

博士論文

数理統計学と情報理論を用いた
限定的な事前情報下での
意思決定補助手法に関する研究

2025年3月

棚 橋 秀 斗

岡山大学大学院
自然科学研究科

目次

本論文で使用される主な記号	iii
第1章 序論	1
第2章 ポアソン分布を想定した複数変化点検出手法	7
2.1 緒言	7
2.2 ポアソン過程に対する複数変化点検出手法の設計	9
2.3 数値検証	13
2.4 提案手法のリアルタイムな工程解析への活用	15
2.5 結言	19
第3章 線形変化を想定した複数変化点検出手法	21
3.1 緒言	21
3.2 線形変化モデルの概要	22
3.3 複数線形変化点検出手法の設計	24
3.4 数値検証	27
3.4.1 単調変化点モデルの概略	27
3.4.2 複数ステップ変化点検出手法の概要	28
3.4.3 シミュレーション結果	29
3.5 結言	37
第4章 Hoeffding の確率不等式の性能向上	39
4.1 緒言	39
4.2 Hoeffding の確率不等式	40
4.3 確率不等式の性能改善に向けた考察	43
4.4 確率変数が2個の場合における性能改善	47
4.5 確率変数の個数が任意の場合における性能改善	49
4.6 数値検証	56
4.7 結言	64

第 5 章 結論	65
付 録 A 初期値を既知とした場合のポアソン分布を想定した 複数変化点検出手法	71
付 録 B Hoeffding の確率不等式の改良に用いる 定理 1 の証明	73
付 録 C KKT 条件を用いたエッジ解の導出	75
謝辞	79

本論文で使用される主な記号

e	: ネイピア数
$\log(x)$	: x の自然対数
$x_i, \forall i$	: 全ての i に対する x_i
$\max a(x)$	: 関数 $a(x)$ の最大値
$\arg \max_x a(x)$	: 関数 $a(x)$ の最大値を与える引数 x
$\min a(x)$	: 関数 $a(x)$ の最小値
$\arg \min_x a(x)$	: 関数 $a(x)$ の最小値を与える引数 x
$\lim_{x \rightarrow \infty} a(x)$	: 関数 $a(x)$ の $x \rightarrow \infty$ での極限值
$D(k, t)$	: 動的計画法の目的関数
$AIC(k)$	: 赤池情報量規準 (AIC)
$B(k)$	: 赤池情報量規準におけるペナルティ項
$\mathcal{L}(Z_1, \dots, Z_n, l)$	: ラグランジュ関数
l	: ラグランジュ乗数
$\Pr\{X \geq x\}$	: 確率変数 X の値が x 以上である確率 (上側確率)
$E_X[g(X)]$	: 確率変数 X に対する関数 $g(X)$ の期待値
$V_X[g(X)]$	: 確率変数 X に対する関数 $g(X)$ の分散
$f_X(X)$	: 確率変数 X の確率質量関数 (離散確率分布の場合) あるいは確率密度関数 (連続確率分布の場合)
$L(\boldsymbol{\theta} \mathbf{X})$	: パラメータ $\boldsymbol{\theta}$ に対しデータ $\mathbf{X}$ で評価する尤度関数
$\log L(\boldsymbol{\theta} \mathbf{X})$	: 自然対数を用いた対数尤度関数

変化点検出に関する数式記号

UCL	: 管理図の上側管理限界
LCL	: 管理図の下側管理限界
μ_0	: 管理状態における工程平均
σ_0^2	: 管理状態における工程の分散
t	: 期数

T	: データを取得した最後の期数 (最大期数)
$\mathbf{X}$	: サンプルデータ全体
x_t	: t 期目のサンプルから得た統計量 (各期ごとに1つのサンプルを得る場合)
$x_{t,i}$	: t 期目のサンプルのうち i 番目から得た統計量 (各期ごとに複数のサンプルを得る場合)
n	: 1 期あたりのサンプルサイズ
$\bar{X}_t$	: t 期目のサンプルの平均
$\sigma_{\bar{X}}^2$	: サンプルの平均 $\bar{X}_t$ の分散
K	: 複数変化点モデルにおける変化点数
$\hat{K}$	: 変化点数の推定値
τ_k	: k 番目の変化点
$\hat{\tau}_k$	: k 番目の変化点の最尤推定量
$\boldsymbol{\tau}^{(k)}$	: k 番目までの変化点の集合
$\hat{\boldsymbol{\tau}}^{(k)}$	: k 番目までの変化点の最尤推定量の集合
λ_k	: k 回目の変化点以後のポアソン分布の強度パラメータ
$\hat{\lambda}_k$	: k 回目の変化点以後の強度パラメータの最尤推定量
$\boldsymbol{\lambda}^{(k)}$	: k 回目までの強度パラメータの集合
$\hat{\boldsymbol{\lambda}}^{(k)}$	: k 回目までの強度パラメータの最尤推定量の集合
$\nu_k(t)$	: k 回目の変化以降である第 t 期における正規分布の平均
ξ_k	: k 回目の変化以降の線形変化の傾き
$\hat{\xi}_k$	: k 回目の変化以降の線形変化の傾きの最尤推定量
$\boldsymbol{\xi}^{(k)}$	: k 回目までの線形変化の傾きの集合
$\hat{\boldsymbol{\xi}}^{(k)}$	: k 回目までの線形変化の傾きの最尤推定量の集合
$\Lambda(\tau_{k-1}, \tau_k, \boldsymbol{\theta} \mid \mathbf{X})$	: 変化点 τ_{k-1}, τ_k 間の対数尤度関数

確率不等式に関する数式記号

Y_i	: i 番目の確率変数
n	: 確率変数の個数
μ_i	: i 番目の確率変数 Y_i の平均
Z_i	: 期待値が 0 である i 番目の確率変数 ($Z_i = Y_i - \mu_i$)
$\bar{Z}$	: 確率変数 $Z_i, \forall i$ の算術平均
σ_i^2	: i 番目の確率変数 Z_i の分散
σ^2	: 全ての確率変数 $Z_i, \forall i$ それぞれの分散の算術平均
δ	: 全ての確率変数 $Z_i, \forall i$ に対する値域の上限値

θ	: Hoeffding の確率不等式に用いる非負の定数
$\eta(Z)$	: $e^{\frac{Z_i}{n}\theta}$ を上から抑え込む二次関数
α	: 二次関数 $\eta(Z)$ における二次の項の係数
β	: 二次関数 $\eta(Z)$ における一次の項の係数
γ	: 二次関数 $\eta(Z)$ における定数項
$\Phi(\bar{Z})$	: 平均 $\bar{Z}$ となる n 個の Z_i を用いた $\prod_{i=1}^n \eta(Z_i)$ の最小値
$\Psi(\zeta, N)$	: 平均 ζ となる N 個の Z_j を用いた $\prod_{j=1}^N \eta(Z_j)$ の最小値
m	: 複号根解における一方の解の個数
$M(m)$	: m の複号根解が存在できるかを判別する条件式
H	: $Z_i = \delta$ となる確率変数の個数

第1章 序論

データを活用し科学的根拠に基づいて客観的に意思決定を行うことは、経験や勘などの暗黙的知見を補助し、また思い込みによる誤った判断を避けるために重要である。特に近年では、Internet of Things (IoT) 技術の発達などにより、多様かつ大量のデータを取得し、それをリアルタイムで共有することが可能となった。技術革新に呼応する形で、企業ではデジタル・トランスフォーメーション (DX) のもとデータの取得とその活用を前提としたサービスや企業自体の変革が進められている。このような昨今の状況において、取得したデータから意思決定に役立つ有用な情報を抽出する方法への期待が高まっている。

モノづくりの現場では、従来より抜取検査や QC 七つ道具 [24] をはじめとした統計的品質管理手法・統計的工程管理手法を利用してデータに基づく意思決定が行われてきた。DX の波が押し寄せている中で、モノづくりの現場においても、より高度なデータ活用が実施されている。例えば、工程内の多様なデータを IoT 機器によって自動的に収集することを通して、より詳細で多様なデータが意思決定に利用されている。さらに、製造現場から得たデータをクラウド上に集約することで工場間をまたいだ巨大なデータベースの構築も行われている [35]。収集したデータを品質管理や工程管理に活用する際には、数理統計学に加えてベイズ理論や情報理論も活用されるようになりつつある [27]。また計算機の性能向上の恩恵もあり、事前データが十分に存在する場合には人工ニューラルネットワーク (ANN) も利用されている。データの収集・解析技術が向上したことでデータドリブンな判断が増加し、従来では作業者の経験や勘に頼っていた判断をコンピュータが担うようになりつつある。

IoT 技術の発達で大量のデータを得られるようになる一方で、必ずしも意思決定の対象に関する事前情報が十分に得られるとは限らない。例えばモノづくりの現場では需要の多様化によって同一品種の大量生産から多品種・小ロットでの生産へ移行し、さらに同一品種を製造する期間も短くなる傾向にある [32]。このため新規に立ち上げられた工程の管理に携わる機会が増加し、他方で一つの工程を継続的に改善する機会は減少している。新たなシステムを対象とする場合には、そのシステムに関するデータが不足する中で意思決定を求められる。このような場合、ANN のような大量の事前データを前提とした手法を意思決定に用いることは

難しい。

ただし事前情報が不足する場合においても、限られたデータを最大限活用することで、データに基づいた客観的な意思決定を行うことは重要である。ここで数理統計学や情報理論を活用して管理対象の状態を可能な限り正確に推察する手法は、限定的な事前情報下での客観的な意思決定を可能にする。本論文では、変化点検出と確率不等式という数理統計学の基づいた二つの手法を取り上げ、その性能改善を行った。

まず変化点検出については、二つの複数変化点検出手法を提案する。変化点検出は時系列に並んだ一連の確率変数に対して、確率分布のパラメータが変化した時点を推定する手法である。意思決定を補助する手法としては、工程から得たデータに適用することで、製造工程において生じた異常原因を究明する作業に活用できる。

製造工程の異常原因究明には、QC 七つ道具の一つである管理図 [36] を用いた統計的工程管理が広く実施されている。管理図は工程から得たデータがある確率分布に従うと仮定したうえで、取得したデータの変動が偶発要因と異常要因のいずれに起因するか判別する手法である。工程に異常が生じたならば、その影響により工程から得られるデータは通常とは異なる値に変化するであろう。管理図はこれを利用して、工程から得た統計量があらかじめ定められた管理限界を超えた場合に工程で異常が生じたと判定する。管理限界は統計的仮説検定の考えに基づいて決定されるため、異常発生をデータに基づき客観的に判断できる手法である。管理図を用いて工程管理を実施する場合は、工程から定期的にデータを取得して管理図へ打点し、もし管理限界を超えたデータが出現すれば工程に生じた異常原因の究明作業に取り掛かる。

しかしながら、異常の発生を検知できたとしても、工程内には異常原因となりうる多数の要因が存在するため、真の異常原因を特定することは容易でない。管理図によって工程異常が検知されたとしても、第2種の誤りの存在から管理図が工程異常を検知した時点が工程異常の発生時点であるとは限らない。管理図では管理限界の他に点列の傾向を見て工程異常を判断する方法が例示されているものの、こちらについても異常発生 of 具体的な時点を示すものではない。例えば稲本と細野 [37] は、管理図が管理外れ状態を検知した際に技術的立場から異常原因を検討する手順を次のように紹介している。

- 使用原材料に異常はなかったかどうか。
- 使用機械に異常はなかったかどうか。
- 作業標準どおりの作業がなされたかどうか。
- 作業標準が適切なものであったかどうか。

以上をよく調べる．いままでの経験や技術的知識により，製品の品質特性に影響を与える原因について，起こりやすいものから順に調べていくようにするとよい．

すなわち，異常原因の究明では想定される原因を列挙してしらみつぶしに検証していくことになり，効率的に実施するには経験や勘といった暗黙的知見が必要になる．現状の工程管理では，異常検知の段階では管理図によりデータを用いた判断が実施されているが，その後の異常原因の究明においてはデータ活用による効率化が十分であると言えない．

異常原因を検討する作業において変化点検出を用いれば，データを活用した科学的根拠によるアプローチで異常原因に関する手がかりを得ることができる．管理図の打点に用いられる一連のデータに対して変化点検出を適用することで，工程から得られるデータの傾向が変化した時点（変化点）を推定できる．この変化点以前において，工程は理想的な状態（管理状態）を維持していたと考えられる．逆に変化点以後では工程が通常とは異なる状態（管理外れ状態）に陥っているが，その変化は異常原因により引き起こされたと推察される．変化点の前後で変更した要因（原材料，機械の状態，作業者など）があれば，それが工程異常の原因である可能性が高い．推定した変化点と工程に関する記録を照らし合わせて異常原因の候補を絞り込むことで，異常原因の究明作業を効率的に実施できる．

変化点検出手法は対象とする工程および統計量の特徴，想定される変化をはじめとした諸々の条件に応じて様々な手法が提案されている [11]．諸条件の中でも，想定される変化は推定に用いる変化点検出手法の選定において大きな影響がある．異常によりデータの傾向が変化する際，その変化が突発的であるか持続的であるかなど，変化点以後のデータの推移の様相には様々なパターンが存在し，各々のパターンに対応した統計モデルが提案されている [31]．以降はこのような変化点以後の変化の様相を変化パターンと呼称する．例えば突発的な変化に対応した変化パターンとしてステップ変化 [16]，持続的な変化に対応した変化パターンとして線形変化 [12] が存在する．Perry and Pignatiello [12] はシミュレーションを通して，データが線形関数に基づいて持続的に変化する場合にはステップ変化を想定した手法 [16] より線形変化を想定した手法の方がより正確に変化点を推定できることを示した．実際の変化パターンに対応した変化点検出手法を用いることは正確な推定に必要であるから，想定されうる多様な変化パターンそれぞれに対応した変化点検出手法を開発することは重要である．

より多様な変化への対応の一例として，一連のデータ内に複数回の変化が生じていることを想定した複数変化点検出手法が挙げられる．前述のステップ変化や線形変化を想定した変化点検出手法は，いずれも工程に生じる変化が1回のみで，変化点以降はパラメータ変化の傾向が変わらないことを想定している．一方で，工

程において複数の異常が生じる，あるいは一つの異常原因の影響が複数回にわたり生じることで，工程の状態が複数回変化することも想定される．このような場合，1個の変化点を推定するだけでは，管理状態から管理外れ状態へ移行した箇所しか取り上げることができない．もしパラメータが変化した各々の時点を推定できれば，各々の変化点ごとに異常原因の候補を探す作業が可能となるため，異常原因究明のさらなる効率化が実現できる．これを踏まえて，パラメータが変化した時点を変化点と定義し，複数回のパラメータ変化を推定する複数変化点検出が提案された．Hawkins [6] は指数型分布族に対し，尤度理論に基づいて複数個の変化点を推定する手法を提案した．この手法はあらかじめ定められた変化点数に応じた複数個の変化点を推定するものであり，変化点数が同じモデルの比較に利用できる．ただし当該手法を用いて変化点数が異なるモデルを比較して的確に推定することは難しいため，データから最適なモデルを一意に定めることはできない．この課題に対して，竹本と有蘭 [26] は変化点数が異なるモデルをパラメータ数が異なる統計モデルとして捉え，尤度理論より派生した情報量規準に基づいて推定する手法を提案した．具体的には赤池情報量規準 (AIC) [1] を用いることで，統計量が正規分布に従う場合を対象にステップ変化による複数個の変化点を最適な変化点数とともに推定する手法を実現した．

前述のように変化点検出手法は統計量の特徴や変化パターンに応じて手法を設計する必要がある．複数変化点検出手法としては統計量が二項分布に従う状況を対象とした Takemoto and Arizono [18] 等が提案されているものの，製造工程で想定される多様な状況に対応しきれていない．これを踏まえて，本論文では異なる二つの状況について，それぞれに対応した複数変化点検出手法を提案する．

第2章ではポアソン分布を想定した複数変化点検出を検討する [30, 19]．ポアソン分布は希少事象の発生回数の分布として利用される離散確率分布である．JIS 規格 [36] にも統計量がポアソン分布に従う場合に利用する管理図である c 管理図や u 管理図が規定されるなど，統計的工程管理において重要な確率分布の一つである．第2章では統計量がポアソン分布に従う場合を対象として，複数回のステップ変化を変化回数とともに推定する手法を提案するとともに，その活用方法について考察する．

第3章では複数回の線形変化を想定した変化点検出手法を提案する [17]．これまでの複数変化点検出手法は複数回の突発的な変化を想定していたが，持続的な変化が複数回生じる場合も想定されうる．変化点の正確な推定には変化パターンに対応した変化点検出手法を用いることが重要であることを鑑み，複数回の持続的な変化を対象とした統計モデルならびに情報量規準を用いた推定手法を提案する．その具体例として，正規分布に従うデータの平均の変化を推定する複数線形変化点検出手法を構築する．

また、本論文では Hoeffding の確率不等式の性能改善も行った。対象とする確率変数が従う分布形が規定されない状況において、意思決定を求められる場合がある。このような状況であっても合理的な意思決定を実現するには、確率変数に関する限られた情報に基づいて意思決定を行う具体的な方法が重要となる。そのような方法の一つが確率不等式である。

確率不等式は確率分布の情報を必要とせずに、平均や分散などの限られた情報を用いて生起確率の上界（あるいは下界）を計算する方法である。もっとも基本的な確率不等式として知られる Markov の不等式 [38] は、確率変数の期待値のみを用いて、上側確率の上界を与える不等式である。確率変数の期待値と分散を利用する手法には、期待値の偏差に関する両側確率の上界を評価する方法として Chebychev の確率不等式、期待値の正の偏差がある基準値よりも大きくなる確率の上界を評価する方法として片側 Chebychev の確率不等式が存在する [5]。複数個の確率変数について期待値、分散および定義域の上限が既知であれば、確率変数の算術平均の上側確率の上界を評価する方法として Hoeffding の確率不等式 [7] が利用できる。

前述の通り、確率不等式が意思決定に用いられるのは、確率変数として捉えられるデータの確率分布は特定できないが一部の統計量は取得可能な場合である。特に、上側確率が何かしらの危険率に相当する場合には、確率不等式に基づき行われる意思決定は安全サイドに立脚して余裕を持たせた意思決定として有用である。竹本ら [28] はリード・タイム期間の需要分布が未知である状況における発注点在庫管理方式での発注点の決定方法に関して Hoeffding の確率不等式に基づく考察を与えている。また有蘭と竹本 [20] および Arizono and Takemoto [2] は、電力需給バランスが逼迫する状況において適切な節電要請を行うことを目的として、確率不等式に基づく適正節電要請率決定のための評価システムを提案している。

確率不等式により与えられる上側確率の上界は、実際の上側確率（真値）に対して必ず余裕のある値として評価される。ここで確率不等式が与える上側確率の上界値を真値に対して上側からよりタイトに与えることができれば、安全性を維持しながらより効率的な意思決定が実現できる。

これを踏まえて第 4 章で Hoeffding の確率不等式の性能改善を行う [29, 21]。Hoeffding の確率不等式の導出過程を精査し一部の計算を置き換えることで、確率変数の平均・分散・定義域の上限という同じ情報を用いながらも、よりタイトな上側確率の上界値を与える確率不等式を実現する。第 4 章の内容は、棚橋ら [29] および有蘭ら [21] にて提案した確率不等式について、その導出課程の一部を見直し洗練させたものである。

最後に第 5 章で結論として各章の内容を要約し、本論文を総括する。

第2章 ポアソン分布を想定した複数 変化点検出手法

2.1 緒言

ポアソン分布は離散確率分布の一種であり、希少現象の発生回数を表す分布として知られる。製造工程においては、製品表面に生じた傷の数などの欠点数がポアソン分布に従う。そのため、ポアソン分布は統計的工程管理において主要な確率分布の一つである。例えば、ポアソン分布に従うデータを対象とした管理図として c 管理図や u 管理図が JIS Z9020 [36] に規定されている。観測データがポアソン分布に従う過程はポアソン過程と呼ばれる。

変化点検出においてもポアソン過程を対象とした手法が提案されている。Samuel and Pignatiello [15] はステップ変化モデルに基づく変化点を尤度理論を用いて推定する手法を提案している。ステップ変化モデルは異常発生により確率分布のパラメータの値が突発的に別の値に変化するモデルである。Perry et al. [13] は線形変化モデルに基づく変化点を尤度理論を用いて推定する手法を提案している。線形変化モデルは管理外れ状態にて経時的にパラメータの値が変化することを想定したモデルであり、変化点以後のパラメータが線形関数にしたがって増加（あるいは減少）するものとして推定する。Perry et al. [14] は単調変化モデルに基づく変化点を尤度理論を用いて推定する手法を提案している。単調変化モデルは管理外れ状態においてパラメータが単調増加（あるいは単調減少）すると仮定して、管理状態を維持する最後の時点を変化点として推定するものである。例えば、ステップ変化や線形変化も単調変化に包含される。

これらの尤度理論に基づく変化点推定手法には二つの欠点がある。一つは、管理状態におけるデータの確率分布のパラメータが既知でなければ利用できない点である。基本的に、管理状態の工程から得られるデータの特徴は、これまで蓄積されてきたデータより明らかになっている。これを踏まえて、変化点検出では管理状態のデータが従う確率分布のパラメータの値を利用して、変化点を推定している。しかし、工程の操業開始直後など、工程が従う確率分布のパラメータの正確な値が未知である場合も存在しうる。その場合、管理状態のパラメータの値に

依存する手法は利用できない。

もう一つは、管理状態と管理外れ状態を峻別する時点しか推定できない点である。ポアソン過程においても、工程に複数回の変化が生じる状況は想定されうる。そのような場合には、管理状態を維持した最後の時点だけを推定するのではなく、変化が生じた時点全てを变化点として推定したほうが、異常原因の究明作業に役立つより多くの情報を得られる。しかしステップ変化モデルに基づく Samuel and Pignatiello [15] は変化が1回のみ生じることを想定している。線形変化モデルに基づく Perry et al. [13] は管理外れ状態においてパラメータが変化し続けるものの、管理状態における変化量は一定のまま変化しないため、工程に生じる変化は実質的に1回のみと解釈される。複数回の変化が生じた工程のデータに上記のような単一の変化を想定した変化点検出を適用した場合、目的とすべき時点とは異なる時点を推定してしまい、かえって混乱を招く恐れがある。もし複数回の変化によってパラメータが単調増加（あるいは単調減少）するのであれば単調変化モデルに基づく Perry et al. [14] を用いて推定可能である。しかし、単調変化モデルは、管理状態と管理外れ状態を峻別する1個の変化点しか推定できない。工程に生じた変化の詳細を明らかにするには、複数個の変化点を推定可能な変化点検出手法が必要である。

ポアソン過程を対象とした複数変化点検出手法として、Assareh et al. [3] はベイズ統計学に基づいて複数の変化点を変化回数を含めて推定する手法を提案している。この手法では一連のデータにおいて複数回のステップ変化が生じる状況を想定している。以降ではこのようなモデルを複数ステップ変化モデルと呼称する。当該手法ではマルコフ連鎖モンテカルロ法 (MCMC) をパラメータ数の異なる統計モデルの比較に拡張したリバーシブル・ジャンプ・マルコフ連鎖モンテカルロ法 (RJMCMC) を用いることで、変化点数の異なるモデルを比較している。これにより一連のデータから変化回数を含めた変化点に関する事後分布を計算している。

ただし、この手法は計算量という観点において課題がある。MCMC は大量の乱数を用いた計算が必要であり、モデル推定に時間がかかる。計算量が多いという点は、特に昨今注目されている IoT を用いたシステムにおいて当該手法を用いる妨げとなる。IoT で用いられるシステム構成として、従来のように生データ全てをサーバに集約してサーバ側で処理するのではなく、データ収集に用いる端末にて処理を行い要約された結果をサーバに送信するフォグコンピューティング [8] が存在する。末端側には比較的性能の低い計算機が用いられることもあり、計算量の多いアルゴリズムを用いるとサーバに要約されたデータが届くまでのラグが大きくなってしまう。この観点からできる限り高速に計算可能な推定方法には価値がある。

本章では以上のことを踏まえて、尤度理論や情報理論を用いて変化点とその個

数を代数的に求めることで高速なモデル推定を目指す．また幅広い状況での活用を目指し，工程の管理状態におけるパラメータが未知であっても適用可能な手法を実現する．

また高速なモデル推定ができるならば，工程から新たなデータを得るたびに変化点検出手法を適用することで，リアルタイムに工程の状態を推定することが可能となる．このように変化点検出を工程の監視に用いる利用方法についても検討する．

