏 2 点法と呼ぶことにする。信頼性の分野では、偏りなく推定することが望まれる場合が多い。とくに m の推定ではその目的からして単に MSE を基準とし縮小された推定量を用いることは必ずしも得策でない。不偏 2 点法の最適組合せと係数 K_{ij} を $n=5 \sim 20$ について示したのが表 7.5 である。最適組合せは表 7.1 での b に対する最適組合せとほぼ同じものになっている。 K_{ij} については、式(7.10)の分子の値が 2.989

表 7.5 m に対する不偏2点法の
最適組合せと係数

n	i, j	K_{ij}
5	1, 5	2.3038
6	1, 6	2.6303
7	1, 7	2.8925
8	1, 8	3.1116
9	2, 9	2.3942
10	2, 10	2.5666
11	2, 11	2.7213
12	2, 12	2.8537
13	2, 13	2.9762
14	2, 14	3.0819
15	3, 15	2.6488
16	3, 16	2.8056
17	3, 17	2.8988
18	3, 18	2.9851
19	4, 19	2.6134
20	4, 20	2.7204

表 7.6 m に対する2点推定法の比較

n	推定法	偏り	分散	MSE
5	2点法 A	0.2486	0.5304	0.5992
	2点法 B	-0.0081	0.3357	0.3358
	不偏2点法	0	0.3411	0.3411
10	2点法 A	0.0971	0.1338	0.1432
	2点法 B	-0.0335	0.1038	0.1039
	不偏2点法	0	0.1111	0.1111
15	2点法 A	0.0715	0.0715	0.0770
	2点法 B	-0.0215	0.0593	0.0597
	不偏2点法	0	0.0619	0.0619
20	2点法 A	0.0593	0.0553	0.0588
	2点法 B	-0.0193	0.0474	0.0478
	不偏2点法	0	0.0493	0.0493

ぐらいであるが、 K_{ij} の値もほぼこれに対応したものになっている。また、 \hat{m}_A , \hat{m}_B , \hat{m}^* の偏り、分散、平均2乗誤差を計算した結果が表7.6である。これらの計算方法については付録3を参照したい。なお、これらはすべて $m = 1$ の場合の値であり、一般の場合には偏りには m を、分散と平均2乗誤差には m^2 を乗じればよい。

表7.6より m に対する2点推定法について、以下のことがわかる。

- 1) 2点法Aは正の偏りを持ち、2点法Bはわずかであるが負の偏りをもつ。
- 2) 平均2乗誤差で比べる限りでは、2点法Bが最も優れているが、不偏2点法もほとんど変わらない。2点法Aは明らかに劣る。

すなわち7.3節の極値分布の母数推定とは逆に2点法Bが優れた性質をもっているということである。なお、不偏2点法の $n = 10 \sim 20$ でのクラーメル・ラオの下限に対する効率は55~65%程度の値である。

また、柴田の報告^[45]と信頼性研究部会の報告^[46]を引用し、他の推定量とここで与えた不偏2点法の平均2乗誤差を比較したのが図7.1である。図7.1より、不偏2点法は柴田の提案した1次の偏り修正をした最尤推定量には劣るが、モーメント推定量、通常的最尤推定量やワイブル確率紙上での最小2乗推定量よりも優れたものであることがわかる。このような結果の理由としては、モーメント推定量や最尤推定量が小標本ではかなり大きな正の偏りをもったため(第3章の競合ワイブル分布でも、図3.2, 図3.3で同様の傾向がみられる)、2点推定量でも不偏にすることにより平均2乗誤差における有効性が出てくるものと思われる。