2.2 ポアソン過程に対する複数変化点検出手法の設計

確率変数 X の確率密度関数（あるいは確率質量関数）を $f_X(X)$ と表記する．ポアソン分布の確率質量関数は

$$f_x(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad (2.1)$$

で定義される．ここで λ はポアソン分布のパラメータであり，強度と呼ばれる．また確率変数 x の値域は非負の整数である．

工程の評価に用いる統計量として，事象の発生数などのポアソン分布に従うデータ x_t を利用する状況を扱う．ここで t はサンプルを取得した期数であり，第 T 期までに各期に対応する T 個のデータ $\mathbf{X} = (x_1, x_2, \dots, x_T)$ を得たとする． $k+1$ 個目の変化点を τ_k ，変化回数を K とする．ポアソン分布は強度 λ で特徴づけられることから， k 回目の変化以降の強度を λ_k を表記すれば，複数ステップ変化を想定した統計モデルは

$$\left\{ \begin{array}{ll} \lambda_0, & t = 1, \dots, \tau_0 \\ \lambda_1, & t = \tau_0 + 1, \dots, \tau_1 \\ \vdots & \\ \lambda_k, & t = \tau_{k-1} + 1, \dots, \tau_k \\ \vdots & \\ \lambda_K, & t = \tau_{K-1} + 1, \dots, \tau_K (= T) \end{array} \right. \quad (2.2)$$

で定義できる． $\tau_K = T$ はデータを取得した最後の時点であるから，必ずしも変化点を意味するものではない．この統計モデルにおける推定量は，変化点数 K ，変化点 $\boldsymbol{\tau}^{(K)} = (\tau_0, \dots, \tau_{K-1})$ およびそれに対応した強度 $\boldsymbol{\lambda}^{(K)} = (\lambda_0, \dots, \lambda_K)$ である．ここで尤度を利用して推定する従来の1回変化点検出手法 [15, 13, 14] や複数変化点検出手法 [26] では初期のパラメータを管理状態の値として既知で与えていた．しかしながら本手法では，初期のパラメータ λ_0 も未知の値として推定する．

式 (2.2) の統計モデルを想定した場合、尤度関数は

$$L(\boldsymbol{\tau}^{(K)}, \boldsymbol{\lambda}^{(K)} | \mathbf{X}, T) = \prod_{k=0}^K \prod_{t=\tau_{k-1}+1}^{\tau_k} \frac{\lambda_k^{x_t} e^{-\lambda_k}}{x_t!} \quad (2.3)$$

である。ただし $\tau_{-1} \equiv 0$ と定義する。さらに対数尤度関数は

$$\log L(\boldsymbol{\tau}^{(K)}, \boldsymbol{\lambda}^{(K)} | \mathbf{X}, T) = \sum_{k=0}^K \sum_{t=\tau_{k-1}+1}^{\tau_k} \{x_t \log(\lambda_k) - \lambda_k - \log(x_t!)\} \quad (2.4)$$

である。

式 (2.2) の複数変化点モデルに基づいて、尤度理論を利用して変化点ならびに強度を推定する。任意の $\lambda_k \in \boldsymbol{\lambda}^{(K)}$ について、 λ_k を用いた対数尤度関数の偏微分より

$$\begin{aligned} \frac{\partial}{\partial \lambda_k} \log L(\boldsymbol{\tau}^{(K)}, \boldsymbol{\lambda}^{(K)} | \mathbf{X}, T) &= \sum_{t=\tau_{k-1}+1}^{\tau_k} \left(\frac{x_t}{\lambda_k} - 1 \right) \\ &= 0 \end{aligned} \quad (2.5)$$

が定義できる。したがって強度の最尤推定量 $\hat{\lambda}_k$ は、その前後の変化点 τ_{k-1}, τ_k が与えられれば

$$\hat{\lambda}_k = \frac{1}{\tau_k - \tau_{k-1}} \sum_{t=\tau_{k-1}+1}^{\tau_k} x_t \quad (2.6)$$

で計算できる。式 (2.6) を用いることで、任意の変化点の組合せ $\boldsymbol{\tau}^{(K)}$ に対する強度の最尤推定量 $\hat{\boldsymbol{\lambda}}^{(K)} = (\hat{\lambda}_0, \dots, \hat{\lambda}_K)$ を一意に定めることができる。

変化点の組合せ $\boldsymbol{\tau}^{(K)}$ に対応した強度の最尤推定量 $\hat{\boldsymbol{\lambda}}^{(K)}$ を式 (2.4) に代入することで全てのデータに対する対数尤度を計算できる。したがって個々の変化点の最尤推定量 $\hat{\tau}_k$ は独立して計算されるのではなく、変化点の組合せの最尤推定量 $\hat{\boldsymbol{\tau}}^{(K)} \equiv (\hat{\tau}_0, \dots, \hat{\tau}_{K-1})$ として

$$\hat{\boldsymbol{\tau}}^{(K)} = \arg \max_{\boldsymbol{\tau}^{(K)}} \log L(\boldsymbol{\tau}^{(K)}, \hat{\boldsymbol{\lambda}}^{(K)} | \mathbf{X}, T) \quad (2.7)$$

の関係より求められる。しかしながら、変化点数によっては考え得る変化点の組合せが膨大になることから、全ての変化点の組合せに対して対数尤度を計算することは計算量の観点から課題がある。ここで、対数尤度ならびに強度の最尤推定量は連続する2個の変化点 τ_{k-1}, τ_k 毎に独立して計算可能であることを利用し、動的計画法を導入することで計算量の削減を図る。

対数尤度関数のうち、2 個の連続する変化点 τ_{k-1}, τ_k に挟まれた区間を抜き出した

$$\Lambda(\tau_{k-1}, \tau_k, \lambda_k \mid \mathbf{X}) = \sum_{t=\tau_{k-1}+1}^{\tau_k} (x_t \log(\lambda_k) - \lambda_k) \quad (2.8)$$

を定義する．ここで式 (2.4) に含まれる $\sum_{k=0}^K \sum_{t=\tau_{k-1}+1}^{\tau_k} \log(x_t!)$ の項は変化点の組合せ $\boldsymbol{\tau}^{(K)}$ に依存しないため省略した．また式 (2.8) の最大値は、式 (2.6) で求めた強度の最尤推定量

$$\begin{aligned} \max \Lambda(\tau_{k-1}, \tau_k, \lambda_k \mid \mathbf{X}) &= \Lambda(\tau_{k-1}, \tau_k, \hat{\lambda}_k \mid \mathbf{X}) \\ &= \sum_{t=\tau_{k-1}+1}^{\tau_k} (x_t \log(\hat{\lambda}_k) - \hat{\lambda}_k) \end{aligned} \quad (2.9)$$

である．式 (2.9) を用いれば、ある時点 t までに k 回の変化が生じた場合の対数尤度関数の最大化問題は

$$D(k, t) = \max_{k \leq \tau_{k-1} < t} \left\{ D(k-1, \tau_{k-1}) + \Lambda(\tau_{k-1}, t, \hat{\lambda}_k \mid \mathbf{X}) \right\} \quad (2.10)$$

となる．式 (2.10) の境界条件は

$$D(-1, \tau_{-1}) \equiv 0 \quad (2.11)$$

である．所与の変化点数 K のもとでの変化点の組合せの最尤推定量は式 (2.7) の代わりに式 (2.10) を用いた

$$\hat{\boldsymbol{\tau}}^{(K)} = \arg \max_{\boldsymbol{\tau}^{(K)}} D(K, T) \quad (2.12)$$

で計算できる．

式 (2.10) の計算には動的計画法が適用できる．小問題 $D(k-1, \tau_{k-1})$ を計算すれば、変化点数と最後の時点の組合せ $k-1, \tau_{k-1}$ における最適な変化点の組合せ $\hat{\boldsymbol{\tau}}^{(k-1)}$ が与えられる．これにより $D(k, \tau_k)$ を計算する際には、変化点の組合せ $\boldsymbol{\tau}^{(k)}$ 全てを網羅せずとも、 τ_{k-1} の値それぞれに対応する $\hat{\boldsymbol{\tau}}^{(k-1)}$ が含まれる組合せだけ検討すればよい．また変化点数 K が異なるモデルに対して対数尤度を計算する際に小問題 $D(k, t)$ の計算結果を流用可能である．動的計画法を用いれば、計算を必要とする変化点の組合せの削減と対数尤度の計算結果の流用により、対数尤度関数の最大化問題に必要な計算量を削減できる．具体的には、変化点の組合せを全探索するなら $O(2^T)$ のところ、動的計画法を用いれば $O(T^2)$ となる．特に、既に

推定したデータ列に新たなデータを1個追加した際には、 $D(k, t)$ をデータ追加前の計算から流用できるため、 $O(T)$ の計算量でデータ追加後の結果を得られる。

以上のように、式 (2.6) と (2.12) を用いることで、任意の変化点数 K における強度と変化点の最尤推定量を得ることができる。しかしながら、変化点数の異なるモデルの比較を尤度関数に基づいて行うことはできない。尤度関数の値はモデルのパラメータの値が観測値に近しいほど高くなる。したがって、尤度関数の最大化のみを考えた場合には観測データ毎に変化点を想定するモデル、すなわち変化点数 $K = T - 1$ のモデルが選択される。その場合、全ての時点が変化点であると推定され、また強度の推定値 $\hat{\lambda}^{(T-1)}$ は各観測時点でのデータの値そのものとなる。推定値をデータそのものに一致させてモデルの適合度を向上させることは、過適合な状態であるといえる。過適合な統計モデルの推定は、データに基づいてプロセスの変化を解釈するうえで意味をもつ変化点を抽出するという目的に合致しているとはいえない。

過適合が生じる一因に、尤度計算におけるデータ $\mathbf{X}$ の扱い方がある。式 (2.9) の右辺にて $\hat{\lambda}_k$ が利用されているように、尤度理論に基づいた比較の際にパラメータの最尤推定量が用いられる。データ $\mathbf{X}$ より求めたパラメータの推定値に対してまったく同じデータ $\mathbf{X}$ を用いて尤度を計算するため、真のパラメータを用いる場合より尤度が高めに計算される。この同じデータを二重に使用することで尤度が過大評価される傾向は、データへのフィッティングが高くなるパラメータ数の多いモデルでより顕著に現れる。これが原因となりパラメータ数の異なるモデルの比較が難しくなっている。

そこで情報量規準を援用して変化点数の異なるモデルを比較する。情報量規準では、尤度計算に最尤推定量を用いることで生じるバイアスの補正を、尤度からペナルティ項を差し引く形で実現する。今回は計算の簡便さに優れる AIC を用いて変化点数の異なるモデルを比較する。

AIC は式 (2.4) を用いて

$$AIC(K) = -2 \log L(\hat{\tau}^{(K)}, \hat{\lambda}^{(K)} | \mathbf{X}, T) + 2B(K) \quad (2.13)$$

で定義され、その値が小さいモデルほどデータに対して適切なモデルとなる。ここで $B(K)$ がペナルティ項であり、AIC ではモデルのパラメータ数として定義される。式 (2.2) の統計モデルで推定されるパラメータは $\tau^{(K)} = (\tau_0, \dots, \tau_{K-1})$ と $\lambda^{(K)} = (\lambda_0, \dots, \lambda_K)$ であるから、提案モデルにおいては

$$B(K) = 2K + 1 \quad (2.14)$$

となる。AIC に含まれる対数尤度の計算に式 (2.10) を代用することで、変化点数は

$$\hat{K} = \arg \min_{0 \leq K < T} \{-D(K, T) + 2K + 1\} \quad (2.15)$$

で推定できる。

以上のように，式 (2.2) の統計モデルに基づけば，ポアソン分布のパラメータを式 (2.6)，変化点数を式 (2.15)，変化点の位置を式 (2.12) で推定できる。

なお式 (2.2) で定義した統計モデルを用いて，従来の変化点検出手法と同様に管理状態の値を既知とした変化点推定手法も設計可能である．その詳細は付録 A に示す。

2.3 数値検証

本章では提案した複数変化点検出手法の性能をモンテカルロシミュレーションにより確認する．具体的には，複数回のステップ変化を想定してポアソン分布に従う乱数列を生成し，これに対して提案手法で変化点を推定する。

まず c 管理図で管理される工程を想定したシミュレーションを実施した．乱数生成に用いる強度変化のシナリオとして

$$\begin{cases} \lambda_0 = 20, & t = 1, \dots, \tau_0 = 25 \\ \lambda_1 = 24, & t = \tau_0 + 1, \dots, \tau_1 = 35 \\ \lambda_2 = 28, & t = \tau_1 + 1, \dots, T \end{cases} \quad (2.16)$$

を定義した．ここで時点 T は c 管理図が異常を検知した時点である．ここで管理状態における強度は $\lambda_0 = 20$ であるから， c 管理図の上側管理限界と下側管理限界はそれぞれ $UCL = 33.4, LCL = 6.6$ となる。

図 2.1 に式 (2.16) のモデルから生成したデータを提案手法で解析した結果を示す．図では推定された変化点の時期をグレーの縦線で示している．提案手法は実際の変化点数と同じ $\hat{K} = 2$ を推定した．最初の変化点の推定値は $\hat{\tau}_0 = 25$ であり，これは実際の最初の変化点 τ_0 と一致している．2 個目の変化点の推定値は $\hat{\tau}_1 = 38$ であり，実際の 2 個目の変化点 $\tau_1 = 35$ に近い値を推定している．変化点検出の先行研究 [15, 13, 14, 26] では管理状態のパラメータを既知として与えているが，手法提案手法では最初の変化点以前のパラメータも推定すべき未知変数としている．今回のデータに対しては $\hat{\lambda}_0 = 19.2$ と実際の管理状態に近い値を推定している．また変化点以後の強度の推定値は $\hat{\lambda}_1 = 23.6, \hat{\lambda}_2 = 29.0$ であり，2 個の変化点の両方において強度が上昇する様子を推定できている。

次に，提案手法は管理状態の値 λ_0 が未知の場合でも利用可能であることを鑑みて，管理図等による工程管理が行われていない状況を想定したシミュレーション

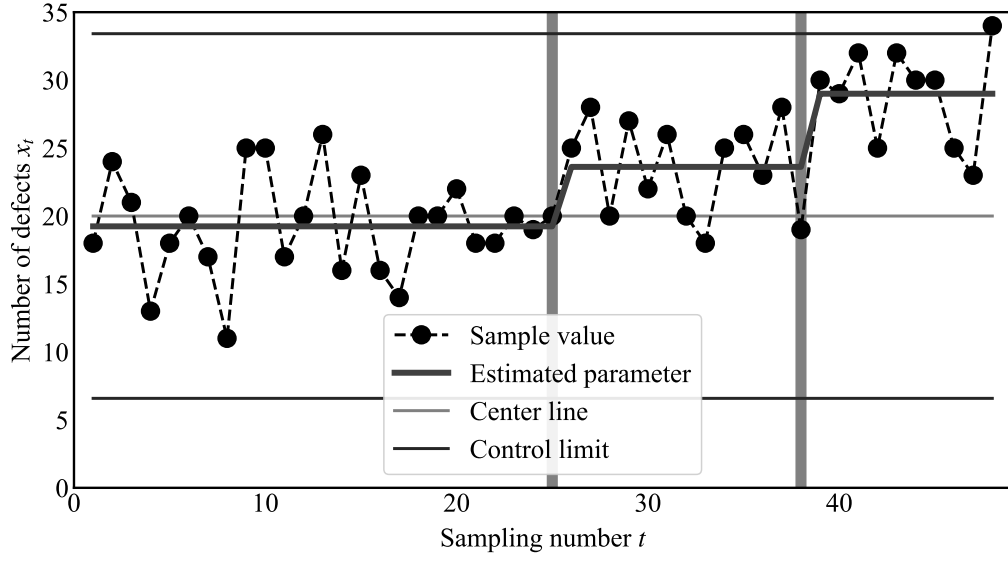


図 2.1: 式 (2.16) に従う乱数列に対する提案手法での解析結果

を実施した。乱数生成に用いる強度変化のシナリオとして

$$\begin{cases} \lambda_0 = 3, & t = 1, \dots, \tau_0 = 15 \\ \lambda_1 = 6, & t = \tau_0 + 1, \dots, \tau_1 = 20 \\ \lambda_2 = 10, & t = \tau_1 + 1, \dots, \tau_2 = 25 \\ \lambda_3 = 14, & t = \tau_2 + 1, \dots, \tau_3 = 35 \\ \lambda_4 = 20, & t = \tau_3 + 1, \dots, T = 40 \end{cases} \quad (2.17)$$

を定義した。

図 2.2 に式 (2.17) のモデルから生成したデータを提案手法で解析した結果を示す。提案手法は実際の変化回数と同じ $\hat{K} = 4$ を推定した。変化点とそれに対応する強度の推定値はそれぞれ $\hat{\tau}^{(K)} = (12, 20, 26, 35)$, $\hat{\lambda}^{(K)} = (2.7, 5.5, 9.0, 14.2, 21.2)$ であり、複数回の変化点にて強度が徐々に増加していく様子が正しく推定されている。

最後に、これまでの 2 例とは異なり、パラメータが途中で減少する状況を想定したシミュレーションを実施した。乱数生成に用いる強度変化のシナリオとして

$$\begin{cases} \lambda_0 = 3, & t = 1, \dots, \tau_0 = 15 \\ \lambda_1 = 10, & t = \tau_0 + 1, \dots, \tau_1 = 25 \\ \lambda_2 = 15, & t = \tau_1 + 1, \dots, \tau_2 = 35 \\ \lambda_3 = 5, & t = \tau_2 + 1, \dots, T = 40 \end{cases} \quad (2.18)$$

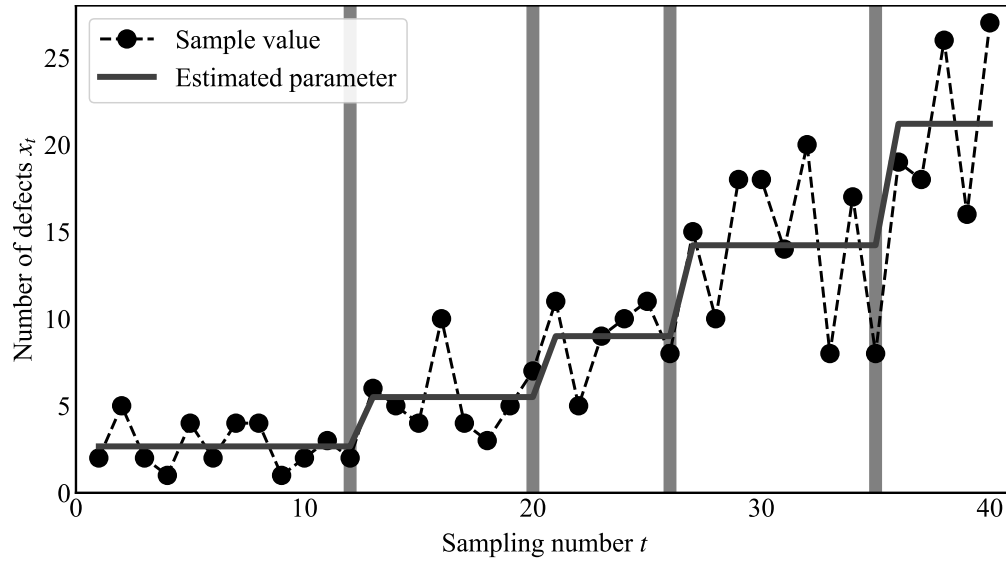


図 2.2: 式 (2.17) に従う乱数列に対する提案手法での解析結果

を定義した.

図 2.3 に式 (2.18) のモデルから生成したデータを提案手法で解析した結果を示す. 提案手法では変化点数を $\hat{K} = 3$, それぞれの変化点を $\hat{\tau}^{(K)} = (15, 25, 35)$ と正確に推定している. 強度の推定値については $\hat{\lambda}^{(K)} = (2.8, 9.8, 14.2, 4.2)$ と最後の変化点で強度が減少したことを正しく推定している.

以上のシミュレーションを通して, 提案手法がポアソン過程における複数回のステップ変化を, 変化回数も含めて推定可能であることが確認できた.

2.4 提案手法のリアルタイムな工程解析への活用

変化点検出手法を統計的工程管理に用いる場合, 従来は管理図によって異常が検知された後に適用されてきた. これに対し竹本と有蘭 [26] は, 変化点検出手法を逐次適用することで, リアルタイムに工程の状況を推定することを提案している. このような運用方法は状態変化追跡方法と呼ばれる. 本節では提案手法を状態変化追跡方法として運用することについて考察する.

ポアソン分布を想定した従来の変化点検出手法では, 統計モデルの特性あるいは計算時間に懸念があるため, 状態変化追跡方法として運用することが困難である. それに対し提案手法は, この 2 点について好ましい特性を持っている.

1 点目は変化点検出手法が適用できる条件である. 従来の変化点検出手法のうち単一の変化点を推定する手法は, 想定する統計モデルの特徴からリアルタイムで

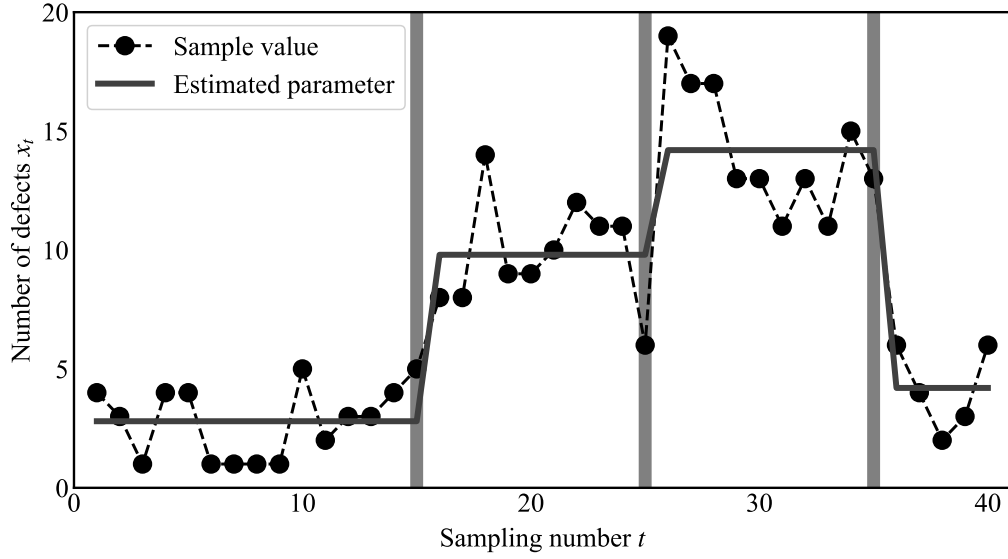


図 2.3: 式 (2.18) に従う乱数列に対する提案手法での解析結果

の運用に用いることが困難であった．変化点検出に関する研究は，管理図による工程管理を補完する役割を担うものとして提案されてきた．そのため，1 回ステップ変化 [15] や 1 回線形変化 [13]，単調変化 [14] などの 1 個の変化点で管理状態と管理外れ状態を峻別する変化点検出モデルは，解析対象のデータに必ずパラメータ変化が含まれることを前提としている．したがってこれらの手法が利用できるのは，管理図による管理外れ状態の検知後など，異常発生が確実視される場合に限られる．

これらの先行研究とは異なり，提案手法は変化点数を推定可能である．ここで式 (2.15) にあるとおり，提案手法は $K = 0$ の変化点が存在しない，すなわち工程が初期の状態を維持し続けている状況を推定することも可能である．したがって提案手法は異常が含まれないデータを扱うことが可能であり，管理図などの異常検知とともにリアルタイムのデータ解析に利用できる．

2 点目は計算量である．尤度理論を用いない複数変化点検出手法は，その計算の複雑さから短時間での推定が困難である．ベイズ理論に基づき統計モデルを推定する手法 [3] では，MCMC を利用するため大量の乱数を用いた計算が必要であり，計算に時間がかかる．さらに，リアルタイムでの解析ではサンプルを取得するたびに解析対象のデータ列 $\mathbf{X}$ が更新されるが，その度に新たなデータ列 $\mathbf{X}$ を利用して計算を再度やり直すことになる．

これに対して，提案手法は式 (2.10) の計算に動的計画法を導入したため，リアルタイムでの解析において高速なモデル推定が可能である．まず動的計画法を用いることで，直近の変化点 τ_{K-1} より前の評価値 $D(\tau_{K-1}, K-1)$ が τ_{K-1} より後

のデータとは独立して計算できる．したがって，第 $T+1$ 期のデータ x_{T+1} を追加して改めてモデルを推定する際に，第 T 期までのデータによるモデル推定で計算した $D(t, k), t \leq T, k < T$ を再利用できる．さらに動的計画法を用いることで，検討すべき変化点の組合せを削減できる．具体的に， $0 \leq K < T$ の全ての K に対して $D(T, K)$ が計算されていれば，新たなデータ x_{T+1} を追加した場合に検討すべき変化点の組合せは

- 追加されたデータ x_{T+1} にて新たに K 回目のパラメータ変化が起きた場合 ($1 \leq K \leq T$)
- 追加されたデータは直近の状態 λ_K を保ち続けている場合 ($0 \leq K \leq T-1$)

の高々 $2T$ 通りとなる．このように提案手法では，直前の計算結果の流用することで高速なモデル推定が可能である．

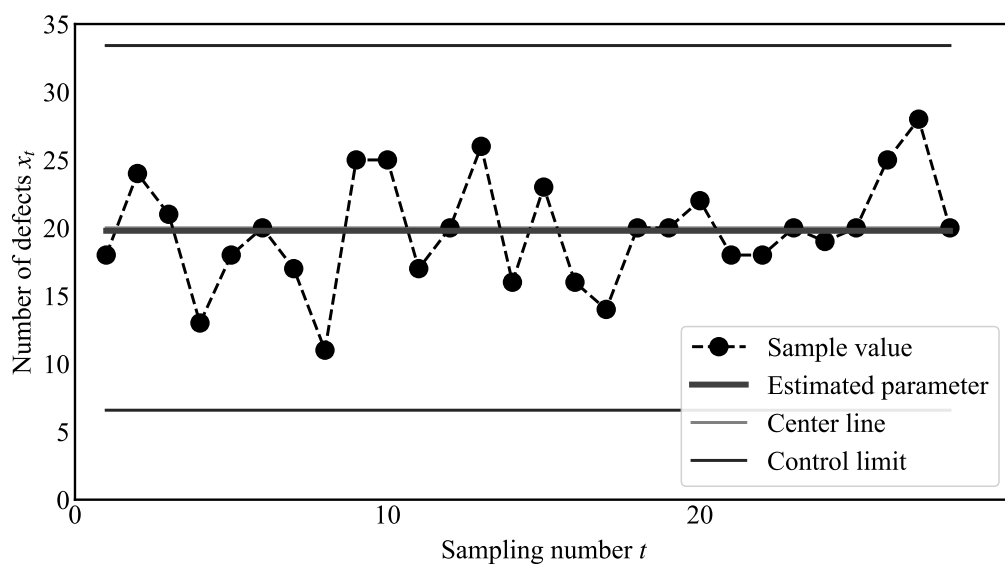
上記のように，提案手法は変化点数 $K=0$ のモデルを推定可能であること，動的計画法の利用によって高速なモデル推定が可能であることから，提案手法を状態追跡方法として利用できる．実際に状態追跡方法として運用した場合の性能を確認するため，モンテカルロシミュレーションを実施した．

図 2.4 に提案手法を状態変化追跡方法として運用した例を一部抜粋して示す．シミュレーションには図 2.1 と同じデータを利用した．乱数列は式 (2.16) の統計モデルに従うため，最初の変化点が $\tau_0 = 25$ に存在する．

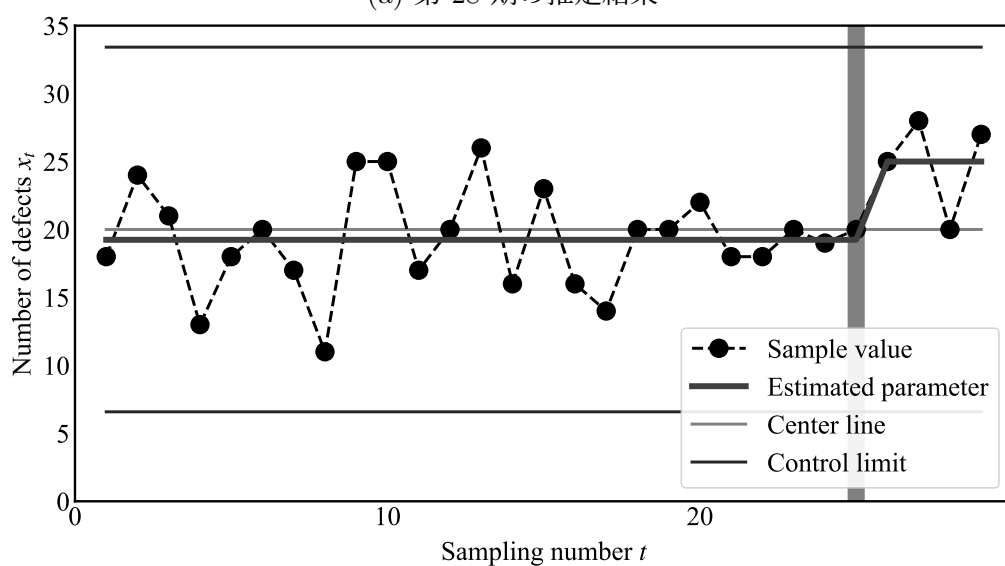
図 2.4(a) に第 28 期までに取得したデータに対して提案手法を適用した推定結果を示す．この時点では変化点数 $\hat{K} = 0$ すなわち変化点が存在しないと推定された．この結果からプロセス管理者は工程が初期状態を保ち続けていると判断し，作業を続行することになる．図 2.4(b) には次のデータが追加された，第 29 期までのデータに対して提案手法を適用した推定結果を示す．この時点における変化点検出では，変化点数 $\hat{K} = 1$ のモデルが選択された．変化点の位置は $\hat{\tau}_0 = 25$ であり，強度が変化点以後において上昇したと推定されている．

図 2.4 に示すとおり，データを逐次追加しながら提案手法を適用することで，工程平均の推移を可視化できる．これらの図には c 管理図による管理限界を併記しているが，図 2.4(b) に示した第 29 期の時点ではデータは管理限界を超えていない．このように，状態変化追跡方法として提案手法を運用すれば，工程に生じた異常による深刻な影響が出る前に，工程異常に気づくための情報を与えることが可能である．

ただし，統計的仮説検定に基づいて管理限界が定められている管理図とは異なり，提案手法による AIC を用いたモデル選択においては，誤推定の確率を保障するには至っていない．例えば変化点が検出されたがパラメータの変化量が小さいなど，誤推定の可能性を考慮する必要がある場合には，原因究明や工程改善のた



(a) 第 28 期の推定結果



(b) 第 29 期の推定結果

図 2.4: 提案手法の状態変化追跡方法としての運用例

めのコストが品質の変化による損失を超えるか判断し難い。したがって実際の運用においては、図 2.4(b) のように管理限界を超えない範囲で発見された状態変化への対処は現場の判断によるところが大きい。例えば、工程を運転させ続けて様子を見ることも一つの選択肢である。その場合は新たなデータを順次追加していくことで、強度や変化点、変化点数の推定値の推移から変化の傾向が観察できる。また管理図における点列の傾向に関する考察 [36] と照らし合わせることも考えられる。

2.5 結言

本章ではポアソン過程において複数回のステップ変化が生じる状況を想定して、尤度理論と情報量規準に基づく複数変化点検出手法を設計した。尤度理論を用いることで、任意の変化点数における変化点とポアソン分布の強度パラメータを推定した。特に、複数ステップ変化に基づく統計モデルの尤度計算に動的計画法を用いることで、検討すべき変化点の組合せの絞り込みと計算結果の再利用が可能となり、高速なモデル推定を実現した。尤度理論のみでは推定ができない変化点数の異なるモデル間の比較については、AIC を用いることで少ない計算量でのモデル推定を可能とした。これにより、一連のデータから複数個の変化点とその個数、ならびにポアソン分布の強度パラメータを一意に推定できる。また従来の尤度に基づく変化点検出と異なり、管理状態におけるパラメータの値を用いずにモデルを推定できる。

設計した複数変化点検出手法の性能はモンテカルロシミュレーションを用いて確認した。数種類の異なるシナリオを想定したシミュレーションを通して、提案手法が複数個の変化点を変化点数とともに推定可能であることが示された。

さらに提案手法の特徴を活かして、状態変化追跡方法としてリアルタイムの工程監視に利用する方法について考察した。工程から逐次データを取得してモデル推定を繰り返すシミュレーションから、提案手法により工程の変化を管理図より早く検知する様子が確認された。

第3章 線形変化を想定した 複数変化点検出手法

3.1 緒言

変化点検出手法において変化パターンは重要な要素である。これは製造工程に生じる様々な異常とそれに伴う工程の変化に対応するためである。工程に生じる異常のうち、原材料のロットが切り替わり材料自体の品質が変化した場合や工程内の機械が故障した場合には工程から得られる統計量の傾向が突発的に変化する。Samuel et al. [16] に用いられるステップ変化モデルはこのような突発的な変化に対応した統計モデルである。他方で作業に伴う工具の摩耗や作業員の疲労が生じた場合、異常原因自体が徐々に変化していくことに対応して統計量の傾向も経時的に変化する。また変化は1回だけ生じるのではなく、異常が検知されるまでに複数回生じる場合もある。このような場合に1回の突発的な変化を想定したモデルで推定しても正確な変化点を得ることができず、かえって異常原因の究明作業をミスリーディングする可能性が生じる。したがって、経時的な変化を想定したモデルを用いて変化点を推定する必要がある。

経時的な変化を対象としたモデルとして Perry and Pignatiello [12] は線形変化モデルを提案した。これは異常発生により一定の変化率で状態が変化することを想定して、管理外れ状態における確率分布のパラメータに線形関数を当てはめ、変化が生じ始めた時点を変化点とするモデルである。線形変化モデルは経時的な変化を扱う代表的なモデルであり、Perry and Pignatiello [12] による正規分布に従うデータを対象とした手法の他にも、Perry and Pignatiello [13] によるポアソン分布を対象とした手法、Niaki and Khedmati [9] による幾何分布を対象とした手法などが提案されている。

また、Noorossana and Shadman [10] による単調変化モデルも経時的な変化を扱うモデルである。これは異常発生により特性値が従う確率分布のパラメータが単調増加（あるいは単調減少）すると仮定したモデルである。このモデルの利点は、パラメータが単調増加するという緩い条件であるため、変化率が一定ではない場合や複数回の変化が生じる場合を含む幅広い状況を扱えることである。ただし、管

管理外れ状態におけるパラメータの値を細かく合わせこむ手法であるため、管理外れ状態における工程の状態変化の詳細をデータから抽出することは難しい。また線形変化モデルと単調変化モデルはどちらも管理状態と管理外れ状態を峻別する変化点を推定するためのモデルである。したがって工程に複数回の変化が生じたとしても、最初の変化点以外の詳細は明らかにできない。

一連のデータにおいて複数回の変化が生じた場合を対象として、管理状態と管理外れ状態を峻別するのみならず、全ての変化を変化点として推定する手法が提案されている。例えば竹本と有蘭 [26] は尤度理論や情報理論を用いて複数回の変化を複数の変化点として推定する手法を提案している。Assareh et al. [3] はベイズ理論を用いて複数の変化点を推定する手法を提案している。これらの手法を用いれば、管理外れ状態における工程の変化をより正確に推定することが可能となる。ひいては、各々の変化点に対応する異常原因を探すことを通して、より一層の工程改善が実現できる。ただし、これらの手法は各変化点においてステップ変化が生じることを想定している。前述の通り、異常原因によっては経時的な変化が生じるため、これをステップ変化モデルに当てはめると誤った変化点を推定するリスクが高まる。

製造工程では複数の原因により持続的な変化が複数回発生する場合も想定される。同じ原因であったとしても、持続的に変化する要因が変化の限界に達する、あるいは変化がより一層顕著になる等で変化量が途中で変わる場合が存在する。この場合において持続的な変化の傾向が変わった各々の時点を推定することは工程変化の原因を推定する一助となる。

これを踏まえて、各々の変化点において線形変化が生じることを想定した複数変化点モデルを提案する。 $\bar{X}$ 管理図が扱うような正規分布に従う統計量の平均が異常発生によって変化する状況を対象に、線形変化による複数個の変化点を、尤度関数と情報理論を用いて推定する変化点検出手法を設計する。また提案手法の性能をモンテカルロシミュレーションを用いて確認する。

3.2 線形変化モデルの概要

異常発生後に工程平均が経時的に変化する状況を想定したモデルの例として、Perry and Pignatiello [12] による 1 回線形変化モデルを紹介する。

検査に用いるデータが正規分布に従う工程を想定する。工程からは各期ごとに n 個のサンプルを取得し、第 t 期の i 個目のサンプルを $x_{t,i}$ と表記する。管理状態におけるデータ $x_{t,i}$ の分布の平均 μ_0 と分散 σ_0^2 は既知である。正規分布の再生性より、第 t 期のサンプルの平均 $\bar{X}_t = \frac{1}{n} \sum_{i=1}^n x_{t,i}$ も正規分布に従う。 $\bar{X}_t$ の平均は個々のサンプルと同じく μ_0 であり、 $\bar{X}_t$ の分散 $\sigma_{\bar{X}}^2$ とサンプルの分散 σ_0^2 の間

には $\sigma_{\bar{X}}^2 = \frac{\sigma_0^2}{n}$ の関係が成り立つ．この性質を利用し，工程の管理には最後にデータを取得した第 T 期までのサンプル平均 $\mathbf{X} = (\bar{X}_1, \dots, \bar{X}_T)$ を用いる．

線形変化モデルでは，変化点 τ にて工程に異常が生じた後に，確率分布のパラメータが線形関数にしたがって増加（あるいは減少）し続けることを想定する．具体的に，第 t 期における工程平均 μ_t が

$$\mu_t = \begin{cases} \mu_0 & t \leq \tau \\ \mu_0 + \xi \times (t - \tau) & t > \tau \end{cases} \quad (3.1)$$

であるとする．ここで ξ は線形関数の傾きで，持続的な変化の変化量を定めるパラメータとなる．式 (3.1) で与えられる 1 回線形変化モデルでは，変化点 τ と傾き ξ を推定する．

Perry and Pignatiello [12] は 1 回線形変化モデルを尤度理論に基づいて推定する手法を提案した． $\bar{X}$ 管理図が異常を検知した時点を T とすれば，管理図が異常を検知するまでのデータに対する式 (3.1) を利用した対数尤度関数は

$$\begin{aligned} \log L(\tau, \xi | \mathbf{X}) = C - \frac{1}{2\sigma_{\bar{X}}^2} & \left[\sum_{t=1}^T \bar{X}_t^2 - 2\mu_0 \sum_{t=1}^T \bar{X}_t + T\mu_0^2 \right. \\ & \left. + \xi \times (T - \tau)(T - \tau + 1) \left(\mu_0 + \frac{1}{6}\xi \times (2T - 2\tau + 1) \right) - 2\xi \times \sum_{t=\tau+1}^T \bar{X}_t(t - \tau) \right] \end{aligned} \quad (3.2)$$

となる．ただし C は τ, ξ の値に寄らない定数項である．式 (3.2) の ξ についての偏微分より，変化点が τ の場合における ξ の最尤推定量は

$$\hat{\xi}(\tau) = \frac{6 \left(\sum_{t=\tau+1}^T \bar{X}_t(t - \tau) - \frac{1}{2}\mu_0(T - \tau)(T - \tau + 1) \right)}{(T - \tau)(T - \tau + 1)(2T - 2\tau + 1)} \quad (3.3)$$

であると導出される．式 (3.3) を式 (3.2) に代入することで，変化点の最尤推定量を

$$\hat{\tau} = \arg \max_{0 \leq \tau < T} \left[-\frac{1}{\sigma_{\bar{X}}^2} \left(\frac{(T - \tau)(T - \tau + 1)}{2} \hat{\xi}(\tau) \times \left(\mu_0 + \frac{2T - 2\tau + 1}{6} \hat{\xi}(\tau) \right) - \hat{\xi}(\tau) \sum_{t=\tau+1}^T \bar{X}_t(t - \tau) \right) \right] \quad (3.4)$$

にて推定できる．したがって式 (3.3) と (3.4) にて計算した $\hat{\tau}$ と $\hat{\xi}(\hat{\tau})$ が 1 回線形変化モデルにおける変化点と傾きの推定値となる．

3.3 複数線形変化点検出手法の設計

Perry and Pignatiello [12] による単一線形変化モデルを拡張し、複数回の線形変化が発生する場合の統計モデルを構築する。複数線形変化モデルでは経時的な変化の傾向が変わった時点の推定を目的として、線形変化の傾きが変わる時点を変化点と定義する。具体的に、 k 回目の線形変化が生じる直前の時点をも τ_{k-1} と表記する。特に τ_0 は管理状態 μ_0 を保った最後の時点とする。変化点 τ_{k-1} にて線形変化の傾きが変化点以前とは異なる値 ξ_k に変化したとすれば、 τ_{k-1} 以後の工程平均を

$$\nu_k(t) = \nu_{k-1}(\tau_{k-1}) + \xi_k \times (t - \tau_{k-1}) \quad (3.5)$$

と定義できる。ただし最初の変化点 τ_0 以前は管理状態であるから

$$\nu_0(t) = \mu_0 \quad (3.6)$$

である。式 (3.5) にて $\nu_k(\tau_{k-1}) = \nu_{k-1}(\tau_{k-1})$ となるとおり、平均値は変化点にて突発的な変化を生じないものとする。複数回の線形変化を想定した場合の統計モデルは各変化点に対応した傾き ξ_k を用いて

$$\begin{cases} \xi_1 & t = \tau_0 + 1, \dots, \tau_1 \\ \xi_2 & t = \tau_1 + 1, \dots, \tau_2 \\ \vdots & \\ \xi_k & t = \tau_{k-1} + 1, \dots, \tau_k \\ \vdots & \\ \xi_K & t = \tau_{K-1} + 1, \dots, \tau_K (= T) \end{cases} \quad (3.7)$$

と表記できる。ここで T はデータを取得した最後の時点である。この統計モデルで推定対象となるのは $\tau_K = T$ を除く変化点 $\boldsymbol{\tau}^{(K)} \equiv (\tau_0, \dots, \tau_{K-1})$ とそれに対応する傾き $\boldsymbol{\xi}^{(K)} \equiv (\xi_1, \dots, \xi_K)$ 、変化点数 K である。

データのサンプル平均 $\bar{X}_t$ が正規分布に従うことから、式 (3.7) の統計モデルを想定した場合の尤度関数は

$$\begin{aligned} L(\boldsymbol{\tau}^{(K)}, \boldsymbol{\xi}^{(K)} | \mathbf{X}, \mu_0, \sigma_{\bar{X}}^2, T) &= \prod_{t=1}^{\tau_0} \frac{1}{\sqrt{2\pi\sigma_{\bar{X}}}} e^{-\frac{1}{2\sigma_{\bar{X}}^2}(\bar{X}_t - \mu_0)^2} \\ &\times \prod_{k=1}^K \prod_{t=\tau_{k-1}+1}^{\tau_k} \frac{1}{\sqrt{2\pi\sigma_{\bar{X}}}} e^{-\frac{1}{2\sigma_{\bar{X}}^2}(\bar{X}_t - \nu_k(t))^2} \end{aligned} \quad (3.8)$$

である．また対数尤度関数は

$$\begin{aligned} & \log L(\boldsymbol{\tau}^{(K)}, \boldsymbol{\xi}^{(K)} \mid \mathbf{X}, \mu_0, \sigma_{\bar{X}}^2, T) \\ &= C + \frac{1}{2\sigma_{\bar{X}}^2} \left\{ -\mu_0^2 \tau_0 + \sum_{k=1}^K \sum_{t=\tau_{k-1}+1}^{\tau_k} \left\{ 2(\nu_{k-1}(\tau_{k-1}) - \mu_0) \bar{X}_t \right. \right. \\ & \quad \left. \left. - (\nu_{k-1}(\tau_{k-1}))^2 - \xi_k^2 \times (t - \tau_{k-1})^2 + 2\xi_k \times (t - \tau_{k-1})(\bar{X}_t - \nu_{k-1}(\tau_{k-1})) \right\} \right\} \end{aligned} \quad (3.9)$$

となる．ただし C はモデルによって変化しない定数項である．

変化点数が同じであれば，尤度理論に基づいて最適なモデルを推定可能である．しかしながら，複数線形変化を想定したモデルではパラメータの最尤推定量の計算が難しい．式 (3.5) が示すように， $\nu_k(t), \tau_{k-1} < t \leq \tau_k$ は ξ_k と $\nu_{k-1}(\tau_{k-1})$ で構成され，変化点 τ_{k-1} で前後の線形関数 $\nu_{k-1}(t), \nu_k(t)$ が連続である．したがって ξ_k の推定問題は $\nu_{k-1}(\tau_{k-1})$ の値の影響を受ける．さらに $\nu_{k-1}(\tau_{k-1})$ の値は $\nu_{k-2}(\tau_{k-2})$ および ξ_{k-1} の値に影響される．同様に ξ_k の推定値は $\nu_k(\tau_k)$ の値ならびに次の線形の傾き ξ_{k+1} の推定問題に影響を与える．このように， ξ_k の値は工程平均が $\nu_k(t)$ となる区間 τ_{k-1}, τ_k の情報だけでは一意に定まらず，他の変化点 $\tau_0, \dots, \tau_{k-2}, \tau_{k+1}, \dots, \tau_K$ の影響を受ける．変化点の組合せが得られても傾き $\boldsymbol{\xi}^{(K)}$ の最尤推定量の算出が難しいこと，変化点の組合せ $\boldsymbol{\tau}^{(K)}$ ごとに傾きを再度計算する必要があることから，対数尤度関数による最適なモデルの選定には膨大な計算量を必要とする．

計算量を削減し現実的な計算時間で処理を行うために，Bellman and Roth [4] による動的計画法を ξ_k の推定に援用する．式 (3.6) より， τ_0 における平均値 $\nu_1(\tau_0) = \nu_0(\tau_0)$ は定数で与えられる．その値を線形関数の切片とすれば，続く変化点 τ_1 に対して，二つの変化点間の傾き ξ_1 を推定できる．この推定値を用いれば， τ_1 までの尤度と τ_1 時点での線形関数の値 $\nu_1(\tau_1)$ が一意に定まる．この $\nu_1(\tau_1)$ を切片として次の変化点 τ_2 までの線形関数の傾き ξ_2 を推定できる．繰り返して計算することで，任意の τ_{k-1}, τ_k の区間における線形関数の傾き ξ_k を推定できる． ξ_k の推定にあたっては，全ての変化点 $\boldsymbol{\tau}^{(K)}$ ではなく， τ_k 以前の変化点 $\boldsymbol{\tau}^{(k)}$ だけが用いられる．上記の計算には動的計画法が適用できるため，変化点の組合せ $\boldsymbol{\tau}^{(K)}$ を全て検討する必要がなくなり，計算量を削減できる．

動的計画法に用いる小問題への分割のために，式 (3.9) の対数尤度関数から連続する変化点 τ_{k-1}, τ_k に挟まれた区間を抜き出した

$$\Lambda(\tau_{k-1}, \tau_k, \xi_k \mid \nu_{k-1}(\tau_{k-1}), \mathbf{X}, \mu_0) = \sum_{t=\tau_{k-1}+1}^{\tau_k} \left\{ 2(\nu_{k-1}(\tau_{k-1}) - \mu_0) \bar{X}_t \right.$$

$$-(\nu_{k-1}(\tau_{k-1}))^2 - \xi_k^2 \times (t - \tau_{k-1})^2 + 2\xi_k \times (t - \tau_{k-1})(\bar{X}_t - \nu_{k-1}(\tau_{k-1}))\} \quad (3.10)$$

を定義する．ここで $\nu_{k-1}(\tau_{k-1})$ が定数として与えられれば，傾きの最尤推定量 $\hat{\xi}_k$ は

$$\begin{aligned} & \frac{\partial}{\partial \xi_k} \Lambda(\tau_{k-1}, \tau_k, \xi_k \mid \nu_{k-1}(\tau_{k-1}), \mathbf{X}, \mu_0) \\ &= \sum_{t=\tau_{k-1}+1}^{\tau_k} \{-2\xi_k \times (t - \tau_{k-1})^2 + 2(t - \tau_{k-1})(\bar{X}_t - \nu_{k-1}(\tau_{k-1}))\} \\ &= 0 \end{aligned} \quad (3.11)$$

より

$$\hat{\xi}_k = \frac{-3\nu_{k-1}(\tau_{k-1})}{2(\tau_k - \tau_{k-1}) + 1} + \frac{6 \sum_{t=\tau_{k-1}+1}^{\tau_k} (t - \tau_k) \bar{X}_t}{(\tau_k - \tau_{k-1})(\tau_k - \tau_{k-1} + 1)(2(\tau_k - \tau_{k-1}) + 1)} \quad (3.12)$$

で計算できる．式 (3.10) に式 (3.12) を代入すれば Λ の最大値が得られる．したがってある時点 t までに k 回の変化が生じた場合の対数尤度関数の最大値は

$$D(k, t) = \max_{k \leq \tau_{k-1} < t} \left\{ D(k-1, \tau_{k-1}) + \Lambda(\tau_{k-1}, t, \hat{\xi}_k \mid \mathbf{X}, \mu_0, \nu_{k-1}(\tau_{k-1})) \right\} \quad (3.13)$$

で求まる．ただし式 (3.13) の境界条件は

$$D(0, \tau_0) \equiv -\mu_0^2 \tau_0 \quad (3.14)$$

である．式 (3.12) と (3.13) を用いることで，所与の変化点数 K のもとでの変化点の組合せの最尤推定量は

$$\hat{\tau}^{(K)} = \arg \max_{\tau^{(K)}} D(K, T) \quad (3.15)$$

で与えられる．そして $\hat{\tau}^{(K)}$ を用いて計算した傾き $\hat{\xi}^{(K)} = (\hat{\xi}_1, \dots, \hat{\xi}_K)$ が，変化回数が K における傾きの最尤推定値となる．