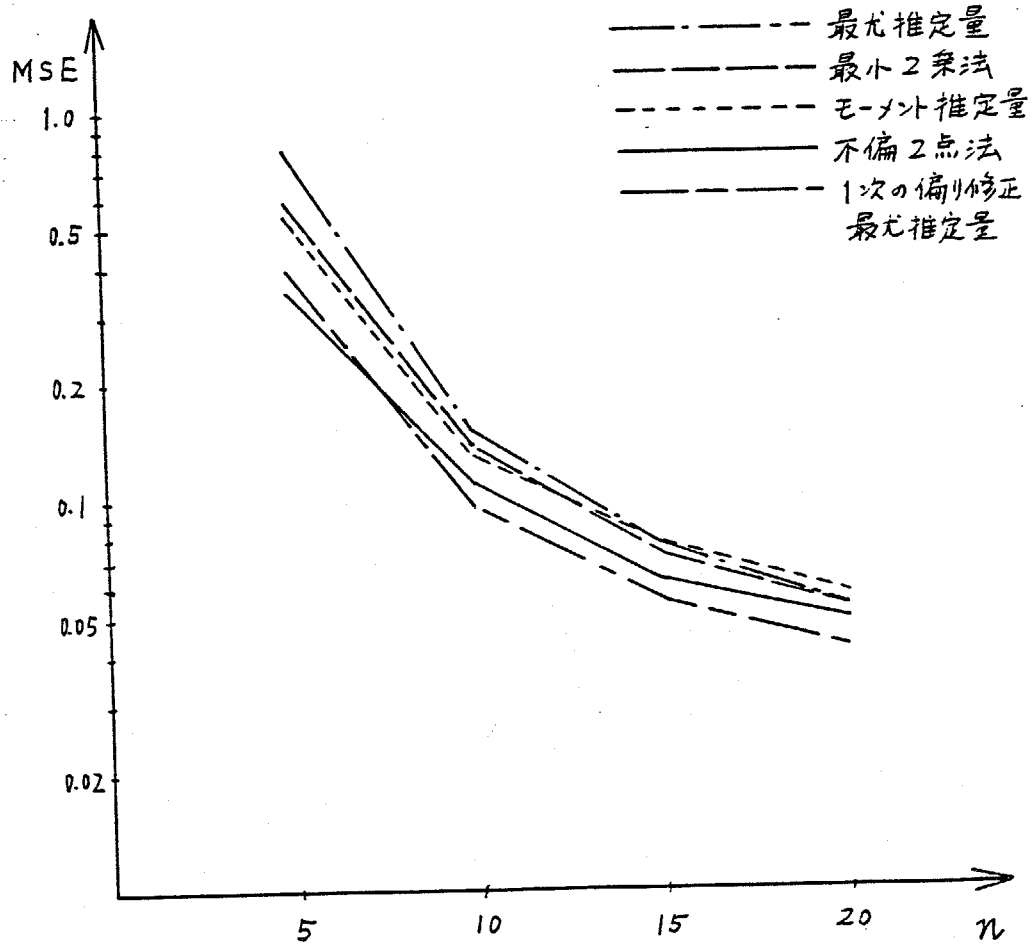


図7.1 m の推定量の比較

7.5 結 言

本章では、ワイブル分布の母数推定法としてワイブル確率紙による簡易推定法を取り上げ、打点の方法、直線の引き方による推定精度の比較を行い、以下の結果を得た[33]。

- 1) ワイブル確率紙による推定の精度を客観的に調べるために、ワイブル確率紙の数学的構造を明らかにし、線形推定理論の適用を可能にした。
- 2) 1) のもとで、極値分布の位置母数と尺度母数の推定については、打点の仕方と直線の引き方から4種類の推定法を考え、この中では最適な2点を結んだ2点法Aが優れていることがわかった。2点法Aの最良線形不偏推定量に対する推定効率、位置母数、尺度母数でそれぞれ90%、80~85%程度と高い値を示すことが明らかになった。
- 3) ワイブル分布の形状母数の推定については、3種類の2点推定法を考え、この中では2)とは逆に2点法Bが優れていることがわかった。また他の推定量との比較からも、小標本では最適な2点を結んだ2点法B、不偏2点法が優れたものであり、クラメル・ラオの下限に対する効率は55~65%であることがわかった。
- 4) ここで提示した2点推定法は、確率紙を用いなくとも計算できるものであり、計算機による繰り返し計算の初期値として用いるなど、予備的な推定法としても有効であることが示された。

第8章 結 論

本論文では、信頼性工学における統計学的手法について理論的考察を行い、システムの実使用段階における多くの信頼性データから、企画・設計段階で活用しうる有用な情報を抽出し、さらに試作段階の寿命試験へと展開するための方法論について論じた。本章では結論として、本論文で提唱した一連の統計的手法の体系的な運用方法について総括するとともに、残された問題、今後の展望について考察する。

フィールドにおけるシステムの信頼性解析においては、まず故障のメカニズム、システムの構造を明確にし、対象とする故障モードを定める必要がある。この際、第6章に述べた故障原因データからの構造関数の推定法は、フィールドでのシステムの信頼性データに基づく1つの統計的手法になり得る。また、何を知りたいのかを明確にし、そのためにはどのようなデータを採取するかを決めることもこの段階の重要な課題であるが、第6章に述べた信頼性データ型と解析手法の関係はこの決定に指針を与えるものと言える。

故障メカニズム、故障モードが定まれば、競合モデルを基にした解析により各コンポーネント、各故障モードの寿命特性もシステムの寿命データから推定することができる。この際、データの質、量、さらに予備的情報の程度に応じて、パラメトリック推定とノンパラメトリック推定を使い分けていくことが望ましい。第2章から第4章までに提唱した不完全な競合データに基づくパラメトリックおよびノンパラメトリック推定法は競合モデルに基づく統計的方法の適用範囲を拡張させたものと言える。このような競合モデルに基づく解析によ