式 (3.13) を用いた尤度の計算においても動的計画法が利用できる．小問題 $D(k-1, \tau_{k-1})$ を計算すれば，変化点数と最後の時点の組合せ $k-1, \tau_{k-1}$ における最適な変化点の組合せ $\hat{\tau}^{(k-1)}$ と対応する傾き $\hat{\xi}^{(k-1)}$ が与えられる．これにより検討すべき変化点の組合せ $\tau^{(k)}$ を減らせるとともに，式 (3.10) と (3.12) の計算に必要な $\nu_{k-1}(\tau_{k-1})$ の値を一意に決定できる．

尤度理論に基づいてモデルを比較した場合、尤度は変化点数 K の多いモデルの方が高くなる。したがって最も変化点数の多い $K = T - 1$ のモデルが推定されてしまう。これを回避するため、AIC [1] を用いて尤度とモデルの複雑性のバランスが取れたモデルを選定する。AIC は

$$AIC(K) = -2 \log L(\boldsymbol{\tau}^{(K)}, \boldsymbol{\xi}^{(K)} \mid \mathbf{X}, \mu_0, \sigma_X^2, T) + 2B(K) \quad (3.16)$$

で定義される。ここで AIC のモデルのパラメータ数によるペナルティ $B(K)$ は、1 回の変化につき変化点 τ_k と傾き ξ_k の推定が行われることから

$$B(K) = 2K \quad (3.17)$$

で与えられる。AIC に含まれる対数尤度の計算に式 (3.13) を代用することで、変化点数は

$$\hat{K} = \arg \min_{0 \leq K < T} \left\{ -\frac{D(K, T)}{2\sigma_X^2} + 2K \right\} \quad (3.18)$$

で推定できる。

3.4 数値検証

本章では、提案した複数線形変化点検出手法の性能をモンテカルロシミュレーションにより確認する。具体的には、複数回の線形変化を想定し生成した乱数列に対して提案手法を適用し、変化点を推定する。比較対象として、Perry and Pignatiello [12] の 1 回線形変化点検出手法、Noorossana and Shadman [10] の単調変化点検出手法および竹本と有蘭 [25] の複数ステップ変化点検出手法による推定結果を用いる。

3.4.1 単調変化点モデルの概略

単調変化モデルは変化点以後にパラメータが単調増加する、すなわち第 t 期の平均 μ_t について

$$\mu_{\tau+1} > \mu_0 \quad (3.19)$$

$$\mu_t \geq \mu_{t-1}, \quad t = \tau + 2, \dots, T \quad (3.20)$$

が成り立つ状況を対象とする。また、式 (3.19) と (3.20) の両方の不等号を逆向きにすることで、パラメータが単調減少する状況を扱える。例えば複数回の線形変

化にて傾きが全て正（あるいは全て負）であれば，単調増加（単調減少）であるため単調変化モデルの想定する状況となる．

式 (3.19) と (3.20) を用いた場合の対数尤度関数は

$$\begin{aligned} \log L(\tau, \mu_{\tau+1}, \dots, \mu_T \mid \mathbf{X}, \mu_0) \\ = C - \frac{1}{2\sigma_{\bar{X}}} \left(\sum_{t=1}^{\tau} (\bar{X}_t - \mu_0)^2 + \sum_{t=\tau+1}^T (\bar{X}_t - \mu_t)^2 \right) \end{aligned} \quad (3.21)$$

である．ただし C は τ の値によらない定数項である．所定の τ に対して式 (3.21) の最大値を与える $\mu_t, t = \tau + 1, \dots, T$ を求める問題は

$$\begin{aligned} \text{minimize} \quad & \sum_{i=t+1}^T (\bar{X}_t - \mu_i)^2 \\ \text{subject to} \quad & \mu_i \leq \mu_{i+1} \text{ for } i = t + 1, \dots, T - 1 \end{aligned} \quad (3.22)$$

と定式化できる．式 (3.22) は isotonic 回帰問題へ帰着されるため，isotonic 回帰ソルバを用いて各期における平均の最尤値 $\hat{\mu}_t, t = \tau + 1, \dots, T$ を計算できる．平均の推定値 $\hat{\mu}_t$ を用いれば，変化点の最尤推定量を

$$\hat{\tau} = \arg \min_{0 \leq \tau < T} \left\{ \sum_{t=1}^{\tau} (\bar{X}_t - \mu_0)^2 + \sum_{t=\tau+1}^T (\bar{X}_t - \hat{\mu}_t)^2 \right\} \quad (3.23)$$

で求めることができる．

3.4.2 複数ステップ変化点検出手法の概要

統計量に従う正規分布について，分散 $\sigma_{\bar{X}}^2$ は一定のまま平均だけにステップ変化が生じることを想定する．具体的に， K 回の状態変化を想定した統計モデル

$$\left\{ \begin{array}{ll} \mu_0 & t = 1, \dots, \tau_0 \\ \mu_1 & t = \tau_0 + 1, \dots, \tau_1 \\ \vdots & \\ \mu_k & t = \tau_{k-1} + 1, \dots, \tau_k \\ \vdots & \\ \mu_K & t = \tau_{K-1} + 1, \dots, \tau_K (= T) \end{array} \right. \quad (3.24)$$

を定義する．ここで τ_k は $k - 1$ 回目の変化点， μ_k は k 回目の変化以降の工程平均である．特に， μ_0 は管理状態における工程平均， τ_K はサンプルを取得した最後の

時点である．式 (3.24) のうち推定量となるのは工程平均 $\boldsymbol{\mu}^{(K)} \equiv (\mu_1, \dots, \mu_K)$ ，変化点 $\boldsymbol{\tau}^{(K)} \equiv (\tau_0, \dots, \tau_{K-1})$ および変化回数 K である．

式 (3.24) の統計モデルを用いた場合，尤度関数は

$$L(\boldsymbol{\mu}^{(K)}, \boldsymbol{\tau}^{(K)} \mid \mathbf{X}, \sigma_{\bar{X}}^2) = (2\pi\sigma_{\bar{X}}^2)^{-\frac{T}{2}} e^{-\frac{1}{2\sigma_{\bar{X}}^2} \sum_{t=1}^{\tau_0} (\bar{X}_t - \mu_0)^2 - \frac{1}{2\sigma_{\bar{X}}^2} \sum_{k=1}^K \sum_{t=\tau_{k-1}+1}^{\tau_k} (\bar{X}_t - \mu_k)^2} \quad (3.25)$$

である．任意の二つの変化点間における工程平均の最尤推定量は

$$\hat{\mu}_k = \frac{1}{\tau_k - \tau_{k-1}} \sum_{t=\tau_{k-1}+1}^{\tau_k} \bar{X}_t \quad (3.26)$$

で計算できる．このように求めた工程平均 $\hat{\boldsymbol{\mu}}^{(K)} \equiv (\hat{\mu}_1, \dots, \hat{\mu}_K)$ を用いれば，変化回数を K 回と想定した場合の変化点の最尤推定量 $\hat{\boldsymbol{\tau}}^{(K)} = (\hat{\tau}_0, \dots, \hat{\tau}_{K-1})$ が

$$\hat{\boldsymbol{\tau}}^{(K)} = \arg \max_{\boldsymbol{\tau}^{(K)}} \left\{ \sum_{k=1}^K (\tau_k - \tau_{k-1}) (\hat{\mu}_k - \mu_0) \right\} \quad (3.27)$$

の関係より求められる．さらに AIC を用いることで，変化回数の推定量 $\hat{K}$ が

$$\hat{K} = \arg \min_K \left\{ -2 \log L(\hat{\boldsymbol{\mu}}^{(K)}, \hat{\boldsymbol{\tau}}^{(K)} \mid \mathbf{X}, \sigma_{\bar{X}}^2) + 2B(K) \right\} \quad (3.28)$$

の関係より求められる．ただし

$$B(K) = \begin{cases} 2K, & \text{if } 2K < T \\ T, & \text{otherwise} \end{cases} \quad (3.29)$$

である．

3.4.3 シミュレーション結果

シミュレーションの設定条件は以下の通りである．各サンプルは正規分布に従い，サンプルサイズは $n = 10$ とする．管理状態では，正規分布のパラメータは $\mu_0 = 0.0, \sigma_0^2 = 10.0$ とする．ここで $\sigma_{\bar{X}} = \frac{\sigma_0^2}{n} = 1.0$ である．したがって $\bar{X}$ 管理図における上側管理限界および下側管理限界はそれぞれ $UCL = 3.0, LCL = -3.0$ とする．複数回の線形変化の例として，単調増加の場合と非単調増加の 2 種類のデータを確認する．それぞれ $\bar{X}$ 管理図が異常を検知した時点までのデータに対して変化点検出手法を適用する．それに加えて単調増加かつ変化量が比較的大きな

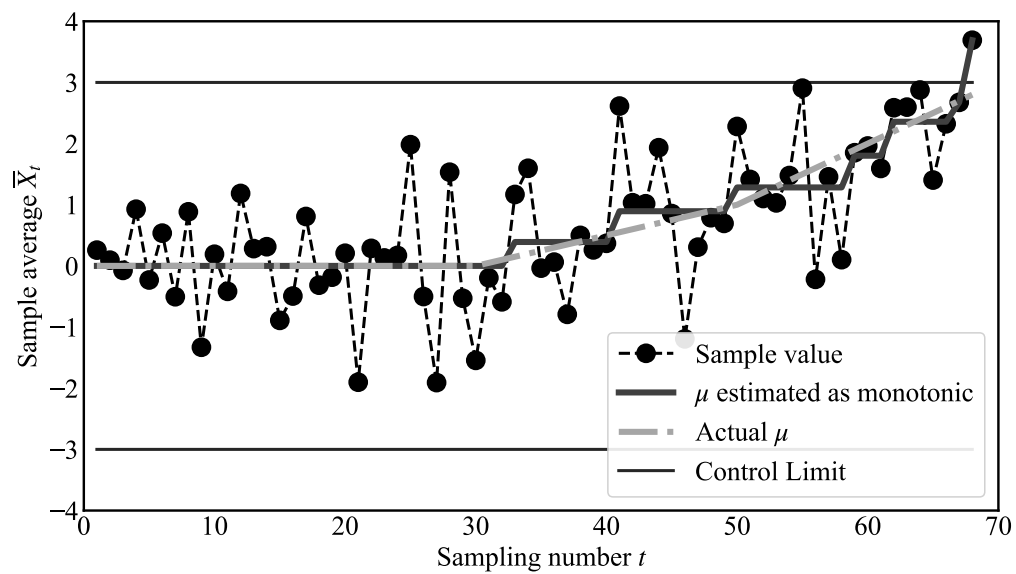


図 3.1: 式 (3.30) に従う乱数列に対する単調変化点検出手法による解析結果

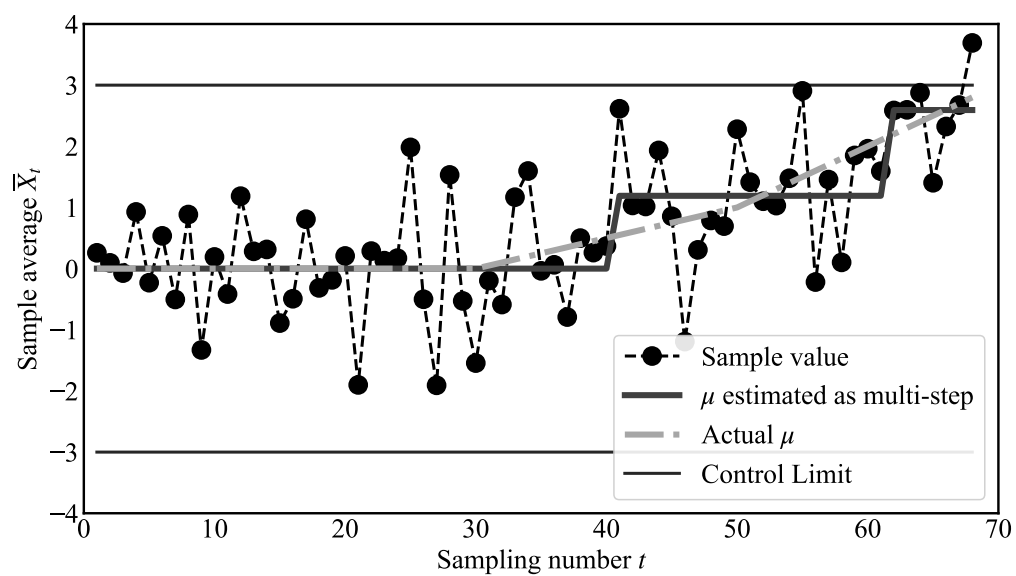


図 3.2: 式 (3.30) に従う乱数列に対する複数ステップ変化点検出手法による解析結果

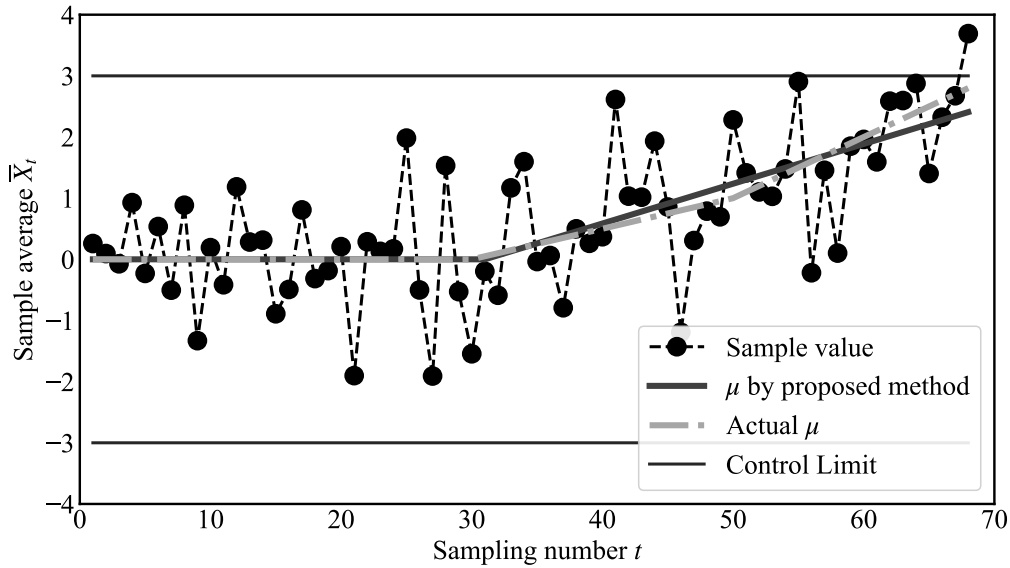


図 3.3: 式 (3.30) に従う乱数列に対する提案手法による解析結果

事例を確認するために、管理図を併用せずに所定の期数まで乱数を生成した事例についても確認した。

まず単調増加な複数線形変化が生じる例として

$$\begin{cases} \tau_0 = 30, & \xi_1 = 0.05 \\ \tau_1 = 50, & \xi_2 = 0.1 \end{cases} \quad (3.30)$$

というシナリオに従う乱数列を生成し、それぞれの変化点検出手法を適用した。

図 3.1 に式 (3.30) に従うデータを単調変化点検出手法で解析した結果を示す。単調変化点検出手法を用いる場合には、パラメータに生じた変化が単調増加か単調減少のどちらであるかを定める必要がある。今回の例では $\bar{X}$ 管理図での異常検知においてデータが上側管理限界を超えたため、パラメータが単調増加したと仮定して手法を適用した。単調変化点検出手法では、変化点を $\hat{\tau} = 32$ と推定した。実際の最初の変化点は $\tau_0 = 30$ であるため、単調変化点検出法では管理状態と管理外れ状態を峻別できている。図 3.2 に式 (3.30) に従うデータを複数ステップ変化点検出手法で解析した結果を示す。複数ステップ変化点検出手法では、2 個の変化点を推定した。1 個目の変化点の推定値 $\hat{\tau}_0 = 40$ は、実際の二つの変化点 $\tau_0 = 30, \tau_1 = 50$ のちょうど中間を示している。また 2 個目の変化点の推定値 $\hat{\tau}_1 = 61$ は、最後の変化点 τ_1 から 10 期以上離れている。このように、ステップ変化に基づく変化点検出手法では、持続的な変化が生じた時点を的確に推定することは難しい。

図 3.3 に式 (3.30) に従うデータを提案手法で解析した結果を示す。提案手法は 1 個の変化点 $\hat{\tau}_0 = 31$ を推定した。提案手法においても、管理状態と管理外れ状

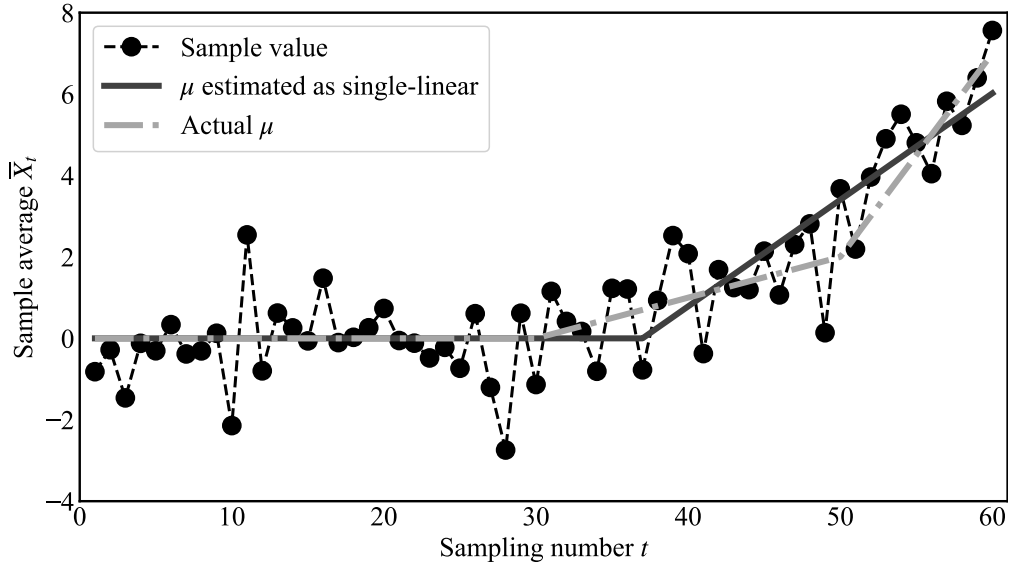


図 3.4: 式 (3.31) に従う乱数列に対する 1 回線形変化点検出手法による解析結果

態を峻別する時点を推定できている．なお，1 回線形変化点検出手法においても図 3.3 と同じ $\hat{\tau} = 31$ が推定された．

提案手法は複数個の変化点を推定できるよう設計されているため，本来であれば 2 個目の変化点も推定されるはずであるが，今回の推定結果ではそのようになっていない．その原因はデータのばらつきに対して変化量が微小であるからと考えられる．AIC では 2 個のモデルで尤度の差がほとんどない場合には，モデルに含まれるパラメータ数がより少ないモデルが選ばれる．式 (3.30) の事例では 2 個目の変化点における傾きの変化量は $\xi_2 - \xi_1 = 0.05$ と，サンプル平均の標準偏差 $\sigma_{\bar{X}}$ に対して 5%，データ自体の標準偏差に対しては 1.6% と小さな値である．このため，変化点数 $K = 1$ のモデルでも $K = 2$ のモデルと尤度に大きな差が生じなかったため，変化点が過少に推定されたと考えられる．

上記の考察を踏まえて，線形変化の傾きがより大きな場合の例として

$$\begin{cases} \tau_0 = 30, & \xi_1 = 0.1 \\ \tau_1 = 50, & \xi_2 = 0.5 \end{cases} \quad \text{and} \quad T = 60 \quad (3.31)$$

のシナリオを用意し，これに従う乱数列に対して変化点検出手法を適用した．この例では $\bar{X}$ 管理図は用いず，データは $T = 60$ 期まで生成した．

図 3.4 に式 (3.31) に従うデータを 1 回線形変化点検出手法で解析した結果を示す．1 回線形変化点検出手法で推定された変化点は $\hat{\tau} = 37$ であり，1 個目の変化点 $\tau_0 = 30$ および 2 個目の変化点 $\tau_1 = 50$ とは少し差がある．図 3.5 に式 (3.31) に従うデータを単調変化点検出手法で解析した結果を示す．パラメータの単調増

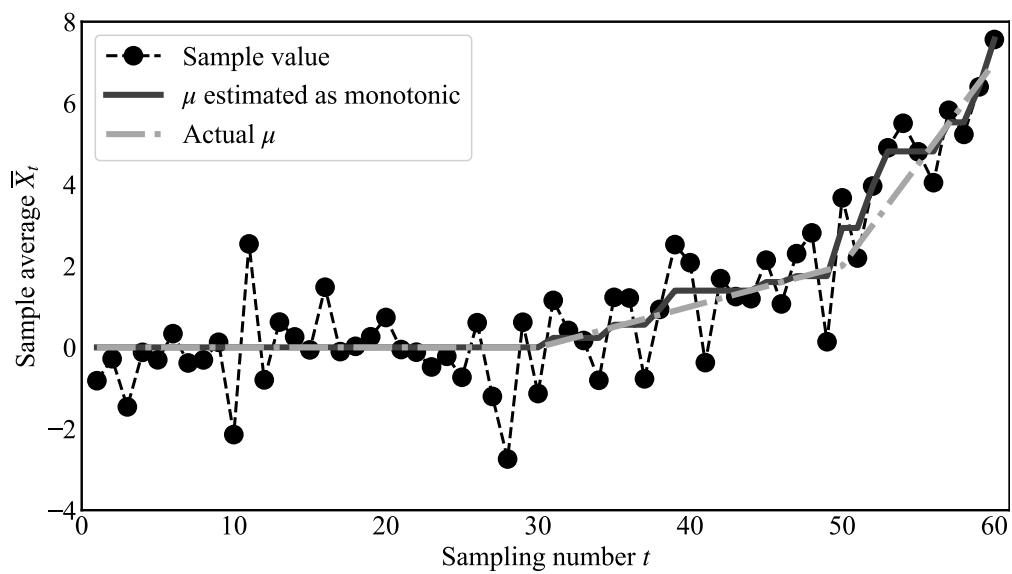


図 3.5: 式 (3.31) に従う乱数列に対する単調変化点検出手法による解析結果

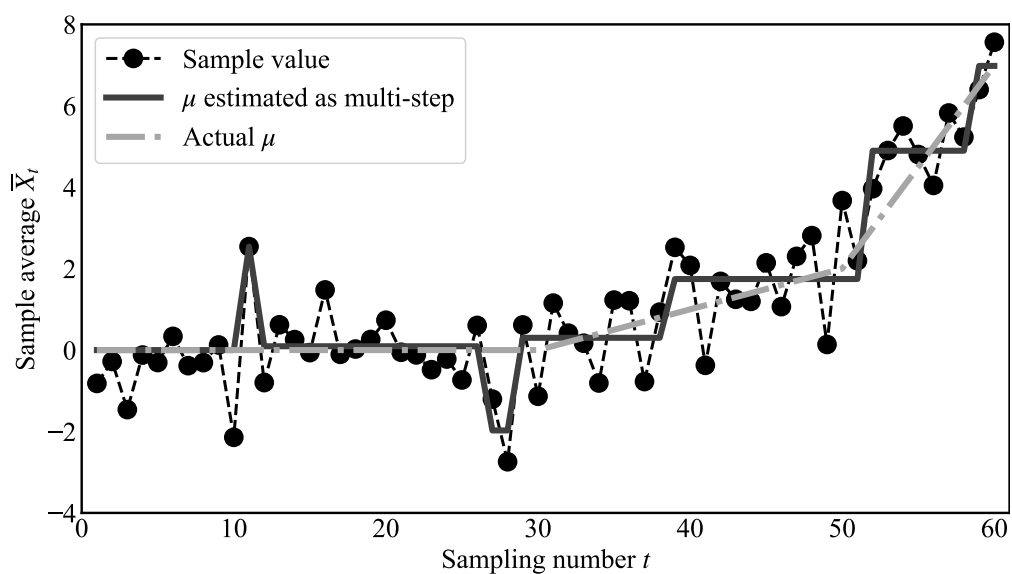


図 3.6: 式 (3.31) に従う乱数列に対する複数ステップ変化点検出手法による解析結果

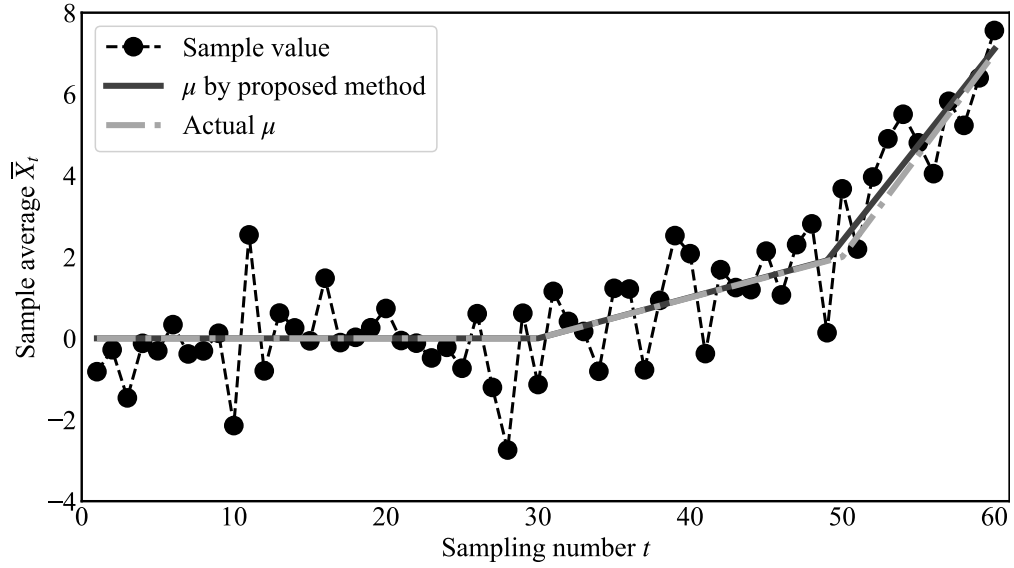


図 3.7: 式 (3.31) に従う乱数列に対する提案手法による解析結果

加を想定した場合における変化点の推定値は $\hat{\tau} = 30$ であり、実際に管理外れ状態に移行した時点である 1 個目の変化点 $\tau_0 = 30$ と一致している．その一方で、単調変化モデルはパラメータを多数のステップ変化に分割して推定するため、図 3.5 の解析結果から 2 個目以降の変化点を推定することは困難である．実際にでは変化点の推定値 $\hat{\tau}$ 以降に 10 回以上パラメータの値が変化している．これらの単調モデル上でのパラメータ変化が実際の変化点であるか否かを判別することは難しい．図 3.6 に式 (3.31) に従うデータを複数ステップ変化点検出手法で解析した結果を示す．複数ステップ変化点では、 $t = 11$ および $t = 28$ の二つの外れ値も変化点として推定しているため、全部で 7 個の変化点を推定した． $\hat{\tau}_4 = 38$ は、実際の変化点 $\tau_0 = 30, \tau_1 = 50$ のほぼ中間の値となっているなど、変化点を正確に推定できているとは言い難い．