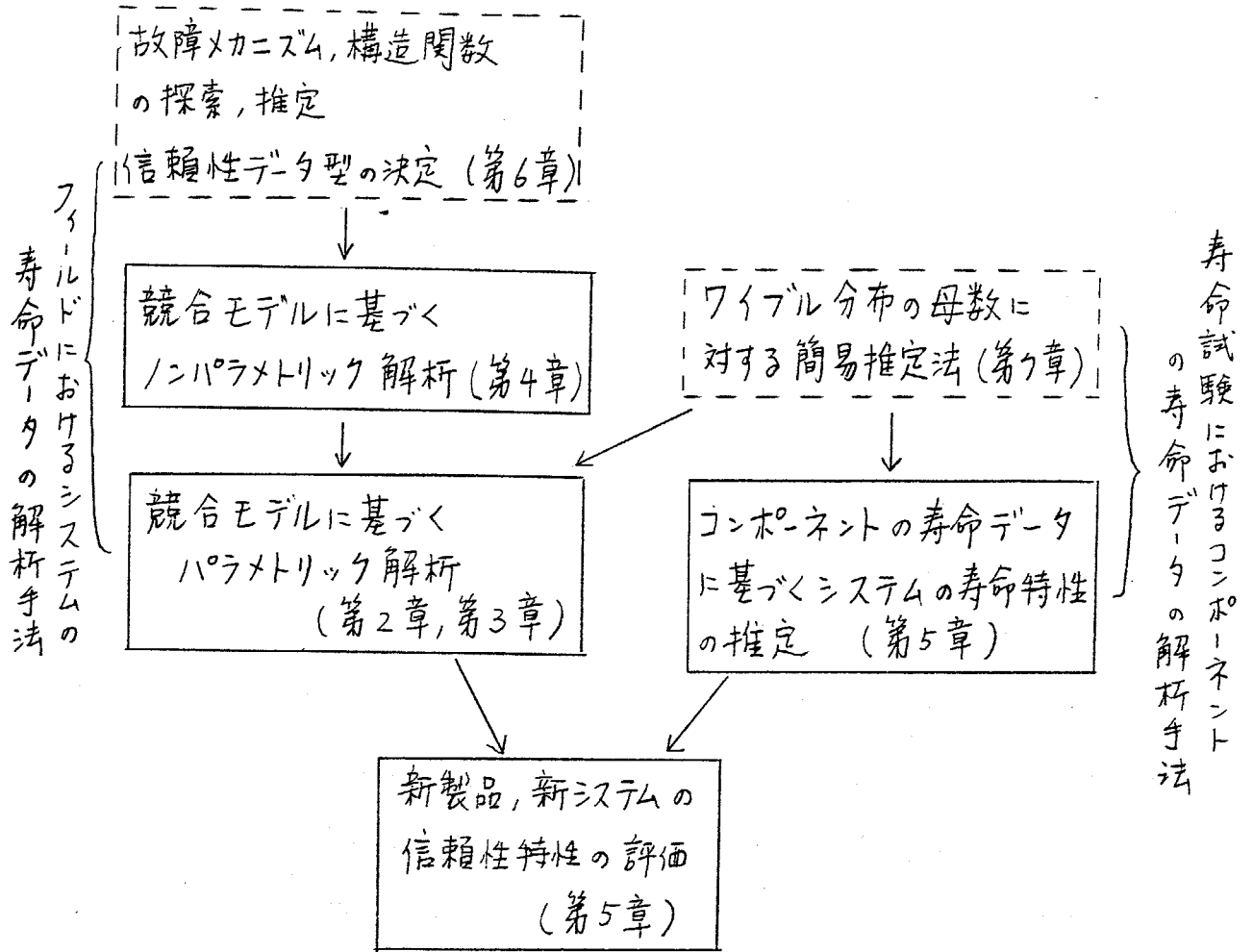


図 8.1 本論文で提唱した統計的方法の体系

り、既存のシステムの信頼性を向上させる上で、とくに重視すべき故障モード、コンポーネントを絞り込むことができる。

一方、新製品、新システムの開発において、フィールドにおける既存のシステムの寿命データからは得られない寿命情報については、コンポーネントレベルでの寿命試験を実施し、その寿命データを採取する必要があるが、このとき第5章で与えた方法により、各コンポーネントの寿命データからシステム全体の寿命特性を効率的に推定することが可能になる。この方法は、各コンポーネントの寿命特性の推定が基本となるが、第7章で与えたワイブル分布の母数に

対する簡易推定法は、コンポーネントの寿命分布にワイブル分布を仮定したとき、有効な手段となり得る。また同時に、ワイブル分布を仮定した競合モデル（第3章）においても役立つものである。

さらに、上記2つの方法群を受けて、フィールドにおける既存のシステムから得た寿命情報とコンポーネントの寿命試験から得た寿命情報を併合し、新システムあるいは改善されたシステムの信頼性特性を評価することになるが、これについては第5章で、コンポーネントの寿命データ解析の枠組みの中で行う1つの方法を与えた。

以上が本論文で提唱した統計的方法の運用体系であり、図8.1にその概念図を示した。

最後に残された問題、今後の展望について述べる。本論文で与えた統計的方法は、システムの信頼性に関する結果系のデータに対する詳細な検討を意図したものであり、要因系データとの対応づけについては論じていない。フィールドにおいても寿命試験においても、故障時間、故障モードといった結果系データとともに多くの要因系データを的確に採取し、これらの要因解析をすることは極めて重要であるが、このときいわゆる分散分析型の手法は平均値も特性とするものであるため、信頼性の分野では必ずしも適切でないことが多い。とくに修理を伴うシステムでは、システムの運用上重要な指標は平均値よりもむしろ各時点での故障率であり、各モードごとの故障率関数に対する様々な要因の影響を定量的に調べるこここそ肝要である。Coxの比例ハザードモデルはこのような観点からは評価されるが、比例性という強い仮定や数値計算上の複雑さなどの問題点がある。また修理を伴うシステムでは、履歴を要因として考慮したモデルとその解析法が必要不可欠である。このように信頼性における統計的方法は、その問題の特異性から、要因解析手法については他の分野に比