図 3.7 に式 (3.31) に従うデータを提案手法で解析した結果を示す．提案手法では変化点数を $\hat{K} = 2$ 、変化点を $\hat{\tau}_0 = 30, \hat{\tau}_1 = 49$ と推定した．1 個目の変化点は実際の値 $\tau_0 = 30$ に一致し、2 個目の変化点は実際の値 $\tau_1 = 50$ と 1 期差と、どちらも高い精度で推定できている．また傾きの推定値は $\hat{\xi}_0 = 0.101, \hat{\xi}_1 = 0.472$ であり、2 個目の変化点で傾きが増大したと推定されている．この結果から、傾きの変化量がある程度大きい場合においてはパラメータが単調増加であっても線形変化に起因する複数個の変化点をその個数とともに推定できることが確認できる．

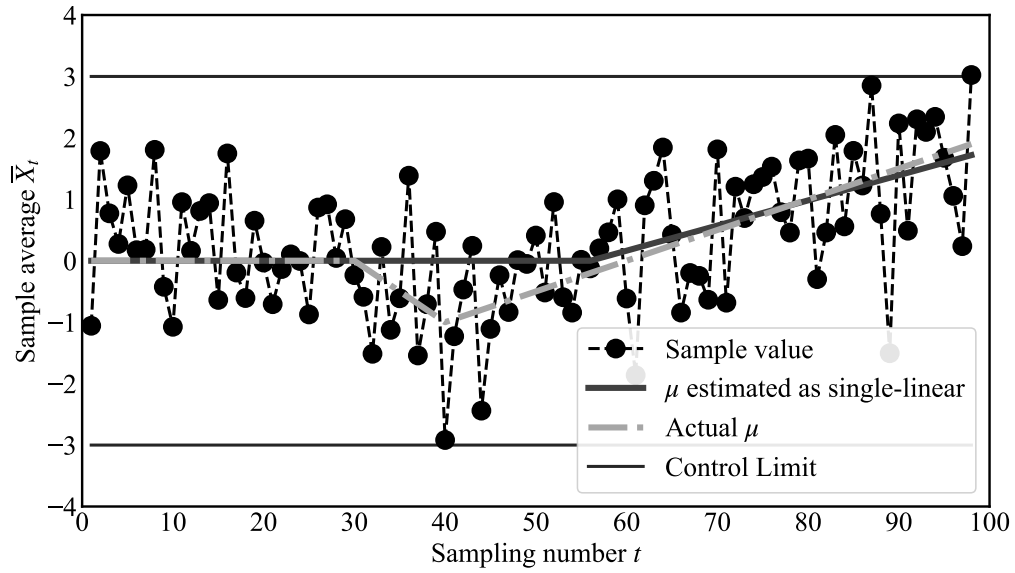


図 3.8: 式 (3.32) に従う乱数列に対する 1 回線形変化点検出手法による解析結果

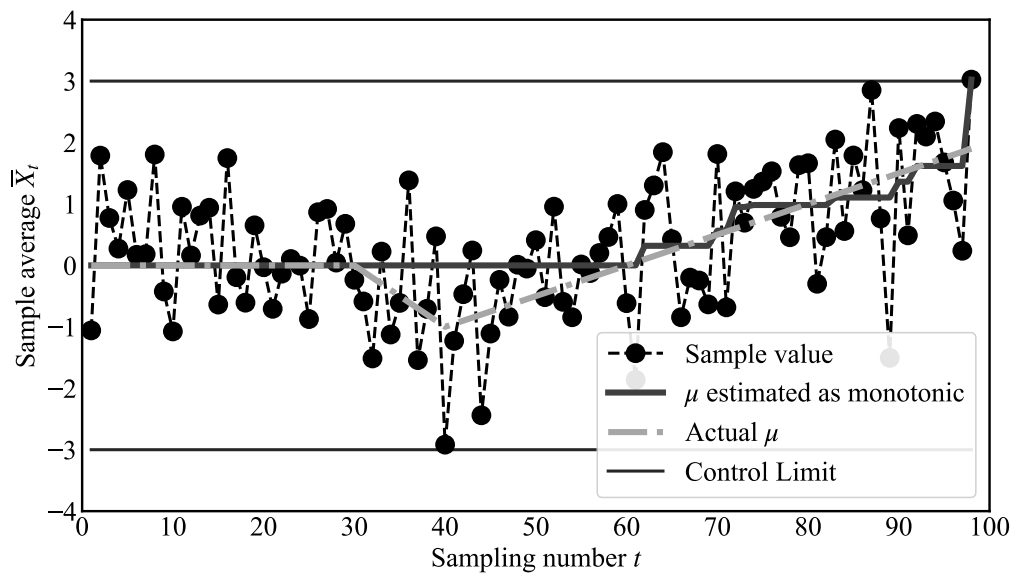


図 3.9: 式 (3.32) に従う乱数列に対する単調変化点検出手法による解析結果

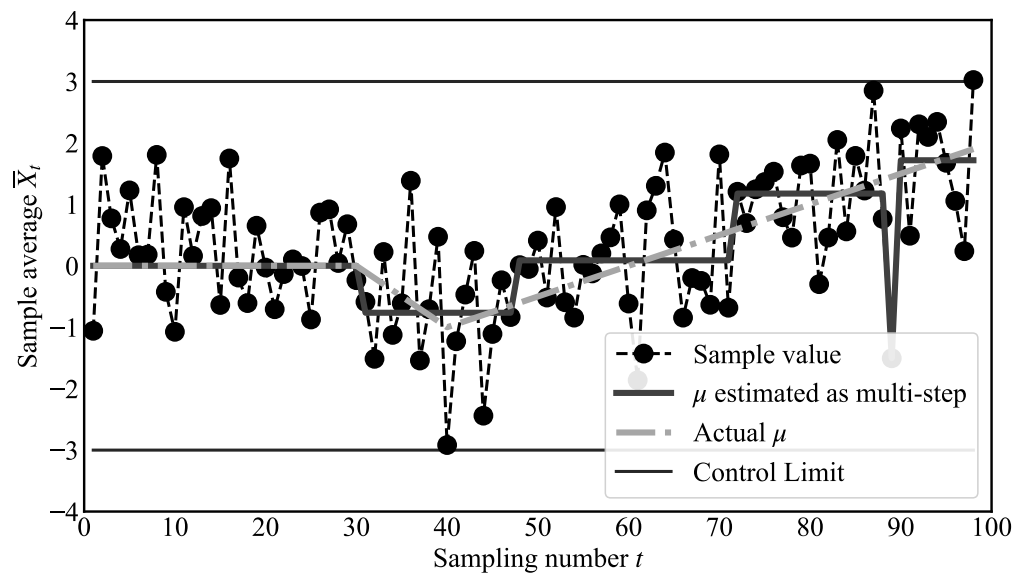


図 3.10: 式 (3.32) に従う乱数列に対する複数ステップ変化点検出手法による解析結果

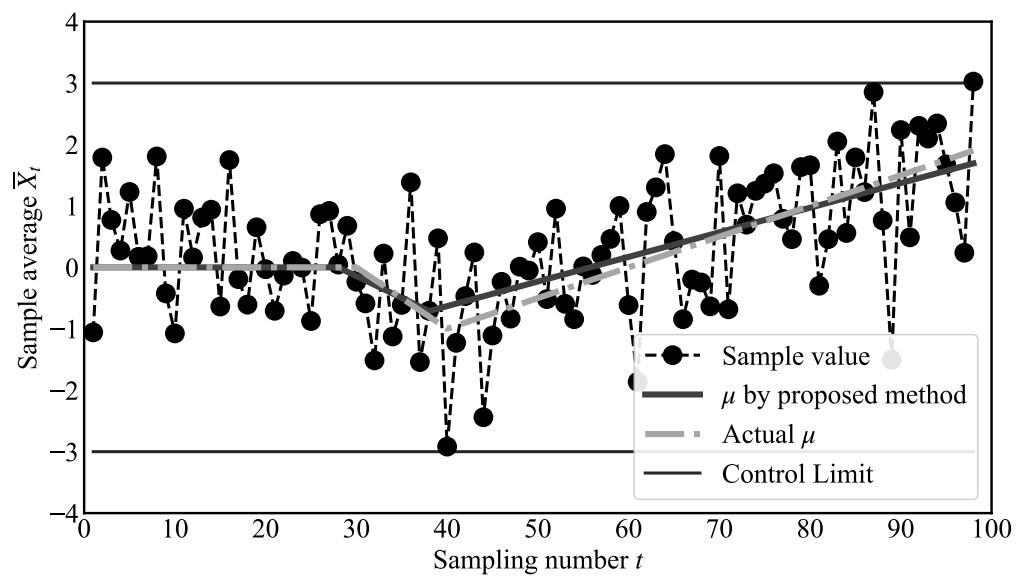


図 3.11: 式 (3.32) に従う乱数列に対する提案手法による解析結果

最後に、パラメータが単調増加や単調減少ではない例として

$$\begin{cases} \tau_0 = 30, & \xi_1 = -0.1 \\ \tau_1 = 40, & \xi_2 = 0.05 \end{cases} \quad (3.32)$$

というシナリオに従う乱数列を生成し、それぞれの変化点検出手法を適用した。

図 3.8 に式 (3.32) に従うデータを 1 回線形変化点検出手法で解析した結果を示す。この手法では $\hat{\tau} = 56$ に変化点が推定されている。図 3.9 に式 (3.32) に従うデータを単調変化点検出手法で解析した結果を示す。ここで $\bar{X}$ 管理図はデータが上側管理限界を超えたことで異常を検知していることから、単調増加を想定して単調変化点検出手法を適用した。その結果変化点は $\hat{\tau} = 61$ と推定された。今回のシナリオにおいて $t = 61$ は工程平均が管理状態の値 μ_0 を初めて超えた時点である。したがって、単調変化点検出手法ではパラメータが管理状態より大きくなる時点を正確に推定していると言える。1 回線形変化点検出手法についてもその推定結果が $t = 61$ に近いことから、パラメータが管理状態の値より大きくなった時点を推定していると考えることができよう。しかしながら、実際の変化点は $\tau_0 = 30$ および $\tau_1 = 40$ であり、前述の 2 手法による推定結果は実際の変化点から大きく外れている。図 3.10 に式 (3.32) に従うデータを複数ステップ変化点検出手法で解析した結果を示す。最初の変化点 $\hat{\tau}_0 = 30$ は実際の変化点 $\tau_0 = 30$ と一致している。しかしながら、 $\tau_1 = 40$ 以降のパラメータの増加を複数回のステップ変化に当てはめているため、 $\hat{\tau}_2 = 71$ のような実際の変化点から大きく離れた時点を変化点として推定している。

図 3.11 に式 (3.32) に従うデータを提案手法で解析した結果を示す。提案手法では変化点数を $\hat{K} = 2$ 、変化点を $\hat{\tau}_0 = 28, \hat{\tau}_1 = 38$ と推定している。変化点の推定結果についてはどちらも実際の変化点 $\tau_0 = 30, \tau_1 = 40$ に近い値である。このように提案手法はパラメータ変化が単調増加あるいは単調減少でない場合であっても変化点を推定可能である。

3.5 結言

本章では異常発生によって工程平均が経時的に変化し、さらにその変化量も途中で増減しうる状況を想定して、複数線形変化点検出手法を設計した。具体的には 1 回の線形変化モデルを拡張する形で、複数回の工程変化が生じ各々の変化点にて線形関数の傾きが変わることを想定した統計モデルを定義した。そのうえで、データが正規分布に従う場合の工程平均の変化を対象に、変化点とそれに対応する変化の傾き、ならびに変化点数を尤度理論と情報量規準に基づいて推定する方

法を導出した。特に変化の傾きと尤度の推定については、前方から順に動的計画法を適用することで計算量の削減を図った。

設計した複数線形変化点検出手法の性能はモンテカルロシミュレーションを用いて確認した。その結果、単調増加（単調減少）ではない場合や傾きの変化量がある程度大きな場合においては複数個の変化点を変化点数とともに推定可能であるなど、従来手法よりも正確に変化点を推定できることが確認された。その一方で、同一方向かつ傾きの変化量が少ない場合には変化点数の推定を過少推定する傾向も見られた。その場合においても管理状態と管理外れ状態を峻別可能であるが、変化量が小さい場合における正確な変化点数推定は今後の課題である。

第4章 Hoeffdingの確率不等式の性能向上

4.1 緒言

確率不等式は確率変数の期待値や分散，定義域などの限定された情報から生起確率の上界（あるいは下界）を評価する手法であり，意思決定の対象が従う分布形が規定されない場合の意思決定に用いられる．確率不等式の一つとして Hoeffding の確率不等式 [7] が存在する．これは期待値と分散が既知である定義域が上に有界な複数個の確率変数について，その算術平均の上側確率の上界を計算する方法である．

確率不等式により与えられる上側確率の上界は，実際の上側確率（真値）に対して必ず余裕のある値である．ここで真値に対してよりタイトな上側確率の上界を与える確率不等式が実現すれば，より効率的な意思決定が実現できる．この観点から，上側確率の上界値をよりタイトに与える確率不等式が検討されてきた．有蘭ら [22] は各々の確率変数が同一でない母集団から得られる場合を想定し，定義域の上限が各々の確率変数で異なる場合において Hoeffding の確率不等式よりもタイトな上側確率の上界値を与える手法を導出した．また光岡ら [34] は確率変数の期待値，分散，定義域の上限および評価対象となる確率変数の平均値が一定の条件を満足する場合において，Hoeffding の確率不等式よりもタイトな上側確率の上界値を与える確率不等式を導出した．これらの手法は Hoeffding の確率不等式よりタイトな上側確率の上界を与えるものの，限定的な条件下でしか適用できない．

本章では Hoeffding の確率不等式の性能改善を扱う．具体的に Hoeffding の確率不等式の導出過程を踏まえたうえで，一部の計算を置き換えることでよりタイトな上側確率の上界を与える確率不等式を導出する．特に Hoeffding の確率不等式と全く同じ条件下で利用できるよう，確率変数の定義域の上界および各々の確率変数の期待値と分散が既知であることのみを提案手法の適用条件とする．まずは Hoeffding の確率不等式の導出過程を精査することで，よりタイトな上側確率の上界を与える確率不等式が存在する可能性を示す．次に確率変数が 2 個の場合に限定して，具体的な計算手順を導出する．これを確率変数が任意の個数の場合に拡

張することで、Hoeffding の確率不等式と同じ条件下でよりタイトな上側確率の上界値を与える確率不等式を導出する．また提案手法の Hoeffding の確率不等式に対する優位性を、数値計算を行って確認する．

4.2 Hoeffding の確率不等式

本節では Hoeffding の確率不等式の導出手順を紹介する．以降は $f_X(X)$ で確率変数 X の確率密度関数（または確率質量関数）， $E_X[g(X)]$ で確率変数 X に対する関数 $g(X)$ の期待値， $V_X[g(X)]$ で確率変数 X に対する関数 $g(X)$ の分散を表す．

独立した確率変数 $Y_i, i = 1, \dots, n$ を想定する． Y_i の分布系は未知であるが，その定義域は上に有界で，各々の確率変数の平均 $\mu_i, i = 1, \dots, n$ と分散 $\sigma_i^2, i = 1, \dots, n$ は既知であるとする．ここで各々の確率変数の平均 μ_i ，分散 σ_i^2 および分布形が同一である必要はない．Hoeffding の確率不等式では，確率変数の算術平均 $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ がある値 y を超える確率 $\Pr\{\bar{Y} \geq y\}$ の上界を算出する．

ここで $Z_i = Y_i - \mu_i$ ， $\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i$ および $z = y - \frac{1}{n} \sum_{i=1}^n \mu_i$ を定義する．確率変数 Z_i の定義域も上に有界であり， $\delta \geq Z_i, \forall i$ を満足する δ が既知であるとする．確率変数 Z_i は平均 $E_{Z_i}[Z_i] = 0, \forall i$ ，分散 $V_{Z_i}[Z_i] = \sigma_i^2, i = 1, \dots, n$ である． $\Pr\{\bar{Y} \geq y\} = \Pr\{\bar{Z} \geq z\}$ であるから，確率変数 Z_i を用いて上側確率の上界値を計算可能である．計算式を簡略化するために，確率変数 Z_i を用いて $\Pr\{\bar{Z} \geq z\}$ の上界を求めることにする．

確率変数 $\bar{Z}$ とネイピア数 e ，非負の実数 θ を用いて新たな確率変数 $e^{\bar{Z}\theta}$ を定義する．確率変数の値域は $e^{\bar{Z}\theta} > 0$ である．ここで Markov の確率不等式 [38] を援用すれば

$$\Pr\{\bar{Z} \geq z\} \leq e^{-z\theta} E_{\bar{Z}} \left[e^{\bar{Z}\theta} \right] \quad (4.1)$$

の関係を得る．これは

$$\begin{aligned} E_{\bar{Z}} \left[e^{\bar{Z}\theta} \right] &= \int_{-\infty}^{\delta} e^{\bar{Z}\theta} f_{\bar{Z}}(\bar{Z}) d\bar{Z} \\ &\geq \int_z^{\delta} e^{\bar{Z}\theta} f_{\bar{Z}}(\bar{Z}) d\bar{Z} \\ &\geq \int_z^{\delta} e^{z\theta} f_{\bar{Z}}(\bar{Z}) d\bar{Z} \\ &= e^{z\theta} \int_z^{\delta} f_{\bar{Z}}(\bar{Z}) d\bar{Z} \end{aligned}$$

$$=e^{z\theta} \Pr\{\bar{Z} > z\} \quad (4.2)$$

より導出されている。

ここで式 (4.1) 右辺の $E_{\bar{Z}}[e^{\bar{Z}\theta}]$ を直接計算する代わりに, $E_{\bar{Z}}[e^{\bar{Z}\theta}]$ を上から抑え込む二次関数を用いて評価する. $Z_i \leq \delta$ の範囲において $e^{Z_i\theta}$ 以上の値となる二次関数

$$\phi_i(Z_i) = u_i Z_i^2 + v_i Z_i + w_i \quad (4.3)$$

を定義する. $Z_i \leq \delta$ にて $e^{Z_i\theta} \leq \phi_i(Z_i)$ を満足するために,

条件 1. $Z_i = \delta$ で $e^{Z_i\theta}$ と $\phi_i(Z_i)$ が交差する.

条件 2. $Z_i = -\frac{\sigma_i^2}{\delta}$ で $e^{Z_i\theta}$ と $\phi_i(Z_i)$ が接する.

という二つの十分条件を与える. 条件 2. については, 両関数が接する点は $Z_i \leq 0$ であればよい. ただし, 任意の接点を想定して $\phi_i(Z_i)$ を導出した場合, $E_{Z_i}[\phi_i(Z_i)]$ は接点が $Z_i = -\frac{\sigma_i^2}{\delta}$ の条件で最小化される. これを踏まえて, 条件 2. では接点として $Z_i = -\frac{\sigma_i^2}{\delta}$ が用いられる. 上記の条件を両方とも満足するならば, 全ての $\phi_i(Z_i)$ において

$$\phi_i(\delta) = e^{\delta\theta} \quad (4.4)$$

$$\phi_i\left(-\frac{\sigma_i^2}{\delta}\right) = e^{-\frac{\sigma_i^2\theta}{\delta}} \quad (4.5)$$

$$\frac{d}{dZ_i}\phi_i\left(-\frac{\sigma_i^2}{\delta}\right) = \theta e^{-\frac{\sigma_i^2\theta}{\delta}} \quad (4.6)$$

の三つの等式が成り立つ. 式 (4.4)~(4.6) を連立して解けば, 式 (4.3) の係数は

$$u_i = \frac{\delta^2}{(\delta^2 + \sigma_i^2)^2} \left(e^{\delta\theta} - e^{-\frac{\sigma_i^2\theta}{\delta}} \right) - \frac{\delta}{\delta^2 + \sigma_i^2} \theta e^{-\frac{\sigma_i^2\theta}{\delta}} \quad (4.7)$$

$$v_i = \frac{2\delta\sigma_i^2}{(\delta^2 + \sigma_i^2)^2} \left(e^{\delta\theta} - e^{-\frac{\sigma_i^2\theta}{\delta}} \right) + \frac{\delta^2 - \sigma_i^2}{\delta^2 + \sigma_i^2} \theta e^{-\frac{\sigma_i^2\theta}{\delta}} \quad (4.8)$$

$$w_i = \frac{(\sigma_i^2)^2}{(\delta^2 + \sigma_i^2)^2} e^{\delta\theta} + \left\{ \frac{\delta^2(\delta^2 + 2\sigma_i^2)}{(\delta^2 + \sigma_i^2)^2} + \frac{\delta\sigma_i^2}{\delta^2 + \sigma_i^2} \theta \right\} e^{-\frac{\sigma_i^2\theta}{\delta}} \quad (4.9)$$

となる.

式 (4.1) に含まれる $e^{\bar{Z}\theta}$ には

$$\begin{aligned} e^{\bar{Z}\theta} &= e^{\sum_{i=1}^n \frac{Z_i}{n} \theta} \\ &= \prod_{i=1}^n e^{\frac{Z_i}{n} \theta} \end{aligned} \quad (4.10)$$

が成り立つ．これは $e^{\frac{Z_i}{n}\theta}$ の積の値が， Z_i の組合せではなく，その平均値 $\bar{Z}$ にのみ依存することを示している．式 (4.10) に加え，独立した確率変数の積の期待値は期待値の積に等しいことを利用すれば

$$E_{\bar{Z}} \left[e^{\bar{Z}\theta} \right] = \prod_{i=1}^n E_{Z_i} \left[e^{\left(\frac{Z_i}{n}\right)\theta} \right] \quad (4.11)$$

が成り立つ． $e^{\frac{Z_i}{n}\theta} \leq \phi_i \left(\frac{Z_i}{n} \right)$ の関係を利用したうえで，

$$\begin{aligned} E_{Z_i} \left[e^{\frac{Z_i}{n}\theta} \right] &\leq E_{Z_i} \left[\phi_i \left(\frac{Z_i}{n} \right) \right] \\ &= \frac{\delta^2}{\delta^2 + \sigma_i^2} e^{-\frac{\sigma_i^2}{\delta} \frac{\theta}{n}} + \frac{\sigma_i^2}{\delta^2 + \sigma_i^2} e^{\delta \frac{\theta}{n}} \end{aligned} \quad (4.12)$$

の関係が導かれる．したがって

$$\begin{aligned} \prod_{i=1}^n E_{Z_i} \left[e^{\left(\frac{Z_i}{n}\right)\theta} \right] &\leq \prod_{i=1}^n E_{Z_i} \left[\phi_i \left(\frac{Z_i}{n} \right) \right] \\ &= \prod_{i=1}^n \left\{ \frac{\delta^2}{\delta^2 + \sigma_i^2} e^{-\frac{\sigma_i^2}{\delta} \frac{\theta}{n}} + \frac{\sigma_i^2}{\delta^2 + \sigma_i^2} e^{\delta \frac{\theta}{n}} \right\} \end{aligned} \quad (4.13)$$

である．さらに

$$\sigma^2 \equiv \frac{1}{n} \sum_{i=1}^n \sigma_i^2 \quad (4.14)$$

を定義し，式 (4.13) に Jensen の不等式 [38] を適用することで

$$\prod_{i=1}^n \left\{ \frac{\delta^2}{\delta^2 + \sigma_i^2} e^{-\frac{\sigma_i^2}{\delta} \frac{\theta}{n}} + \frac{\sigma_i^2}{\delta^2 + \sigma_i^2} e^{\delta \frac{\theta}{n}} \right\} \leq \left\{ \frac{\delta^2}{\delta^2 + \sigma^2} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} + \frac{\sigma^2}{\delta^2 + \sigma^2} e^{\delta \frac{\theta}{n}} \right\}^n \quad (4.15)$$

が導かれる．

式 (4.1), (4.11), (4.13) および (4.15) より

$$\left\{ \frac{\delta^2}{\delta^2 + \sigma^2} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} + \frac{\sigma^2}{\delta^2 + \sigma^2} e^{\delta \frac{\theta}{n}} \right\}^n \geq e^{z\theta} \Pr\{\bar{Z} > z\} \quad (4.16)$$

が成立する．式 (4.16) を変形することで Hoeffding の不等式

$$\Pr\{\bar{Z} \geq z\} \leq \frac{1}{e^{z\theta}} \left\{ \frac{\delta^2}{\delta^2 + \sigma^2} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} + \frac{\sigma^2}{\delta^2 + \sigma^2} e^{\delta \frac{\theta}{n}} \right\}^n \quad (4.17)$$

が導出される．式 (4.17) の右辺は上側確率 $\Pr\{\bar{Z} \geq z\}$ の上界値を意味するとともに， θ の関数となっている．ここで上界値を最小にする θ の値は解析的に

$$\theta = \frac{n\delta}{\delta^2 + \sigma^2} \log \left(\frac{1 + \frac{z\delta}{\sigma^2}}{1 - \frac{z}{\delta}} \right) \quad (4.18)$$

と求まる．式 (4.18) は z に対して単調増加であり， $z = 0$ において $\theta = 0$ となる． θ は非負の実数と定義したことから，Hoeffding の確率不等式は $z \geq 0$ に対して適用可能である．

4.3 確率不等式の性能改善に向けた考察

式 (4.17) 右辺の分子にあたる式 (4.15) 右辺は，二次関数 $\phi_i(Z_i)$ と式 (4.14) で定義される σ^2 を利用して導出された．一方で式 (4.16) の右辺，特に $e^{z\theta}$ は式 (4.2) の通り $e^{\frac{Z_i}{n}\theta}$ を利用して導出されており， $\phi_i(Z_i)$ は全く関与していない．ここで分母においても二次関数より導出した値を用いれば， $e^{z\theta}$ を用いるよりもタイトな上側確率の上界を得られる可能性がある．

式 (4.16) の左辺は σ_i^2 に代えて σ^2 が用いられていること，確率不等式の導出過程では $\frac{Z_i}{n}$ が頻出することを鑑みて， $Z_i \leq \delta$ の範囲で $e^{\frac{Z_i}{n}\theta}$ を上から抑える二次関数

$$\eta(Z_i) = \alpha Z_i^2 + \beta Z_i + \gamma \quad (4.19)$$

$$\alpha = \frac{\delta^2}{(\delta^2 + \sigma^2)^2} \left(e^{\delta \frac{\theta}{n}} - e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} \right) - \frac{\delta}{\delta^2 + \sigma^2} \frac{\theta}{n} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} \quad (4.20)$$

$$\beta = \frac{2\delta\sigma^2}{(\delta^2 + \sigma^2)^2} \left(e^{\delta \frac{\theta}{n}} - e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} \right) + \frac{\delta^2 - \sigma^2}{\delta^2 + \sigma^2} \frac{\theta}{n} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} \quad (4.21)$$

$$\gamma = \frac{(\sigma^2)^2}{(\delta^2 + \sigma^2)^2} e^{\delta \frac{\theta}{n}} + \left\{ \frac{\delta^2(\delta^2 + 2\sigma^2)}{(\delta^2 + \sigma^2)^2} + \frac{\delta\sigma^2}{\delta^2 + \sigma^2} \frac{\theta}{n} \right\} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} \quad (4.22)$$

を定義する．

この二次関数の係数を用いれば，式 (4.16) の左辺を

$$\left\{ \frac{\delta^2}{\delta^2 + \sigma^2} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} + \frac{\sigma^2}{\delta^2 + \sigma^2} e^{\delta \frac{\theta}{n}} \right\}^n = (\alpha\sigma^2 + \gamma)^n \quad (4.23)$$

と表記できる．算術平均と幾何平均の関係を用いれば

$$\begin{aligned}
(\alpha\sigma^2 + \gamma)^n &= \left\{ \frac{\sum_{i=1}^n (\alpha\sigma_i^2 + \gamma)}{n} \right\}^n \\
&\geq \left\{ \sqrt[n]{\prod_{i=1}^n (\alpha\sigma_i^2 + \gamma)} \right\}^n \\
&= \prod_{i=1}^n (\alpha\sigma_i^2 + \gamma)
\end{aligned} \tag{4.24}$$

が成り立つ．

ここで $\eta(Z_i)$ の確率変数 Z_i に対する期待値は

$$\begin{aligned}
E_{Z_i} [\eta(Z_i)] &= \alpha E_{Z_i} [Z_i^2] + \beta E_{Z_i} [Z_i] + \gamma \\
&= \alpha\sigma_i^2 + \gamma
\end{aligned} \tag{4.25}$$

である．したがって式 (4.24) の右辺は

$$\prod_{i=1}^n (\alpha\sigma_i^2 + \gamma) = \prod_{i=1}^n E_{Z_i} [\eta(Z_i)] \tag{4.26}$$

である．

2 次関数 $\eta(Z_i)$ は狭義単調増加ではないため，確率変数 $\eta(Z_i)$ の確率密度関数は Z_i の確率密度関数と一意に対応していない．しかしながら， $\eta(Z_i)$ に対する期待値については

$$E_{Z_i} [\eta(Z_i)] = E_{\eta(Z_i)} [\eta(Z_i)] \tag{4.27}$$

で一致するため

$$\prod_{i=1}^n E_{Z_i} [\eta(Z_i)] = \prod_{i=1}^n E_{\eta(Z_i)} [\eta(Z_i)] \tag{4.28}$$

である．確率変数 $\eta(Z_i), i = 1, \dots, n$ は互いに独立であるから，それらの期待値の積は積の期待値と一致する，すなわち

$$\prod_{i=1}^n E_{\eta(Z_i)} [\eta(Z_i)] = E_{\eta(Z_1), \dots, \eta(Z_n)} \left[\prod_{i=1}^n \eta(Z_i) \right] \tag{4.29}$$

が成り立つ．確率変数を $\eta(Z_i)$ から Z_i に戻すことで

$$\begin{aligned}
& E_{\eta(Z_1), \dots, \eta(Z_n)} \left[\prod_{i=1}^n \eta(Z_i) \right] \\
&= E_{Z_1, \dots, Z_n} \left[\prod_{i=1}^n \eta(Z_i) \right] \\
&= \int_{-\infty}^{\delta} \cdots \int_{-\infty}^{\delta} \prod_{i=1}^n \{ \eta(Z_i) f_{Z_i}(Z_i) \} dZ_n \cdots dZ_1 \tag{4.30}
\end{aligned}$$

を得る．

Hoeffding の確率不等式に用いられた $\prod_{i=1}^n e^{\frac{Z_i}{n}\theta}$ は, $\frac{1}{n} \sum_{i=1}^n Z_i = \bar{Z}$ の条件下において, $Z_i, i = 1, \dots, n$ の組合せに依らず $\bar{Z}$ の値のみで一意に定まっていた．他方で, $\prod_{i=1}^n \eta(Z_i)$ は $\frac{1}{n} \sum_{i=1}^n Z_i = \bar{Z}$ の条件下においても $Z_i, i = 1, \dots, n$ の組合せによって値が異なる．ここで, $\bar{Z} = z$ の条件下で計算した α, β, γ を用いる $\prod_{i=1}^n \phi\left(\frac{Z_i}{n}\right)$ の最小値

$$\begin{aligned}
\Phi(\bar{Z}) &= \min_{Z_i, \forall i} \prod_{i=1}^n \eta(Z_i) \\
\text{s.t. } & \frac{1}{n} \sum_{i=1}^n Z_i = \bar{Z} \\
& Z_i \leq \delta, \quad \forall i
\end{aligned} \tag{4.31}$$

を定義する．任意の Z_i の組合せに対して

$$\prod_{i=1}^n \eta(Z_i) f_{Z_i}(Z_i) \geq \Phi\left(\frac{1}{n} \sum_{i=1}^n Z_i\right) \prod_{i=1}^n f_{Z_i}(Z_i) \tag{4.32}$$

であるから, 式 (4.30) の 2 段目に対して

$$\begin{aligned}
& \int_{-\infty}^{\delta} \cdots \int_{-\infty}^{\delta} \prod_{i=1}^n \{ \eta(Z_i) f_{Z_i}(Z_i) \} dZ_n \cdots dZ_1 \\
& \geq \int_{-\infty}^{\delta} \cdots \int_{-\infty}^{\delta} \Phi\left(\frac{1}{n} \sum_{i=1}^n Z_i\right) \prod_{i=1}^n f_{Z_i}(Z_i) dZ_n \cdots dZ_1 \\
& = E_{Z_1, \dots, Z_n} \left[\Phi\left(\frac{1}{n} \sum_{i=1}^n Z_i\right) \right] \tag{4.33}
\end{aligned}$$

が成り立つ。そして $\Phi\left(\frac{1}{n}\sum_{i=1}^n Z_i\right)$ の値は、 $Z_i, i = 1, \dots, n$ の組合せに依らず、その算術平均 $\bar{Z} = \frac{1}{n}\sum_{i=1}^n Z_i$ の値のみで一意に定まる。したがって $Z_i, i = 1, \dots, n$ の同時確率に対する期待値と、確率変数 $\bar{Z}$ に対する期待値の間に

$$E_{Z_1, \dots, Z_n} \left[\Phi \left(\frac{1}{n} \sum_{i=1}^n Z_i \right) \right] = E_{\bar{Z}} [\Phi(\bar{Z})] \quad (4.34)$$

の関係が成り立つ。

ここで $\Phi(\bar{Z})$ に対して以下の定理が成り立つ。

定理 1. $\bar{Z} \geq 0$ において $\Phi(\bar{Z})$ は $\bar{Z}$ に対し狭義単調増加である。

定理 1 の証明を付録 B に示す。

定理 1 が真であるから、Hoeffding の確率不等式と同様に $z \geq 0$ を用いれば、式 (4.2) と同様の式変形が式 (4.34) 右辺に成り立つ。すなわち

$$\begin{aligned} E_{\bar{Z}} [\Phi(\bar{Z})] &= \int_{-\infty}^{\delta} \Phi(\bar{Z}) f_{\bar{Z}}(\bar{Z}) d\bar{Z} \\ &\geq \int_z^{\delta} \Phi(\bar{Z}) f_{\bar{Z}}(\bar{Z}) d\bar{Z} \\ &\geq \int_z^{\delta} \Phi(z) f_{\bar{Z}}(\bar{Z}) d\bar{Z} \\ &= \Phi(z) \int_z^{\delta} f_{\bar{Z}}(\bar{Z}) d\bar{Z} \\ &= \Phi(z) \Pr\{\bar{Z} > z\} \end{aligned} \quad (4.35)$$

である。

式 (4.23), (4.24), (4.26), (4.28)~(4.30) および (4.33)~(4.35) より

$$\left\{ \frac{\delta^2}{\delta^2 + \sigma^2} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} + \frac{\sigma^2}{\delta^2 + \sigma^2} e^{\delta \frac{\theta}{n}} \right\}^n \geq \Phi(z) \Pr\{\bar{Z} > z\} \quad (4.36)$$

の不等式が成り立つ。式 (4.36) を変形することで、 $\bar{Z}$ の上側確率の上界を求める確率不等式

$$\Pr\{\bar{Z} > z\} \leq \frac{1}{\Phi(z)} \left\{ \frac{\delta^2}{\delta^2 + \sigma^2} e^{-\frac{\sigma^2}{\delta} \frac{\theta}{n}} + \frac{\sigma^2}{\delta^2 + \sigma^2} e^{\delta \frac{\theta}{n}} \right\}^n \quad (4.37)$$

が導出される。

式 (4.37) と Hoeffding の確率不等式である式 (4.17) の違いは右辺の分母のみである。 $Z_i \leq \delta$ の範囲において $\eta(Z_i) \geq e^{\frac{Z_i}{n}\theta} > 0$ であるから、 $0 \leq z \leq \delta$ の任意の z に対して

$$\Phi(z) \geq e^{z\theta} \quad (4.38)$$

が成り立つ。 よって式 (4.37) の値は式 (4.17) 以下となる。

このように、もし $\Phi(z)$ の値が得られるならば、 Hoeffding の確率不等式よりもタイトな上界を式 (4.37) で与えることができる。

4.4 確率変数が2個の場合における性能改善

$\Phi(z)$ を求めるにあたり、 $\frac{1}{n} \sum_{i=1}^n Z_i = z$ を満足する Z_i を用いた $\prod_{i=1}^n \eta(Z_i)$ について考察する。 確率変数の個数が $n = 2$ であれば $Z_2 = 2z - Z_1$ であるから、 $\prod_{i=1}^2 \eta(Z_i)$ は Z_1 の四次関数となり、1種類の変数で表すことが可能となる。これを踏まえて、本節では $n = 2$ の場合に限定して Hoeffding の確率不等式の性能向上を考察する。

ここで式 (4.19) は

$$\eta(Z_i) = \frac{1}{4\alpha} \{ (2\alpha Z_i + \beta)^2 + (4\alpha\gamma - \beta^2) \} \quad (4.39)$$

と表すことができる。以降は Z_i に代えて $(2\alpha Z_i + \beta)$ の表現を用いる。

$\Phi(z)$ の探索にあたり、一度 $Z_i \leq \delta$ の条件を緩和して、 $\eta(Z_1) \times \eta(2z - Z_1)$ の大域的最小解を検討する。 $\eta(Z_1) \times \eta(2z - Z_1)$ は z を中心に左右対称で、上に開いた四次関数である。最大3個の停留点を持ち、そのいずれかが大域的最小解となる。したがって、 $\Phi(z)$ を与える可能性がある解の一つは、四次関数の停留点である、 Z_1, Z_2 の両方が

$$(2\alpha Z_i + \beta) = (2\alpha z + \beta) \quad (4.40)$$

となる組合せである。以降の考察ではこの解を重根解と呼称する。

$\Phi(z)$ を与える可能性がある残り二つの停留点は、 Z_1, Z_2 が

$$(2\alpha Z_i + \beta) = (2\alpha z + \beta) + \sqrt{(2\alpha z + \beta)^2 - (4\alpha\gamma - \beta^2)} \quad (4.41)$$

$$(2\alpha Z_i + \beta) = (2\alpha z + \beta) - \sqrt{(2\alpha z + \beta)^2 - (4\alpha\gamma - \beta^2)} \quad (4.42)$$

のペアで与えられる組合せである。以降の考察ではこの解を複号根解と呼称する。ただし式 (4.41) と (4.42) が実数でなければ解として有意ではない。式 (4.41) と

(4.42) が実数となる条件は、 $(2\alpha z + \beta)^2 \geq (4\alpha\gamma - \beta^2)$ を満足することである。このうち $(2\alpha z + \beta)^2 = (4\alpha\gamma - \beta^2)$ において、重根解と複号根解は一致する。

$(2\alpha z + \beta)^2 < (4\alpha\gamma - \beta^2)$ の場合は式 (4.40) が唯一の停留点となる。したがって $\prod_{i=1}^2 \eta(Z_i)$ の最小値は式 (4.40) を用いた

$$\prod_{i=1}^2 \eta(Z_i) = \left(\frac{1}{4\alpha}\right)^2 \{(2\alpha z + \beta)^2 + (4\alpha\gamma - \beta^2)\}^2 \quad (4.43)$$

で与えられる。

他方で式 (4.41) と (4.42) が実数であれば、これを用いた $\prod_{i=1}^2 \eta(Z_i)$ の値は

$$\prod_{i=1}^2 \eta(Z_i) = \left(\frac{1}{2\alpha}\right)^2 (2\alpha z + \beta)^2 (4\alpha\gamma - \beta^2) \quad (4.44)$$

で与えられる。 $\eta(Z_1) \times \eta(2z - Z_1)$ は上に開いた四次関数であることから、式 (4.41) と (4.42) が実数を与える場合には式 (4.43) と (4.44) について

$$\left(\frac{1}{4\alpha}\right)^2 \{(2\alpha z + \beta)^2 + (4\alpha\gamma - \beta^2)\}^2 \geq \left(\frac{1}{2\alpha}\right)^2 (2\alpha z + \beta)^2 (4\alpha\gamma - \beta^2) \quad (4.45)$$

が成立する。 $(2\alpha z + \beta)^2 = (4\alpha\gamma - \beta^2)$ であれば式 (4.45) は等式となる。以上のことから、式 (4.41) と (4.42) が実数を与える場合の $\prod_{i=1}^2 \eta(Z_i)$ の最小値は式 (4.44) となる。

以上は $Z_i \leq \delta$ の条件を緩和した状況での考察であるため、その条件を考慮してあらためて考察する。 $z \leq \delta$ であるから、式 (4.40)~(4.42) で与えられる最小解のうち $Z_i > \delta$ となり得るのは式 (4.41) に限られる。もし式 (4.41) で $Z_i > \delta$ となれば、その解で与えた $\prod_{i=1}^2 \eta(Z_i)$ を $\Phi(z)$ に用いることは不適切である。この場合、 $\eta(Z_1) \times \eta(2z - Z_1)$ は上に開いた四次関数であることから、一方の Z_i の値に δ を用いた $(Z_1, Z_2) = (\delta, 2z - \delta)$ あるいは $(Z_1, Z_2) = (\delta, 2z - \delta)$ が $\prod_{i=1}^2 \eta(Z_i)$ の最小値

$$\begin{aligned} \prod_{i=1}^2 \eta(Z_i) &= \left(\frac{1}{4\alpha}\right)^2 \{(2\alpha\delta + \beta)^2 + (4\alpha\gamma - \beta^2)\} \\ &\quad \times \{(2\alpha(2z - \delta) + \beta)^2 + (4\alpha\gamma - \beta^2)\} \end{aligned} \quad (4.46)$$

の候補となる。以降の考察では Z_i の一方を δ で与えた解をエッジ解と呼称する。

エッジ解を用いる場合、 $\eta(Z_1) \times \eta(2z - Z_1)$ は上に開いた四次関数であるから、式 (4.46) の値は式 (4.43) より小さい。したがって式 (4.41) で $Z_i > \delta$ の場合には、 $\prod_{i=1}^2 \eta(Z_i)$ の最小値はエッジ解を用いた式 (4.46) で与えられる。

以上の考察から、 $\Phi(z)$ とそれを与える Z_1, Z_2 の組合せが明らかになった。したがって、 $n = 2$ の場合は以下の手順を用いて Hoeffding の確率不等式よりもタイトな上側確率の上界を計算できる。

手順 1. 式 (4.14) より σ^2 を求める。

手順 2. z, n, δ, σ^2 を用いて式 (4.18) と (4.20)~(4.22) より $\theta, \alpha, \beta, \gamma$ を計算する。

手順 3. $(2\alpha z + \beta)^2 - (4\alpha\gamma - \beta^2)$ を計算する。その値が負であれば重根解を用いる。 $\Phi(z)$ を式 (4.43) で与えて手順 6. に進む。

手順 4. $(2\alpha z + \beta) + \sqrt{(2\alpha z + \beta)^2 - (4\alpha\gamma - \beta^2)}$ を計算する。その値が $(2\alpha\delta + \beta)$ より小さいならば複号根解を用いる。 $\Phi(z)$ を式 (4.44) で与えて手順 6. に進む。

手順 5. 手順 3. と手順 4. のどちらの条件にも該当しないためエッジ解を用いる。 $\Phi(z)$ を式 (4.46) で与えて手順 6. に進む。

手順 6. $\Phi(z)$ を利用して式 (4.37) で上側確率の上界を計算する。

4.5 確率変数の個数が任意の場合における性能改善

確率変数の個数 n がいかなる値であっても、 $\Phi(z)$ が得られるならば Hoeffding の確率不等式の性能改善が実現できる。ただし $\frac{1}{n} \sum_{i=1}^n Z_i = z$ を満足する Z_i を用いた $\prod_{i=1}^n \eta(Z_i)$ について、 $n = 2$ の場合には $\prod_{i=1}^2 \eta(Z_i)$ を Z_1 による 1 変数の関数として表現できたのに対して、 $n \geq 3$ の場合は $\prod_{i=1}^n \eta(Z_i)$ が $n - 1$ 次元に張る関数となる。したがって $n = 2$ の場合とは異なるアプローチで $\Phi(z)$ を探索する必要がある。

まず $\sum_{i=1}^n Z_i = nz$ の条件のもとで関数 $\prod_{i=1}^n \eta(Z_i)$ に有限の最小値が存在することを確認する。その考察にあたり、一度 $Z_i \leq \delta$ の条件を緩和する。 $\eta(Z_i)$ は $e^{Z_i\theta}$ を上から抑え込む下に凸な二次関数として定義される。したがって $\eta(Z_i)$ は最小値として正の有限値 $\eta(-\frac{\beta}{2\alpha})$ を持つ。これにより、空間 $\{Z_1, \dots, Z_n\}$ の全領域における $\prod_{i=1}^n \eta(Z_i)$ の最小値は $\prod_{i=1}^n \eta(-\frac{\beta}{2\alpha})$ で正の有限値として与えられる。他方で $\eta(Z_i)$ は $Z_i \geq -\frac{\beta}{2\alpha}$ にて単調増加、 $Z_i \leq -\frac{\beta}{2\alpha}$ にて単調減少であり、 $Z_i \rightarrow \pm\infty$ において $\lim_{Z_i \rightarrow \pm\infty} \eta(Z_i) \rightarrow \infty, \forall i$ である。これにより $\prod_{i=1}^n \eta(Z_i)$ は $\sum_{i=1}^n Z_i \rightarrow \infty$ において $\lim_{\sum_{i=1}^n Z_i \rightarrow \infty} \prod_{i=1}^n \eta(Z_i) \rightarrow \infty$ となる上に開いた関数である。以上の特性から、 $\sum_{i=1}^n Z_i = nz$ の条件のもとでの $\prod_{i=1}^n \eta(Z_i)$ は、有限で正の最小値を持つ。

ここで $\sum_{i=1}^n Z_i = nz$ は $Z_1, \dots, Z_n$ の一次式である。したがって $Z_i \leq \delta$ の条件を緩和すれば、 Z_i の算術平均が z であるという条件における $\prod_{i=1}^n \eta(Z_i)$ の最小値

をラグランジュの未定乗数法で求められる．その最小値を与える Z_i が $Z_i \leq \delta, \forall i$ ならば，得られた最小値が $\Phi(z)$ の値となる．これを踏まえて，ラグランジュ関数の停留点を与える Z_i の組合せを明らかにする．

ラグランジュ乗数 l を用いてラグランジュ関数

$$\mathcal{L}(Z_1, \dots, Z_n, l) = \prod_{i=1}^n \eta(Z_i) - l \left(\sum_{i=1}^n Z_i - zn \right) \quad (4.47)$$

を定義する．式 (4.47) より，ラグランジュ関数 $\mathcal{L}(Z_1, \dots, Z_n, l)$ の $Z_i, \forall i$ に関する偏微分での 1 階条件として

$$\frac{\partial \mathcal{L}(Z_1, \dots, Z_n, l)}{\partial Z_i} = (2\alpha Z_i + \beta) \prod_{\substack{j=1 \\ j \neq i}}^n \eta(Z_j) - l = 0 \quad (4.48)$$

を得る．ラグランジュ関数の停留点は式 (4.48) を満足する Z_i の組として与えられる．ここで式 (4.48) に $\eta(Z_i)$ を乗じて合計すれば

$$(2\alpha z + \beta)n \prod_{j=1}^n \eta(Z_j) = l \sum_{j=1}^n \eta(Z_j) \quad (4.49)$$

となり，式変形によって

$$l = (2\alpha z + \beta)n \frac{\prod_{j=1}^n \eta(Z_j)}{\sum_{j=1}^n \eta(Z_j)} \quad (4.50)$$