べ遅れている。この種の手法に関する理論的かつ実践的な研究が今後益々望まれるところである。

一方、結果系データの解析について言えば、従属な競合モデル（これは打ち切り時間が故障時間と従属なランダムセンサリングモデルと同等である）に関する研究が不可欠である。第2章に述べたように、通常の競合データから独立性従属性も同定することはできないので、これには別の情報、仮定が必要である。この際、パラメトリックな多変量モデルよりも、故障率関数に対するノンパラメトリックな従属モデルを基に、再配分概念を用いた解析手法の方が、とくにフィールドデータ解析では有効と考えられる。近年、Sludら^[47]によってそのような手法の方向付けがなされたことは意義深い。競合モデルは故障率関数に対する自然な加法モデルであり、これに基づく解析結果はシステムの設計、改善に有用な情報を与え得るものであるから、原因系データを併合した競合モデルに基づく統計的方法の今後のいっそうの発展が期待される。

謝 辞

昭和50年4月に東京工業大学に入学し、翌51年経営工学科に所属してから57年3月博士課程を中退するまでの6年間、さらにその後も経営工学科に勤務する中で、本学科の先生方には多くの御指導を賜わって頂きました。ここに心より御礼申し上げる次第です。

本論文は筆者の大学院時代の研究が基になっていますが、当時指導教官であられた真壁肇教授にはその後も一貫して懇切なる御指導をいただいてまいりました。また生産管理講座助手に着任して以来、秋庭雅夫教授、圓川隆夫助教授には現業実習等を通じて、問題への取組み方、実データの解析の仕方についてご教示いただくとともに、論文作成にあたっては多大な励ましをいただきました。三人の先生方にはお礼の言葉も見当たりません。

さらに、大学院時代からお世話になっている鈴木久敏助手、西尾敦助手をはじめ、共同研究者になっていただいた無機材料工学科松尾陽太郎助教授、山本久志氏（現(株)東芝勤務）、原田哲也氏（現日本電気(株)勤務）の方々に深く感謝の意を表します。

最後になりましたが、本論文作成にあたり惜しみないご協力をいただいた10年来の僚友伊藤謙治助手に厚く御礼申し上げます。

昭和60年 6月

宮川雅巳

参 考 文 献

- [1] Bain, L.J. (1978): *Statistical Analysis of Reliability and Life Testing Model*, New York : Marcel Dekker.
- [2] Barlow, R.E., Bartholomew, J.M., Bremner, J.M. and Brunk, H.D. (1972): *Statistical Inference under Order Restrictions*, John Wiley and Sons.
- [3] Barlow, R.E. and Proschan, F. (1975): *Statistical Theory of Reliability and Life Testing*, Holt Rinehart and Winston.
- [4] Basu, A.P. (1964): "Estimates of Reliability for Some Distributions Useful in Life Testing", *Technometrics*, 6, 215-219.
- [5] Basu, A.P. and Mawaziny, A.H.El. (1978): "Estimates of Reliability of k-out-of-m Structures in the Independent Exponential Case", *Jour. Amer. Statist. Assoc.*, 73, 850-854.
- [6] Bolotin, V.V. (1969): *Statistical Methods in Structural Mechanics*, Holden-Day.
- [7] Boyles, R.A. (1983): "On the Convergence of the EM Algorithm", *Jour. Roy. Statist. Soc. B*, 47-50.
- [8] Breslow, N. and Crowley, J, (1974): "A Large Sample Study of the Life Table and Product Limit Estimator under Random Censorship", *Ann. Statist.*, 2, 437-453.
- [9] Chan, L.K. and Kabir, A.B.M.L. (1969): "Optimal Quantiles for the Linear Estimation of the Parameters of the Extreme-Value Distribution in Complete and Censored Samples", *Nav. Res. Log. Quart.*, 16, 381-404.
- [10] Chao, A. (1981): "Approximate Mean Square Errors of Estimators of Reliability for k-out-of-m Systems in the Independent Exponential Case", *Jour. Amer. Statist. Assoc.*, 76, 720-724.
- [11] Chen, Y.Y., Hollander, M. and Langberg, N.A. (1982): "Small Sample Results for the Kaplan-Meier Estimator", *Jour. Amer. Statist. Assoc.*, 77, 141-144.
- [12] Chernoff, H. and Lieberman, J. (1965): "The Use of Generalized Probability Paper for Continuous Distributions", *Ann. Math. Statist.*, 27, 806-818.
- [13] David, H.A. and Moeschberger, M.L. (1978): *The Theory of Competing Risks*, Griffin, London.
- [14] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977): "Maximum Likelihood from Incomplete Data via the EM Algorithm", *Jour. Roy. Statist. Soc. B*, 39, 1-22.

- [15] Dinse, G.E. (1982): "Nonparametric Estimation for Partially-Complete Time and Types of Failure Data", *Biometrics*, 38, 417-431.
- [16] Dubey, S.D. (1967): "Some Percentile Estimators of Weibull Parameters", *Technometrics*, 9, 119-129.
- [17] Efron, B. (1967): "The Two Sample Problem with Censored Data", *In Proc. Fifth Berkeley Symp. on Math. Statist. Prob.*, 831-853.
- [18] 市田 嵩 (1981): "対数正規分布の諸特性の確率紙による推定", 「品質」, 11, 233-242.
- [19] Inagaki, T. and Henley, E.J. (1980): "Probabilistic Evaluation of Prime Implicants and Top-Events for Non-Coherent Systems", *IEEE Trans. Reliab.*, R-24(5), 361-367.
- [20] 蟹江 壽, 野中保雄 (1982): "競合分布のパラメータ簡易推定法", 「第12回信頼性・安全性シンポジウム発表報文集」, 日科技連, 267-272.
- [21] Kao, J.H.K. (1959): "A Graphic Estimation of Mixed Weibull Parameters in Life-Testing of Electron Tubes", *Technometrics*, 1, 389-407.
- [22] Kaplan, E.L. and Meier, P. (1958): "Nonparametric Estimation from Incomplete Observations", *Jour. Amer. Statist. Assoc.*, 53, 457-481.
- [23] Kendall, M. and Stuart, A. (1973): *The Advanced Theory of Statistics*, Vol.1, 3rd Edition, C. Griffin, London.
- [24] Kimball, B.F. (1960): "On the Choice of Plotting Position on Probability Paper", *Jour. Amer. Statist. Assoc.*, 55, 546-560.
- [25] Lehmann, E.L. (1959): *Testing Statistical Hypotheses*, John Wiley, New York.
- [26] Lieblein, J. and Zelen, M. (1956): "Statistical Investigation of the Fatigue Life of Deep Groove Ball Bearings", *Jour. Res. Nat. Bur. Stand.*, 57, 273-316.
- [27] Mann, N.R. (1971): "Best Linear Invariant Estimation for Weibull Parameters under Progressive Censoring", *Technometrics*, 13, 521-533.
- [28] 松尾陽太郎, 佐藤秀幸 (1979): "せい性材料からなる回転円板の破壊とその確率論的取扱い", 「日本機械学会論文集」, A-45, 171-178.
- [29] 松尾陽太郎, 木村修七, 安田栄一, 犬飼崇男 (1984): "せい性材料の強度に及ぼす負荷速度の影響と二重モードワイブル分布", 「窯業協会誌」, 92-5, 274-280.
- [30] 松尾陽太郎, 村田博隆 (1984): "セラミックスの曲げ試験における強度分布モデルの選択とモンテカルロシミュレーション", 「材料」, 33, 1545-1551.

- [31] 宮川雅己 (1982): "競合モデルにおける不完全データの統計的解析", 「品質」, 12, 23-29.
- [32] 宮川雅己 (1982): "コンポ-ネントの寿命データに基づくコヒーレントシステムの信頼度の推定", 「応用統計学」, 11, 81-87.
- [33] 宮川雅己 (1983): "ワイブル確率紙による母数推定の小標本での偏りと平均2乗誤差", 「品質」, 13, 47-53.
- [34] 宮川雅己, 山本久志 (1984): "競合モデルにおける信頼度関数のノンパラメトリック推定 ~ 故障原因不明データを含む場合", 「品質」, 14, 59-67.
- [35] Miyakawa, M. (1984): "Analysis of Incomplete Data in Competing Risks Model", *IEEE Trans. Reliab.*, R-33(4), 293-296.
- [36] 宮川雅己, 原田哲也 (1984): "コンポ-ネントの寿命データに基づくコヒーレントシステムのMTTFの推定", 「応用統計学」, 13, 61-70.
- [37] Miyakawa, M. (1984): "On Stochastic Coherent Systems", *Stochastic Models in Reliability Theory, Springer Lecture Notes in Economics and Mathematical Systems*, 235, 1-11.
- [38] 宮川雅己 (1985): "フィールドライフデータの解析手法", 「オペレーションズ・リサーチ」, 30, 184-190.
- [39] 宮川雅己, 鈴木和幸 (1985): "不完全データによる寿命解析法", 「品質」, 15, 123-131.
- [40] 村田博隆, 松尾陽太郎, 宮川雅己, 北上浩一 (1985): "多重モードワイブル分布のパラメータ推定におよぼす破壊原因データの役割", 「日本機会学会論文集」, 掲載予定.
- [41] Nadler, J. (1960): "Inverse Binomial Sampling Plans When an Exponential Distribution is Sampled with Censoring", *Ann. Math. Statist.*, 31, 1201-1204.
- [42] Peterson, A.V.Jr. (1976): "Bounds for a Joint Distribution Function with Fixed Sub-Distribution Functions: Application to Competing Risks", *Proc. Nat. Acad. Sci.*, 72, 20-22.
- [43] Peterson, A.V.Jr. (1977): "Expressing the Kaplan-Meier Estimator as a Function of Empirical Subsurvival Functions", *Jour. Amer. Statist. Assoc.*, 72, 854-858.
- [44] Rutemiller, H.C. (1966): "Point Estimation of Reliability of a System Comprised of k Elements from the Same Exponential Distribution", *Jour. Amer. Statist. Assoc.*, 61, 1033-1051.
- [45] 柴田義貞 (1980): "ワイブル分布の母数推定量の小標本での比較", 「日本品質管理学会第17回研究発表会研究発表要旨集」, 55-58.