を得る．式 (4.50) を式 (4.48) に代入して整理すれば

$$\frac{(2\alpha Z_i + \beta)}{(2\alpha z + \beta)n} = \frac{\eta(Z_i)}{\sum_{j=1}^n \eta(Z_j)}, \quad \forall i \quad (4.51)$$

を得る．

ここで $Z_i = z, \forall i$ であれば，式 (4.51) は両辺とも $\frac{1}{n}$ となるため，全ての i に対して式 (4.51) を満足する．したがってラグランジュ関数の停留点を与える Z_i の組合せの一つを

$$(2\alpha Z_i + \beta) = (2\alpha z + \beta), \quad \forall i \quad (4.52)$$

すなわち $Z_i = z, \forall i$ で得る．以降では式 (4.52) の関係より得られる Z_i の組を重根解と呼称する．

他方で、式 (4.51) を満足する Z_i の組は重根解以外にも考えられる．まず式 (4.51) を変形して

$$\frac{(2\alpha z + \beta)n}{\sum_{j=1}^n \eta(Z_j)} = \frac{(2\alpha Z_i + \beta)}{\eta(Z_i)}, \quad \forall i \quad (4.53)$$

を得る．ここで任意の i に対して式 (4.53) が成り立つことから、 $1 \leq i \leq n, i < j \leq n$ である任意の i, j の組合せに関して

$$\frac{(2\alpha Z_i + \beta)}{\eta(Z_i)} = \frac{(2\alpha Z_j + \beta)}{\eta(Z_j)} \quad (4.54)$$

を満足する必要がある．ここで式 (4.39) を式 (4.54) に代入して整理すれば

$$(2\alpha Z_i + \beta)(2\alpha Z_j + \beta)(Z_j - Z_i) + (4\alpha\gamma - \beta^2)(Z_i - Z_j) = 0 \quad (4.55)$$

を得る． $Z_i = Z_j$ ならば式 (4.55) を満足するのは明らかであるが、 $Z_i \neq Z_j$ であっても式 (4.55) を満足できる可能性がある．

$Z_i \neq Z_j$ の条件のもとで式 (4.55) を満足する Z_i, Z_j の組合せについて考察する．式 (4.55) を $Z_i - Z_j \neq 0$ で除することで

$$(2\alpha Z_i + \beta)(2\alpha Z_j + \beta) = (4\alpha\gamma - \beta^2) \quad (4.56)$$

を得る．ここで二次関数 $\eta(Z_i)$ は $Z_i = -\frac{\beta}{2\alpha}$ で最小値をとるが、 $\eta(Z_i)$ が $e^{Z_i\theta}$ を上から抑え込む関数であるから $\eta(-\frac{\beta}{2\alpha}) > 0$ である．これを式 (4.39) の表現と組み合わせれば

$$4\alpha\gamma - \beta^2 > 0 \quad (4.57)$$

が導かれる．さらに $\eta(Z_i)$ は下に凸であるから $\alpha > 0$ 、 $\eta(Z_i)$ と $e^{Z_i\theta}$ の接点の条件 $Z_i = -\frac{\sigma_i^2}{\delta} < 0$ から $\beta > 0$ も導かれる．したがって式 (4.56) の左辺にある $(2\alpha Z_i + \beta)$ と $(2\alpha Z_j + \beta)$ は同符号でなくてはならず、式 (4.56) を満足する Z_i, Z_j の組は一意的な対応関係となる．

さらに、式 (4.56) を満足する Z_i, Z_j の組が同時に 2 組以上存在することはない．例えば $Z_i \neq Z_j$ かつ式 (4.56) を満足する Z_i, Z_j として、 $Z^{(L_i)} < Z^{(U_i)}$ である $(Z^{(L_1)}, Z^{(U_1)})$ と $(Z^{(L_2)}, Z^{(U_2)})$ の 2 組が同時に存在する場合を考える．

もし $Z^{(L_1)} \neq Z^{(U_2)}$ であれば、 $Z^{(L_1)}$ に対応する値は $Z^{(U_1)}$ で一意に定まることから、式 (4.56) が成立するためには $Z^{(U_2)} = Z^{(U_1)}$ でなければならない．しかれば、式 (4.56) の等式より $Z^{(U_2)}$ に対応する値は $Z^{(L_2)} = Z^{(L_1)}$ で一意に定まる．したがって $(Z^{(L_1)}, Z^{(U_1)}) = (Z^{(L_2)}, Z^{(U_2)})$ と 2 組は同じ値となり、 $Z^{(L_1)} \neq Z^{(U_2)}$ であ

れば式 (4.56) を満足する (Z_i, Z_j) は結局 1 組しか存在しないことが分かる．この考察は $Z^{(L_2)} \neq Z^{(U_1)}$ の関係のもと $Z^{(U_1)} = Z^{(U_2)}$ から検討しても同じ結論に至る．

次に $Z^{(L_1)} \neq Z^{(L_2)}$ の場合を考える．この場合， $Z^{(L_1)}$ に対応する値は $Z^{(U_1)}$ で一意に定まることから， $Z^{(L_2)} = Z^{(U_1)}$ となる． $Z^{(L_2)}$ の値が定まれば式 (4.56) より $Z^{(U_2)} = Z^{(L_1)}$ と一意に定まる．しかし，この場合は $Z^{(U_2)} = Z^{(L_1)} < Z^{(U_1)} = Z^{(L_2)}$ であり， $Z^{(L_2)} < Z^{(U_2)}$ の前提条件と矛盾する．したがって $Z^{(L_1)} \neq Z^{(L_2)}$ の場合，式 (4.56) を満足する 2 組以上の (Z_i, Z_j) は存在し得ない．この考察は $Z^{(U_1)} \neq Z^{(U_2)}$ の関係のもと $Z^{(U_2)} = Z^{(L_1)}$ から検討しても同じ結論に至る．

このように $Z_i \neq Z_j$ かつ式 (4.56) を満足する Z_i, Z_j の組が同時に 2 組以上存在することはない．したがって， $Z_i = z, \forall i$ 以外の停留点が存在するならば，それは $Z^a \neq Z^b$ である 2 種類の値 Z^a, Z^b を組合せたものである．具体的に， $1 \leq m \leq n-1$ の整数 m を用いて， $Z_i = Z^a$ となる Z_i が $(n-m)$ 個， $Z_i = Z^b$ となる Z_i が m 個であると想定する． Z_i の算術平均は z であるから

$$(n-m)(2\alpha Z^a + \beta) + m(2\alpha Z^b + \beta) = n(2\alpha z + \beta) \quad (4.58)$$

が成り立つ．式 (4.56) と式 (4.58) を連立して解くことにより， Z_i の値は

$$M(m) = \{(2\alpha z + \beta)n\}^2 - 4(n-m)m(4\alpha\gamma - \beta^2) \quad (4.59)$$

を用いて

$$\begin{aligned} (2\alpha Z_i + \beta) &= (2\alpha Z^a + \beta) \\ &= \frac{1}{2(n-m)} \left\{ (2\alpha z + \beta)n - \sqrt{M(m)} \right\}, \quad i = 1, \dots, n-m \end{aligned} \quad (4.60)$$

$$\begin{aligned} (2\alpha Z_i + \beta) &= (2\alpha Z^b + \beta) \\ &= \frac{1}{2m} \left\{ (2\alpha z + \beta)n + \sqrt{M(m)} \right\}, \quad i = n-m+1, \dots, n \end{aligned} \quad (4.61)$$

または平方根の項の正負が逆となる

$$\begin{aligned} (2\alpha Z_i + \beta) &= (2\alpha Z^a + \beta) \\ &= \frac{1}{2(n-m)} \left\{ (2\alpha z + \beta)n + \sqrt{M(m)} \right\}, \quad i = 1, \dots, n-m \end{aligned} \quad (4.62)$$

$$\begin{aligned} (2\alpha Z_i + \beta) &= (2\alpha Z^b + \beta) \\ &= \frac{1}{2m} \left\{ (2\alpha z + \beta)n - \sqrt{M(m)} \right\}, \quad i = n-m+1, \dots, n \end{aligned} \quad (4.63)$$

で与えられる．ただし Z_b を $m' = n-m$ 個と想定した場合の式 (4.60) および式 (4.61) は， Z_b を m 個と想定した場合の式 (4.63) および式 (4.62) にそれぞれ一

致する．したがって $m = 1, \dots, n-1$ の各々に対して式 (4.60) と (4.61) の組を計算すれば全ての組合せを網羅できる．

以上のように，重根解以外にラグランジュ関数の停留点となりうる Z_i は式 (4.60) と (4.61) に定義された $Z_i, \forall i$ の組に限定される．以降は式 (4.60) と (4.61) で与えられる Z_i の組合せを複号根解と呼称する．ここで式 (4.60) と (4.61) において平方根の内部が負の値となれば Z_i は虚数となるため，そのような解は $\prod_{i=1}^n \eta(Z_i)$ の最小値を求めるうえで検討すべき Z_i の組合せから外れることになる．したがって， $M(m) \geq 0$ となる m でのみ複号根解を計算する．

以上の考察より，ラグランジュ関数の停留点を与えうる Z_i の組合せを網羅できた．ただし，これらの停留点は $Z_i \leq \delta$ の条件を緩和したうえで導出された．導出された解のうち，式 (4.60) と (4.61) で定義される Z^a, Z^b は δ より大きな値となる可能性がある．もし $\prod_{i=1}^n \eta(Z_i)$ の最小値を与える解が $Z_i > \delta, \exists i$ の複号根解ならば，その最小値を $\Phi(z)$ に用いることは不適切である．

$\prod_{i=1}^n \eta(Z_i)$ の最小値を与える解に $Z_i > \delta$ が含まれるなら，定義域の上限となる $Z_i = \delta$ が含まれる Z_i の組が， $\sum_{i=1}^n Z_i = nz$ の条件下における $\prod_{i=1}^n \eta(Z_i)$ の有意な最小値を与える可能性がある．このことは， $\Phi(z)$ の探索を非線形計画問題として定式化し，Karush-Kuhn-Tucker (KKT) 条件 [33] を利用することで確認できる．その詳細は付録 C に示す．

具体的に H 個の Z_i が

$$Z_i = \delta, \quad i = n - H + 1, \dots, n \quad (4.64)$$

である Z_i の組合せを検討する．このとき，残る $n - H$ 個の Z_i の算術平均は $\frac{nz - H\delta}{n - H}$ である．

ここで正の整数である N と

$$\zeta = \frac{1}{N} \sum_{j=1}^N Z_j, \quad (4.65)$$

を用いて，平均が ζ となる N 個の確率変数による $\eta(Z_j)$ の積の最小値を

$$\begin{aligned} \Psi(\zeta, N) &= \min_{Z_j, \forall j} \prod_{j=1}^N \eta(Z_j) \\ \text{s.t.} \quad &\frac{1}{N} \sum_{j=1}^N Z_j = \zeta \\ &Z_j \leq \delta, \quad \forall j \end{aligned} \quad (4.66)$$

と表記する．式 (4.66) 内の $\eta(Z_j)$ には， $\bar{Z} = z$ の条件下で計算した α, β, γ を用いることに注意されたい．

Z_j の一部に式 (4.64) を利用した場合， $\sum_{i=1}^{n-H} Z_i + H\delta = nz$ を満足する H と $Z_i, i = 1, \dots, n-H$ の組合せは多数存在する．ただし $\prod_{i=n-H+1}^n \eta(Z_i) = \{\eta(\delta)\}^H$ であるから，そのような Z_i の組合せがもし $\prod_{i=1}^n \eta(Z_i)$ の最小値を与えるならば

$$\prod_{i=1}^n \eta(Z_i) = \{\eta(\delta)\}^H \Psi\left(\frac{nz - H\delta}{n-H}, n-H\right) \quad (4.67)$$

となる．つまり，重根解や複号根解以外に $\Phi(z)$ を与える可能性を持つ Z_i の組合せは，式 (4.64) に $\Psi(\frac{nz-H\delta}{n-H}, n-H)$ を与える $Z_i, i = 1, \dots, n-H$ を組合せた解に限定される．このような解を， $n=2$ の場合と同様にエッジ解と呼称する．

式 (4.66) で定義される $\Psi(\zeta, N)$ の計算に用いる Z_j は，式 (4.31) の $\Phi(\bar{Z})$ の計算に用いる解と同様に導出できる．まず重根解と同様に導出される

$$(2\alpha Z_j + \beta) = (2\alpha\zeta + \beta), \quad j = 1, \dots, N \quad (4.68)$$

すなわち $Z_j = \zeta$ が存在する．ここで $Z_i, i = 1, \dots, n-H$ に $\zeta = \frac{nz-H\delta}{n-H}$ を代入した式 (4.68)，残りの Z_i に式 (4.64) を用いた Z_i の組が $\prod_{i=1}^n \eta(Z_i)$ の最小値を与える候補の一つとなる．このように解の一部に重根解を援用したエッジ解を部分重根エッジ解と呼称する．一方で複号根解を援用することで， $\frac{1}{N} \sum_{j=1}^N Z_j = \zeta$ を満足する Z_j の組である

$$\begin{aligned} (2\alpha Z_j + \beta) &= (2\alpha Z^{a'} + \beta) \\ &= \frac{1}{2(N-m')} \left\{ (2\alpha\zeta + \beta)N - \sqrt{M'(m', \zeta, N)} \right\}, \\ j &= 1, \dots, N-m' \end{aligned} \quad (4.69)$$

$$\begin{aligned} (2\alpha Z_j + \beta) &= (2\alpha Z^{b'} + \beta) \\ &= \frac{1}{2m'} \left\{ (2\alpha\zeta + \beta)N + \sqrt{M'(m', \zeta, N)} \right\}, \\ j &= N-m'+1, \dots, N \end{aligned} \quad (4.70)$$

も $\Psi(\zeta, N)$ を与える可能性がある．ただし m' は $1 \leq m' \leq N-1$ の整数，また

$$M'(m', \zeta, N) = \{(2\alpha\zeta + \beta)N\}^2 - 4(N-m')m'(4\alpha\gamma - \beta^2) \quad (4.71)$$

である． $Z_i, i = 1, \dots, n-H$ に $\zeta = \frac{nz-H\delta}{n-H}$ ， $N = n-H$ を代入した式 (4.69) と (4.70)，残りの Z_i に式 (4.64) を用いた Z_i の組が $\prod_{i=1}^n \eta(Z_i)$ の最小値を与える候補の一つとなる．このように解の一部に複号根解を援用したエッジ解を部分複号

根エッジ解と呼称する．複号根解と同様に， $M'(m', \frac{nz-H\delta}{n-H}, n-H) \geq 0$ となる m' に対してのみ部分複号根エッジ解を検討する．

以上の検討から任意の H に対するエッジ解の候補が導出された．しかしながら， $Z^{a'}, Z^{b'}$ において $Z_j \geq \delta$ である可能性がある． $Z_j = \delta$ が h 個見つかった場合，これは $H+h$ 個の Z_i を δ で固定した場合の部分重根エッジ解として扱える．一方で $Z_j > \delta$ ならば確率不等式に用いるには不適切であるから，解の候補に含めない．また部分複号根解の検討にあたり， $n(2\alpha z + \beta) - H(2\alpha\delta + \beta) \geq 0$ でなければならない．したがって H の探索範囲は

$$1 \leq H \leq \frac{n(2\alpha z + \beta)}{2\alpha\delta + \beta} \quad (4.72)$$

の範囲における整数となる．エッジ解の検討が必要な場合には，各 H の値において部分重根エッジ解および部分複号根エッジ解を計算し，その中で最小の $\prod_{i=1}^n \eta(Z_i)$ を与える解を選ぶ．

以上の検討から， $\Phi(z)$ を与えうる Z_i の組合せは，重根解と複号根解，エッジ解の3種類に限定されることが明らかとなった．ただしエッジ解については，一定の条件下でなければ解の候補に含まれない．さらに，エッジ解には部分重根エッジ解と部分複号根エッジ解の2種類が存在する．これを考慮すれば，任意の確率変数の個数 n について，以下の手順で Hoeffding の確率不等式よりもタイトな上側確率の上界を計算できる．

手順 1. 式 (4.14) より σ^2 を求める．

手順 2. z, n, δ, σ^2 を用いて式 (4.18) と (4.20)~(4.22) より $\theta, \alpha, \beta, \gamma$ を計算する．

手順 3. 式 (4.52) を用いて重根解を導出する．

手順 4. $1 \leq m \leq n-1$ の m について式 (4.59) を計算し， $M(m) \geq 0$ を満足する全ての m に対して式 (4.60) と (4.61) を用いて複号根解を導出する．

手順 5. 手順 3. と手順 4. で求めた停留点から $\prod_{i=1}^n \eta(Z_i)$ の最小値を与える Z_i の組を選ぶ．それが $Z_i \leq \delta, \forall i$ を満足するなら，その $\prod_{i=1}^n \eta(Z_i)$ の最小値を $\Phi(z)$ として手順 11. に進む．

手順 6. 最小値を与える Z_i が $Z_i > \delta, \exists i$ であれば，手順 3. と手順 4. で求めた重根解と複号根解のうち $Z_i \leq \delta, \forall i$ を満足する Z_i の中で最小の $\prod_{i=1}^n \eta(Z_i)$ を与える Z_i の組を探す．

手順 7. 式 (4.72) を満たす整数 H に対して，式 (4.64) と (4.68) を用いて部分重根エッジ解を導出する．

手順 8. 式 (4.72) を満たす整数 H と式 (4.71) が正の値となる m' に対して,
式 (4.64),(4.69) および (4.70) を用いて部分複号根エッジ解を導出する.

手順 9. 手順 7. と手順 8. にて求めたエッジ解から式 (4.67) の最小値を与える Z_i
の組を選ぶ.

手順 10. 手順 6. にて選ばれた停留点と手順 9. にて選ばれたエッジ解のうち
 $\prod_{i=1}^n \eta(Z_i)$ が小さいほうを $\Phi(z)$ として手順 11. に進む.

手順 11. $\Phi(z)$ を利用して式 (4.37) で上側確率の上界を計算する.

なお, $n = 2$ であれば上記の手順で求めた重根解・複号根解・エッジ解は 4.4 節で
導出したそれぞれの値と一致する.

4.6 数値検証

具体的な計算例を通して提案手法の計算手順ならびにその計算結果が Hoeffding
の確率不等式よりもタイトな上側確率の上界値を与えることを示す. $E_{Z_i}[Z_i] = 0, \forall i$
および $\delta = 5$ としたうえで, $V_{Z_i}[Z_i] = \sigma_i^2 = 4, \forall i$ と $V_{Z_i}[Z_i] = \sigma_i^2 = 1, \forall i$ の 2 種
類の分散, $n = 2, 5, 10$ の 3 種類の確率変数の個数を組み合わせた 6 通りの条件を
扱う. それぞれの条件において, $0 \leq z < \delta$ の範囲を 0.01 間隔で計算する. 上側
確率の上界値を比較するにあたり, 提案手法と Hoeffding の確率不等式で計算され
る上側確率の上界に加えて, Hoeffding の確率不等式に対する提案手法の相対改善
比率

$$\text{Relative improvement ratio} = \frac{\Phi(z) - e^{z\theta}}{\Phi(z)} \quad (4.73)$$

を用いる. また各条件に対して $z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補を示すことで, 提
案手法の計算手順を例示する.

まず, $\sigma_i^2 = 4, \forall i$ における計算結果を紹介する.

表 4.1: $\sigma_i^2 = 4, n = 2, z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補とその評価値

	$\prod_{i=1}^n \eta(Z_i^\dagger)$	$(2\alpha Z_1 + \beta)$	$(2\alpha Z_2 + \beta)$
重根解	14.4031	1.4004	1.4004
複号根解	7.4937	0.1335	2.6673
エッジ解	11.1369	2.0634	0.7374

表 4.2: $\sigma_i^2 = 4, n = 5, z = 2\sqrt{V_{\bar{Z}}[\bar{Z}]}$ における解の候補とその評価値

		$\prod_{i=1}^n \eta(Z_i)$	$(2\alpha Z_i + \beta)$	
			1 種類目	2 種類目
重根解		17.0603	0.5213	–
	$m = 1$	16.3167	0.5951	0.2261
複号根解	$m = 2$	14.4360	0.7491	0.1796
	$m = 3$	9.9272	1.1236	0.1198
	$m = 4$	3.2183	2.3803	0.0565
エッジ解	$K = 2$ 重根解	12.3115	0.8912	0.2747

$\sigma_i^2 = 4, \forall i$ かつ $n = 2$ の場合について, $z = 2\sqrt{V_{\bar{Z}}[\bar{Z}]}$ における解の候補を表 4.1 に示す. 解の候補の探索にあたっては, まず複号根解が所定の z にて存在するか確認し, 存在するならば重根解と複号根解による $\prod_{i=1}^2 \eta(Z_i)$ の値を比較する. 今回の例において重根解 $(2\alpha Z_i + \beta) = 1.4004, i = 1, 2$ では $\prod_{i=1}^2 \eta(Z_i) = 14.4031$ である. 一方, 複号根解が $(2\alpha Z_1 + \beta) = 0.1335, (2\alpha Z_2 + \beta) = 2.6673$ として存在する. そして複号根解では $\prod_{i=1}^2 \eta(Z_i) = 7.4937$ である. ここで複号根解による $\prod_{i=1}^2 \eta(Z_i)$ の方が小さな値であるが, 複号根解の Z_2 について $(2\alpha\delta + \beta) = 2.0634 < (2\alpha Z_2 + \beta)$ であるから, 複号根解は存在するものの有意な解ではない. このような場合にはエッジ解を計算し, 有意な解である重根解とエッジ解を用いた $\prod_{i=1}^2 \eta(Z_i)$ の値を比較する. エッジ解で $(2\alpha Z_1 + \beta) = (2\alpha\delta + \beta)$ と対になるのは $(2\alpha Z_2 + \beta) = 0.7374$ であり, その値を用いて $\prod_{i=1}^2 \eta(Z_i) = 11.1369$ を得る. ここで $\prod_{i=1}^2 \eta(Z_i)$ の値は重根解よりエッジ解の方が小さい. したがって $\sigma_i^2 = 4, n = 2, z = 2\sqrt{V_{\bar{Z}}[\bar{Z}]}$ では $\Phi(z)$ の計算にエッジ解が採用される. $e^{z\theta} = 9.8555 < \Phi(z) = 11.1369$ であるから, 提案手法による上側確率の上界値は Hoeffding の確率不等式による上側確率の上界値よりタイトな値となる. 実際に Hoeffding の確率不等式では $\Pr\{\bar{Z} \geq z\} = 0.5060$ であるのに対し, 提案手法では $\Pr\{\bar{Z} \geq z\} = 0.5051$ とよりタイトな値を与えている. 図 4.1 に $\sigma_i^2 = 4, n = 2$ における上側確率の上界値と相対改善比率を示す. 計算を行った全ての z において, 提案手法による上側確率の上界値が Hoeffding の確率不等式による上側確率の上界値以下となっている. ただし, $z = 2.1$ において提案手法と Hoeffding の確率不等式が与える上側確率の上界は一致している. 相対改善比率は最大で 34% 程度である.