- [46] 信頼性研究部会報告(1981): "2母数ワイブル分布の点推定", 「日本品質管理学会第11回年次大会研究発表要旨集」, 1-4.
- [47] Slud, F.V. and Rubinstein, L.V. (1983): "Dependent Competing Risks and Summary Curves", *Biometrika*, 70, 643-649.
- [48] 鉄道車両の保守管理委員会(1984): 「昭和58年度鉄道車両の保守管理に関する研究報告書」, 日本鉄道技術協会.
- [49] Tsiatis, A.A. (1975): "A Nonidentifiability Aspect of the Problem of Competing Risks", *Proc. Nat. Acad. Sci.*, 72, 20-22.
- [50] Turnbull, B.W. (1976): "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data", *Jour. Roy. Statist. Soc. B*, 38, 290-295.
- [51] 吉本英明(1984): "鉄道信号システムの信頼性,安全性,保全性 ~ 新幹線ATCにおける事例", 「第12回OR学会シンポジウム要旨集」, 8-16.
- [52] Zacks, S. (1971): *The Theory of Statistical Inference*, John Wiley, New York.
- [53] Zacks, S. and Even, M. (1966a): "The Efficiency in Small Samples of the Maximum Likelihood and Best Unbiased Estimators of Reliability Functions", *Jour. Amer. Statist. Assoc.*, 61, 1033-1051.
- [54] Zacks, S. and Even, M. (1966b): "Minimum Variance Unbiased and Maximum Likelihood Estimators of Reliability Function for Systems in Series and in Parallel", *Jour. Amer. Statist. Assoc.*, 61, 1052-1062.
- [55] Zelen, M. Dannemiller, M. (1961): "The Robustness of Life Testing Procedures Derived from the Exponential Distribution", *Technometrics*, 3, 29-49.

付 録

付録1. 式(2.13), (2.14)の導出

式(2.5)より

$$\frac{\partial^2}{\partial \lambda_1^2} \ln L = -\frac{r}{\lambda_1}$$

$$\frac{\partial^2}{\partial \lambda_1 \partial \lambda_2} \ln L = 0$$

$$\frac{\partial^2}{\partial \lambda_2^2} \ln L = -\frac{n-r}{\lambda_2}$$

となる。ここで $E(r) = n\lambda_1 / (\lambda_1 + \lambda_2)$ であるから Fisher の情報量行列は

$$I_1(\lambda_1, \lambda_2) = \begin{pmatrix} \frac{n}{\lambda_1(\lambda_1 + \lambda_2)} & 0 \\ 0 & \frac{n}{\lambda_2(\lambda_1 + \lambda_2)} \end{pmatrix}$$

となる。また、この逆行列は

$$I_1^{-1}(\lambda_1, \lambda_2) = \begin{pmatrix} \frac{\lambda_1(\lambda_1 + \lambda_2)}{n} & 0 \\ 0 & \frac{\lambda_2(\lambda_1 + \lambda_2)}{n} \end{pmatrix}$$

となり、 $\hat{\lambda}_1$ の漸近分散はこの (1, 1) 要素

$$A. \text{Var}(\hat{\lambda}_1) = \frac{\lambda_1^2 + \lambda_1 \lambda_2}{n} \quad (2.13)$$

で与えられる。一方、式 (2.10), (2.11) より

$$\frac{\partial^2}{\partial \lambda_1^2} \ln L = -\frac{r}{\lambda_1^2} + \frac{n-r}{(\lambda_1 + \lambda_2)^2}$$

$$\frac{\partial^2}{\partial \lambda_1 \partial \lambda_2} \ln L = \frac{n-r}{(\lambda_1 + \lambda_2)^2}$$

$$\frac{\partial^2}{\partial \lambda_2^2} \ln L = -\frac{n-r}{\lambda_2^2} + \frac{n-r}{(\lambda_1 + \lambda_2)^2}$$

となる。ここで、 $E(r) = n\lambda_1 / (\lambda_1 + \lambda_2)$ であるから Fisher の情報量行列は

$$I_2(\lambda_1, \lambda_2) = \begin{pmatrix} \frac{n}{\lambda_1(\lambda_1 + \lambda_2)} - \frac{n\lambda_2}{(\lambda_1 + \lambda_2)^3} & -\frac{n\lambda_2}{(\lambda_1 + \lambda_2)^3} \\ -\frac{n\lambda_2}{(\lambda_1 + \lambda_2)^3} & \frac{n}{\lambda_2(\lambda_1 + \lambda_2)} - \frac{n\lambda_1}{(\lambda_1 + \lambda_2)^3} \end{pmatrix}$$

となる。また、この逆行列は

$$I_2^{-1}(\lambda_1, \lambda_2) = \begin{pmatrix} \frac{\lambda_1^2 + 2\lambda_1\lambda_2}{n} & \frac{\lambda_2^2}{n} \\ \frac{\lambda_2^2}{n} & \frac{\lambda_2(\lambda_1^2 + \lambda_1\lambda_2 + \lambda_2^2)}{n\lambda_1} \end{pmatrix}$$

となり、 $\tilde{\lambda}_1$ の漸近分散はこの (1, 1) 要素

$$A. \text{Var}(\tilde{\lambda}_1) = \frac{\lambda_1^2 + 2\lambda_1\lambda_2}{n} \quad (2.14)$$

で与えられる。

付録2. 式(5.20), (5.21), (5.24) の導出

$\tilde{\theta}_i = Z_i / n_i$ の確率密度関数は

$$f(\tilde{\theta}_i) = \frac{\tilde{\theta}_i^{n_i-1} e^{-(n_i/\theta_i)\tilde{\theta}_i}}{(\theta_i/n_i)^{n_i} \Gamma(n_i)}$$

であるから、 $\tilde{\theta}_i - \theta_i$ のモーメントは

$$E(\tilde{\theta}_i - \theta_i) = 0, \quad E(\tilde{\theta}_i - \theta_i)^2 = \theta_i^2 / n_i,$$

$$E(\tilde{\theta}_i - \theta_i)^3 = 2\theta_i^3 / n_i^2, \quad E(\tilde{\theta}_i - \theta_i)^4 = 3\theta_i^4 / n_i^2 + 6\theta_i^4 / n_i^3,$$

$$E(\tilde{\theta}_i - \theta_i)^5 = O(1/n_i^3), \dots \quad (A.1)$$

となる (たとえば [23], p. 64 を参照). 一方, 2コンポーネント直列システムにおけるMLEの偏りは

$$\begin{aligned} E(\tilde{\mu}_s - \mu_s) &= E\left(\frac{\tilde{\theta}_1 \tilde{\theta}_2}{\tilde{\theta}_1 + \tilde{\theta}_2} - \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right) \\ &= \frac{\theta_1 \theta_2}{\theta_1 + \theta_2} E\left[\frac{\left(1 + \frac{\tilde{\theta}_1 - \theta_1}{\theta_1}\right)\left(1 + \frac{\tilde{\theta}_2 - \theta_2}{\theta_2}\right)}{\left(1 + \frac{\tilde{\theta}_1 - \theta_1}{\theta_1 + \theta_2} + \frac{\tilde{\theta}_2 - \theta_2}{\theta_1 + \theta_2}\right)} - 1\right] \end{aligned}$$

と書けるが, ここで $\varepsilon_1 = \tilde{\theta}_1 - \theta_1$, $\varepsilon_2 = \tilde{\theta}_2 - \theta_2$ として右辺を級数展開すると

$$E(\tilde{\mu}_s - \mu_s) = \frac{\theta_1 \theta_2}{\theta_1 + \theta_2} E\left[\left(1 + \frac{\varepsilon_1}{\theta_1}\right)\left(1 + \frac{\varepsilon_2}{\theta_2}\right) \sum_{r=0}^{\infty} (-1)^r \left(\frac{\varepsilon_1 + \varepsilon_2}{\theta_1 + \theta_2}\right)^r - 1\right]$$

となる. さらに ε_1 と ε_2 が独立であることから, 式(A.1)の結果をもとに n_i^2 のオーダーまで集めると

$$\begin{aligned} E(\tilde{\mu}_s - \mu_s) \doteq & -\frac{(n_1 + n_2) \theta_1^2 \theta_2^2}{n_1 n_2 (\theta_1 + \theta_2)^5} \left[n_2 (n_1 + 1) \theta_1^2 + 2(n_1 n_2 - n_1 - n_2) \theta_1 \theta_2 \right. \\ & \left. + n_1 (n_2 + 1) \theta_2^2 \right] \end{aligned}$$

を得る. ここで $n_2 = a n_1$ とおけば式(5.20)を得る.

同様にして MSE についても

$$E(\tilde{\mu}_S - \mu_S)^2 = \mu_S^2 E \left[\left(1 + \frac{\varepsilon_1}{\theta_1}\right) \left(1 + \frac{\varepsilon_2}{\theta_2}\right) \left(1 + \frac{\varepsilon_1 + \varepsilon_2}{\theta_1 + \theta_2}\right)^{-1} - 1 \right]^2$$

より式(5.21)を得る。

次に, $m=2$ における \bar{U} -統計量 μ_S^* の分散は以下のように求められる。いま, $Y_{ij} = \min(T_{1i}, T_{2j})$ とすれば, μ_S^* の2次のモーメントは

$$\begin{aligned} E(\mu_S^*)^2 &= \frac{1}{n_1^2 n_2^2} E \left[\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} Y_{ij} \right]^2 \\ &= \frac{1}{n_1 n_2} E(Y_{ij}^2) + \frac{n_2-1}{n_1 n_2} E(Y_{ij} Y_{il}) \\ &\quad + \frac{n_1-1}{n_1 n_2} E(Y_{ij} Y_{kj}) + \frac{(n_1-1)(n_2-1)}{n_1 n_2} E(Y_{ij} Y_{kl}) \\ &\hspace{15em} (i \neq k, j \neq l) \end{aligned} \tag{A.2}$$

と書けるが, ここで

$$E(Y_{ij}^2) = \frac{2\theta_1^2 \theta_2^2}{(\theta_1 + \theta_2)^2}, \quad E(Y_{ij} Y_{kl}) = \frac{\theta_1^2 \theta_2^2}{(\theta_1 + \theta_2)^2}$$

であり, さらに

$$\begin{aligned} E(Y_{ij} Y_{il}) &= E[\min(T_{1i}, T_{2j}) \times \min(T_{1i}, T_{2l})] \\ &= 2 E[T_{1i}^2 \mid T_{1i} < T_{2j} < T_{2l}] \times \Pr[T_{1i} < T_{2j} < T_{2l}] \end{aligned}$$

$$\begin{aligned}
& + 2E[T_{1i}T_{2j} | T_{2j} < T_{1i} < T_{2l}] \times P_r[T_{2j} < T_{1i} < T_{2l}] \\
& + 2E[T_{2j}T_{2l} | T_{2j} < T_{2l} < T_{1i}] \times P_r[T_{2j} < T_{2l} < T_{1i}] \\
& = \frac{4\theta_1^2\theta_2^3}{(2\theta_1+\theta_2)^3} + \frac{2\theta_1^3\theta_2^2(4\theta_1+3\theta_2)}{(\theta_1+\theta_2)^2(2\theta_1+\theta_2)^3} + \frac{2\theta_1^4\theta_2^2(4\theta_1+3\theta_2)}{(\theta_1+\theta_2)^2(2\theta_1+\theta_2)^3} \\
& = \frac{2\theta_1^2\theta_2^2}{(\theta_1+\theta_2)(2\theta_1+\theta_2)}
\end{aligned}$$

とほり, 同様にして

$$E(Y_{ij} Y_{kj}) = \frac{2\theta_1^2\theta_2^2}{(\theta_1+\theta_2)(\theta_1+2\theta_2)}$$

であるから, これらを式(A.2)へ代入することにより, 2次のモーメントが求まり, これより直ちに式(5.24)が導出される.

付録3. α_i, v_{ij}, K_{ij} の数値計算

X が式(7.2)に従うとき, $U = (X - u) / b$ の確率密度関数は

$$f(u) = e^u \exp[-e^u]$$

となるから, 順序統計量 $U_{(1)} \leq U_{(2)} \leq \dots \leq U_{(n)}$ に対し, $U_{(i)}$ の確率

密度関数は

$$f_{U_{(i)}}(u) = \frac{n!}{(i-1)!(n-i)!} \exp[u - (n-i+1)e^u] [1 - \exp(-e^u)]^{n-i}$$

となる。これより $U_{(i)}$ の 1 次, 2 次のモーメントは

$$E(U_{(i)}^k) = \int_{-\infty}^{\infty} u^k f_{U_{(i)}}(u) du \quad (k=1, 2) \quad (A.3)$$

で表わされ, α_i, ν_{ii} は

$$\alpha_i = E(U_{(i)})$$

$$\nu_{ii} = E(U_{(i)}^2) - \alpha_i^2$$

となるので, シンプソンの積分公式を用いて式(A.3)の数値積分を行なった。Lieblein and Zelen^[26]は $n=2 \sim 6$ について α_i と ν_{ij} を与えているが, 単積分で計算される α_i, ν_{ii} についてはここで計算した結果と8桁すべて一致している。

次に, $U_{(i)}$ と $U_{(j)}$ ($i < j$) の同時分布は

$$\begin{aligned} f_{U_{(i)}, U_{(j)}}(u, s) &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [1 - \exp(-e^u)]^{i-1} \\ &\quad \times [\exp(-e^u) - \exp(-e^s)]^{j-i-1} \\ &\quad \times \exp[u + s - (n-j+1)e^s - e^u] \end{aligned}$$

となる。ここで $U_{(i)}, \bar{Z} = U_{(j)} - U_{(i)}$ に変数変換すれば, 行列 V の非

対角要素 v_{ij} は

$$v_{ij} = \int_{-\infty}^{\infty} \int_0^{\infty} (s-z-\alpha_i)(s-\alpha_j) f_{Z, \bar{U}_{ij}}(z, s) dz ds \quad (A.4)$$

で与えられるので、シンプソンの公式を繰り返し用いた2重積分で式(A.4)の値を求めた。Lieblein and Zelen^[26]が $n=2 \sim 6$ について与えた v_{ij} の値と、ここで計算した結果は少なくとも5桁の一致をみせているので、 $n \geq 7$ の値についても実用上ほぼ問題のない程度の積分精度があると思われる。なお2点法Aの最適組合せは総当たりで求めたものである。

一方、 m に対する2点法A, 2点法B, 不偏2点法はいずれも

$$\tilde{m} = \frac{\gamma(i, j)}{X_{(j)} - X_{(i)}} = \frac{\gamma(i, j)}{b(\bar{U}_{(j)} - \bar{U}_{(i)})} = \frac{m \gamma(i, j)}{\bar{U}_{(j)} - \bar{U}_{(i)}}$$

という形で与えられるから、 $1/(\bar{U}_{(j)} - \bar{U}_{(i)})$ の期待値と分散を求めれば、2点推定法の評価ならびに不偏2点法の係数 K_{ij} の値が求まることになる。この期待値は

$$E\left(\frac{1}{\bar{U}_{(j)} - \bar{U}_{(i)}}\right) = \int_{-\infty}^{\infty} \int_0^{\infty} \frac{1}{z} f_{Z, \bar{U}_{ij}}(z, s) dz ds \quad (A.5)$$

で表わされるので、式(A.5)も式(A.4)と同様の方法で計算した。さらに分散については

$$\text{Var}\left(\frac{1}{\bar{U}_{(j)} - \bar{U}_{(i)}}\right)$$

$$= \int_{-\infty}^{\infty} \int_0^{\infty} \left[\frac{1}{z} - E \left(\frac{1}{\bar{U}_{(j)} - \bar{U}_{(i)}} \right) \right]^2 f_{Z, \bar{U}_{(j)}}(z, s) dz ds$$

を同様の方法で計算した。この場合にも2点推定法の最適組合せは、総当たりで求めた。