$\sigma_i^2 = 4, \forall i$ かつ $n = 5$ の場合について, $z = 2\sqrt{V_{\bar{Z}}[\bar{Z}]}$ における解の候補を表 4.2 に示す. $n = 5$ であれば複号根解の m の値域は $1 \leq m \leq 4$ であり, 最大 4 種類の複号根解が存在しうる. 今回の状況では 4 種類の m 全てにおいて複号根解が存在する. 重根解 $(2\alpha Z_i + \beta) = 0.5213, \forall i$ および 4 種類の複号根解のうち, $\prod_{i=1}^5 \eta(Z_i)$ の最小

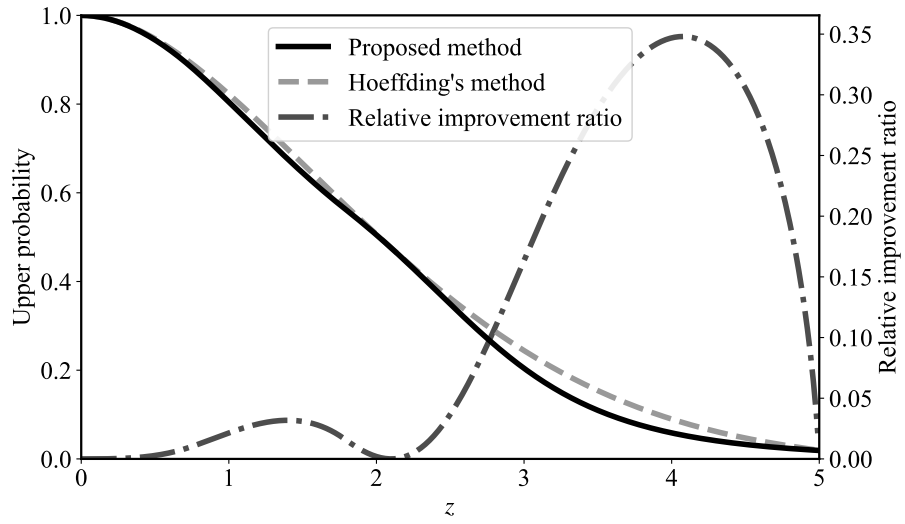


図 4.1: $\sigma_i^2 = 4, n = 2$ における上側確率の上界値と相対改善比率

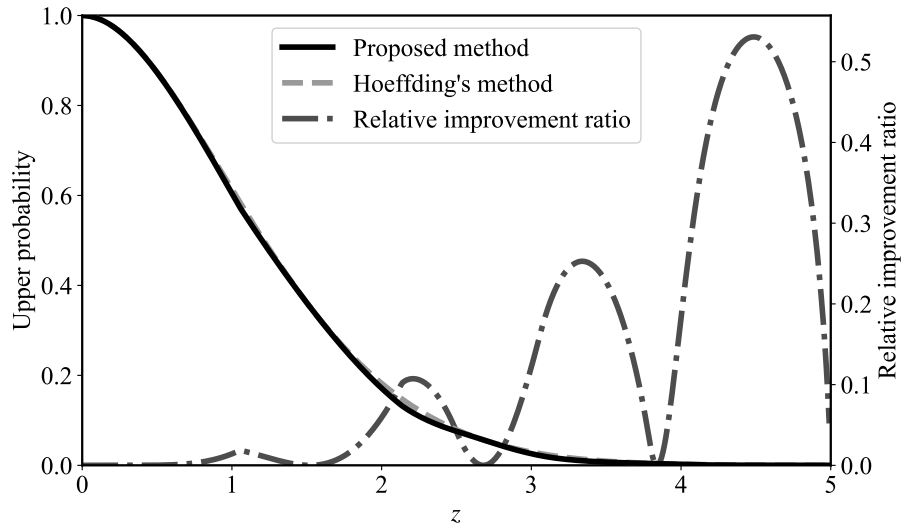


図 4.2: $\sigma_i^2 = 4, n = 5$ における上側確率の上界値と相対改善比率

表 4.3: $\sigma_i^2 = 4, n = 10, z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補とその評価値

		$\prod_{i=1}^n \eta(Z_i)$	$(2\alpha Z_i + \beta)$	
			1 種類目	2 種類目
重根解		19.3270	0.3059	–
	$m = 1$	19.2764	0.3152	0.2271
	$m = 2$	19.1834	0.3291	0.2133
	$m = 3$	18.9925	0.3514	0.1997
	$m = 4$	18.5504	0.3898	0.1800
複号根解	$m = 5$	17.4335	0.4589	0.1530
	$m = 6$	14.7195	0.5848	0.1200
	$m = 7$	9.5865	0.8200	0.0856
	$m = 8$	3.6619	1.3163	0.0533
	$m = 9$	0.5051	2.8365	0.0247
エッジ解	$K = 4$ 重根解	15.1897	0.5334	0.1543

値を与えるのは $m = 4$ における複号根解である。しかしながら、 $(2\alpha\delta + \beta) = 0.8912$ に対して $m = 4$ の複号根解は $(2\alpha Z^a + \beta) = 2.3803, (2\alpha Z^b + \beta) = 0.0565$ であり、 $Z^a \geq \delta$ であるからこの解は有意でない。算出した重根解と複号根解のうち有意な解は重根解と $m = 1, 2$ の複号根解の 3 種類であり、そのうち $\prod_{i=1}^5 \eta(Z_i)$ の最小値は $m = 2$ の複号根解における $\prod_{i=1}^5 \eta(Z_i) = 14.4360$ である。そのためエッジ解を算出し、 $\prod_{i=1}^5 \eta(Z_i)$ の値をエッジ解と $m = 2$ の複号根解で比較することになる。式 (4.72) より $n \frac{2\alpha z + \beta}{2\alpha\delta + \beta} = 5 \times \frac{0.5213}{0.8912} = 2.9 \dots$ であることから、エッジ解の H は 2 以下の値となる。その H の範囲において部分重根エッジ解と部分複号根エッジ解を比較した結果、エッジ解の代表として $H = 2$ における部分重根エッジ解が選ばれた。そのエッジ解を用いれば $\prod_{i=1}^5 \eta(Z_i) = 12.3115$ であり、 $m = 2$ の複号根解より小さな $\prod_{i=1}^5 \eta(Z_i)$ を得られる。したがって、 $\sigma_i^2 = 4, n = 5, z = 2\sqrt{V_Z[\bar{Z}]}$ では $\Phi(z)$ の計算に $H = 2$ の部分重根エッジ解が採用される。ここで $e^{z\theta} = 12.1079 < \prod_{i=1}^5 \eta(Z_i^\dagger)$ であるから、提案手法による上側確率の上界値は Hoeffding の確率不等式による上側確率の上界値よりタイトな値となる。図 4.2 に示す $\sigma_i^2 = 4, n = 5$ における上側確率の上界値と相対改善比率のグラフにおいても、提案手法が Hoeffding の確率不等式よりタイトな上側確率の上界値を与えることが確認できる。

$\sigma_i^2 = 4, \forall i$ かつ $n = 10$ の場合について、 $z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補を表 4.3 に示す。今回の条件では想定されうる $m = 1, \dots, 9$ に対して 9 個全てにおいて複号根解が存在する。重根解と 9 個の複号根解のうち、 $\prod_{i=1}^{10} \eta(Z_i)$ の最小値を与えるのは $m = 9$ の複号根解 $(2\alpha Z^a + \beta) = 2.8365, (2\alpha Z^b + \beta) = 0.0247$ で

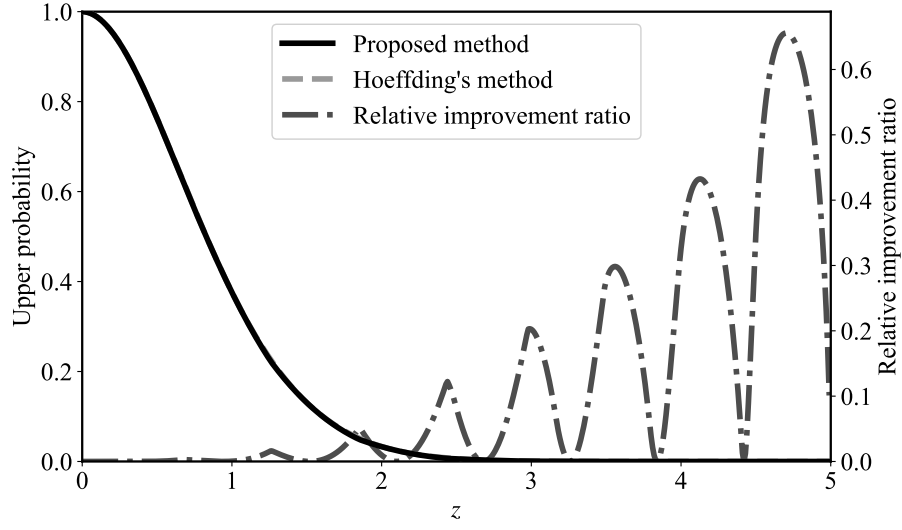


図 4.3: $\sigma_i^2 = 4, n = 10$ における上側確率の上界値と相対改善比率

表 4.4: $\sigma_i^2 = 1, n = 2, z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補とその評価値

	$\prod_{i=1}^n \eta(Z_i^\dagger)$	$(2\alpha Z_1 + \beta)$	$(2\alpha Z_2 + \beta)$
重根解	5.2144	1.2763	1.2763
複号根解	4.5711	0.3560	2.1965

ある．しかしながら， $(2\alpha\delta + \beta) = 0.5334$ より $Z^a > \delta$ であるから， $m = 9$ の複号根解は有意な解ではないため，エッジ解を求めることになる．式 (4.72) より $n \frac{2\alpha z + \beta}{2\alpha\delta + \beta} = 10 \times \frac{0.3059}{0.5334} = 5.7 \dots$ であるから，エッジ解の H は 5 以下の値である．検討の結果，エッジ解のうち $\prod_{i=1}^{10} \eta(Z_i)$ の最小値を与えるのは $H = 4$ の部分重根エッジ解であり，その場合 $\prod_{i=1}^{10} \eta(Z_i) = 15.1897$ となる．ここで重根解と複号根解のうち， $\prod_{i=1}^{10} \eta(Z_i)$ の最小値を与える有意な解は $m = 5$ の複号根解であり，その場合 $\prod_{i=1}^{10} \eta(Z_i) = 17.4335$ となる．したがって $\sigma_i^2 = 4, n = 10, z = 2\sqrt{V_Z[\bar{Z}]}$ では $\Phi(z)$ の計算に $H = 4$ の部分重根エッジ解が採用される． $e^{z\theta} = 14.9404 \leq \prod_{i=1}^{10} \eta(Z_i^\dagger)$ であるから，提案手法による上側確率の上界値は Hoeffding の確率不等式による上側確率の上界値よりタイトな値となる．図 4.3 に示す $\sigma_i^2 = 4, n = 10$ における上側確率の上界値と相対改善比率のグラフにおいても，提案手法が Hoeffding の確率不等式よりタイトな上側確率の上界値を与えることが確認できる．

次に， $\sigma_i^2 = 1, \forall i$ における計算結果を示す．

$\sigma_i^2 = 1, \forall i$ かつ $n = 2$ の場合について， $z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補を表 4.4 に示す．今回の例では複号根解が $(2\alpha Z_1 + \beta) = 0.3560, (2\alpha Z_2 + \beta) = 2.1965$ として存在する．ここで $(2\alpha\delta + \beta) = 3.1692$ であるから，複号根解は $(2\alpha\delta + \beta) \geq$

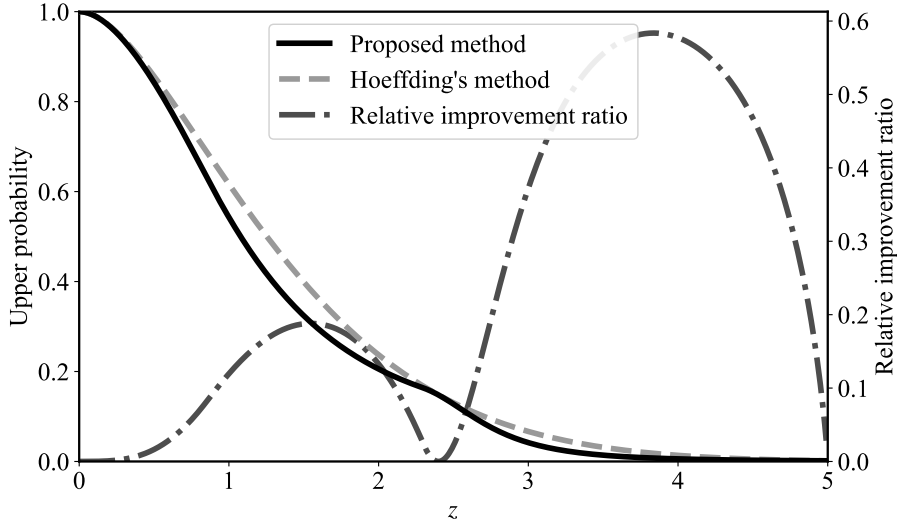


図 4.4: $\sigma_i^2 = 1, n = 2$ における上側確率の上界値と相対改善比率

表 4.5: $\sigma_i^2 = 1, n = 5, z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補とその評価値

		$\prod_{i=1}^n \eta(Z_i)$	$(2\alpha Z_i + \beta)$	
			1 種類目	2 種類目
重根解		6.5414	0.6226	—
	$m = 1$	6.5412	0.6336	0.5787
複号根解	$m = 2$	6.5395	0.6762	0.5423
	$m = 3$	6.4168	1.0143	0.3615
	$m = 4$	4.1048	2.5344	0.1447
エッジ解	$K = 1$ 重根解	5.1170	1.6862	0.3567

$(2\alpha Z_i + \beta), \forall i$ を満足する．したがって, $\sigma_i^2 = 1, n = 2, z = 2\sqrt{V_Z[\bar{Z}]}$ においては, エッジ解を検討する必要はなく, $\Phi(z)$ の計算には複号根解が採用される． $e^{z\theta} = 3.7311 < 5.2144 = \Phi(z)$ であるから, 提案手法による上側確率の上界値は Hoeffding の確率不等式による上側確率の上界値よりタイトな値となる．図 4.4 に $n = 2$ における上側確率の上界値と相対改善比率を示す．相対改善比率は最大で 58% 程度である．また $z = 2.4$ にて提案手法と Hoeffding の確率不等式で上側確率の上界が一致している．分散だけが異なる図 4.1 と図 4.4 を比較すれば, 相対改善比率は分散の小さい後の方が全体的に高い傾向にある．

$\sigma_i^2 = 1, \forall i$ かつ $n = 5$ の場合について, $z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補を表 4.5 に示す．今回の状況では $m = 1, \dots, 4$ の全てにおいて複号根解が存在する．重根解 $(2\alpha Z_i + \beta) = 0.6226, \forall i$ および 4 種類の複号根解のうち, $\prod_{i=1}^5 \eta(Z_i)$ の最小値を与え

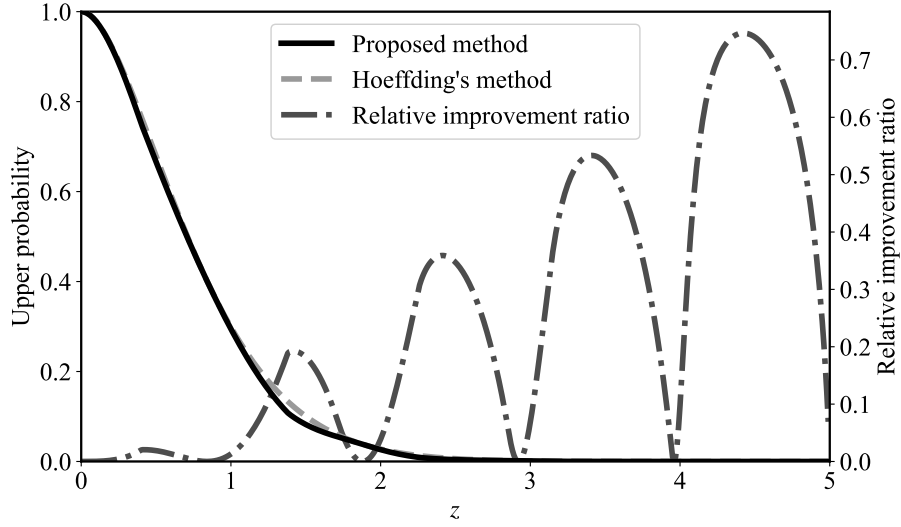


図 4.5: $\sigma_i^2 = 1, n = 5$ における上側確率の上界値と相対改善比率

るのは $m = 4$ における複号根解である．しかしながら、 $(2\alpha\delta + \beta) = 1.6862$ に対して $m = 4$ の複号根解では $(2\alpha Z^a + \beta) = 2.5344$ であり、 $Z^a \geq \delta$ であるからこの解は有意でない．算出した重根解と複号根解のうち有意な解は重根解と $m = 1, 2, 3$ の複号根解の 4 種類であり、そのうち $\prod_{i=1}^5 \eta(Z_i)$ の最小値は $m = 3$ の複号根解における $\prod_{i=1}^5 \eta(Z_i) = 6.4168$ である．式 (4.72) について $n \frac{2\alpha z + \beta}{2\alpha\delta + \beta} = 5 \times \frac{0.6226}{1.6862} = 1.8 \dots$ であるから、今回の条件では $H = 1$ のエッジ解が存在する．部分重根エッジ解と部分複号根エッジ解を比較した結果、エッジ解の代表として部分重根エッジ解が選ばれた．そのエッジ解を用いれば $\prod_{i=1}^5 \eta(Z_i) = 5.1170$ であり、 $m = 3$ の複号根解より小さな $\prod_{i=1}^5 \eta(Z_i)$ を得られる．したがって、 $\sigma_i^2 = 1, n = 5, z = 2\sqrt{V_Z[\bar{Z}]}$ では $\Phi(z)$ の計算に $H = 1$ の部分重根エッジ解が採用される．ここで $e^{z\theta} = 5.1103 < \prod_{i=1}^5 \eta(Z_i^\dagger)$ であるから、提案手法による上側確率の上界値は Hoeffding の確率不等式による上側確率の上界値よりタイトな値となる．図 4.5 に示す $\sigma_i^2 = 1, n = 5$ における上側確率の上界値と相対改善比率のグラフにおいても、提案手法が Hoeffding の確率不等式よりタイトな上側確率の上界値を与えることが確認できる．

$\sigma_i^2 = 1, \forall i$ かつ $n = 10$ の場合について、 $z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補を表 4.6 に示す．今回の条件では、 $m = 3, \dots, 7$ において $M(m) < 0$ となる．したがって複号根解は $m = 1, 2, 8, 9$ に対して検討する．重根解と 4 個の複号根解のうち、 $\prod_{i=1}^{10} \eta(Z_i)$ の最小値を与えるのは $m = 9$ の複号根解 $(2\alpha Z^a + \beta) = 3.5794, (2\alpha Z^b + \beta) = 0.0584$ である．これは、 $(2\alpha\delta + \beta) = 1.0809$ より $Z^a > \delta$ であるため有意な解ではないから、エッジ解を求めることになる．式 (4.72) について $n \frac{2\alpha z + \beta}{2\alpha\delta + \beta} = 10 \times \frac{0.4105}{1.0809} = 3.7 \dots$ であるから、エッジ解の H は 3 以下の値である．検討の結果、エッジ解のうち $\prod_{i=1}^{10} \eta(Z_i)$ の最小値を与えるのは $H = 2$ の部分重根エッジ解であり、その場合は

表 4.6: $\sigma_i^2 = 1, n = 10, z = 2\sqrt{V_Z[\bar{Z}]}$ における解の候補とその評価値

	$\prod_{i=1}^n \eta(Z_i)$	$(2\alpha Z_i + \beta)$	
		1 種類目	2 種類目
重根解	7.9440	0.4105	—
$m = 1$	7.9538	0.3977	0.5259
複号根解 $m = 2$	7.9792	0.3730	0.5608
$m = 8$	5.9898	1.4919	0.1402
$m = 9$	1.5518	3.5794	0.0584
エッジ解 $K = 2$ 重根解	6.8263	1.0809	0.2429

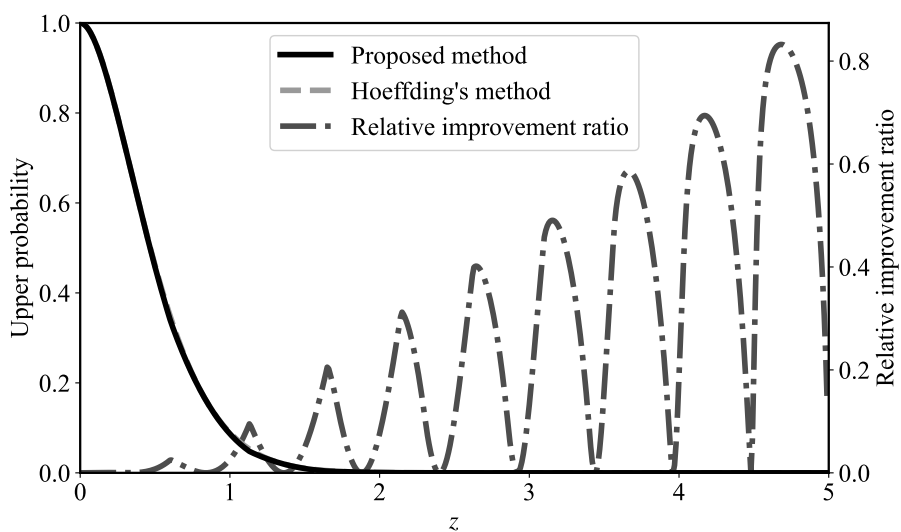


図 4.6: $\sigma_i^2 = 1, n = 10$ における上側確率の上界値と相対改善比率

$\prod_{i=1}^{10} \eta(Z_i) = 6.8263$ となる．ここで重根解と複号根解のうち， $\prod_{i=1}^{10} \eta(Z_i)$ の最小値を与える有意な解は重根解であり，その場合は $\prod_{i=1}^{10} \eta(Z_i) = 7.9440$ とエッジ解を用いるより大きな値となる．したがって $\sigma_i^2 = 1, n = 10, z = 2\sqrt{V_Z[\bar{Z}]}$ では $\Phi(z)$ の計算に $H = 2$ の部分重根エッジ解が採用される． $e^{z\theta} = 6.6789 \leq \prod_{i=1}^{10} \eta(Z_i^\dagger)$ であるから，提案手法による上側確率の上界値は Hoeffding の確率不等式による上側確率の上界値よりタイトな値となる．図 4.6 に示す $n = 10$ における上側確率の上界値と相対改善比率のグラフにおいても，提案手法が Hoeffding の確率不等式よりタイトな上側確率の上界値を与えることが確認できる．

4.7 結言

本章では，確率変数の平均・分散・定義域の上限という情報から上側確率の上界を求める Hoeffding の確率不等式に対して，よりタイトな上側確率の上界値を与える確率不等式を導出した． $e^{z\theta} \leq \prod_{i=1}^n \eta(Z_i)$ の関係を利用し，Hoeffding の確率不等式における $e^{z\theta}$ の代わりに， $\prod_{i=1}^n \eta(Z_i)$ の最小値となる $\Phi(z)$ を用いることでよりタイトな上側確率の上界値を与えた．ここで $\Phi(z)$ を与える Z_i の組合せについては，確率変数の個数が $n = 2$ の場合には $\prod_{i=1}^2 \eta(Z_i)$ が一変数関数と見なせることから，候補となり得る 3 種類の Z_i の組を導出した．このように求められた解のうち，どれが最小値を与えるかについて，その判別方法も明らかにした．任意の $n \geq 3$ の場合については，候補になり得る Z_i の組合せが，重根解，複号根，部分重根エッジ解，部分複号根エッジ解のいずれかに限られることを明らかにした．数値検証を通して，提案手法は Hoeffding の確率不等式よりもタイトな上側確率の上界値を与えることが確認された．

第5章 結論

データを活用して客観的な判断を補助する手法は、効果的な意思決定に寄与する。意思決定の対象に対して限られた情報しか得られない場合であっても、その情報を最大限活用することが求められる。本論文では、数理統計学に基づいて意思決定を補助する手法である変化点検出と確率不等式について、その性能向上を検討した。本論文で得られた成果を要約すると以下の通りである。

第2章ではポアソン分布に従う時系列データを対象とした複数ステップ変化点検出手法を設計した。提案手法は尤度とAICを援用したモデル比較によって、一連のデータから複数個の変化点とその個数、ならびにポアソン分布の強度パラメータを一意に推定可能である。特に尤度計算に動的計画法を用いることでモデル推定を高速に実施できる。この特徴を活かし、提案手法をリアルタイムな状態追跡に利用する利用例を示した。

第3章では線形変化を想定した複数変化点検出手法を設計した。従来の1回線形変化モデルを拡張する形で、持続的な工程変化の変化量が異常発生により複数回変化する統計モデルを提案した。尤度とAICを用いることで、正規分布に従う時系列データの平均を対象とした複数線形変化点検出手法を設計した。シミュレーションを用いて提案手法の性能を確認した結果、単調増加（あるいは単調減少）ではない場合や傾きの変化量がある程度大きな場合においては、複数個の変化点を変化点数とともに推定可能であることが確認できた。

第4章ではHoeffdingの確率不等式について、よりタイトな上側確率の上界を与えるよう改良した。具体的には確率不等式の導出過程でMarkovの不等式を適用する部分において、分子の導出に用いた二次関数を分母にも用いることで性能向上を実現した。Markovの不等式に適用するためには二次関数の積の最大値が必要であり、対象とする確率変数が2個の場合と3個以上の場合に分けて二次関数の積の最大値の導出手順を明らかにした。これにより、Hoeffdingの確率不等式と同じ確率変数の平均・分散・定義域の上限という情報を用いて、Hoeffdingの確率不等式以上にタイトな上側確率の上界値が計算可能となった。

本論文で取り組んだ研究において、今後検討すべき課題は以下の通りである。

第2章と第3章で扱った複数変化点検出における課題として、推定精度の向上と変化パターン自体の推定が挙げられる。

変化点をより正確に推定することは、誤った推定結果に影響されて無駄な作業に時間とコストを浪費するリスクを減らすためにも、重要なことである。第3章の数値検証で見受けられたように、変化量の小さい場合における推定精度の改善が変化点検出手法の課題である。複数変化点検出手法は、尤もらしいモデルを一つ選択する、いわば点推定を行う手法である。これに対して、Perry and Pignatiello [12] は変化点の区間推定について検討している。区間推定を行うことで、真に異常が生じた時点を見落とす可能性を減らすことができる。複数変化点検出では変化点数を推定するために情報量規準を用いているが、例えば下平 [23] は多重比較を用いることで情報量規準の信頼集合を求める方法を提案している。このような手法を援用することで、複数変化点推定においても信頼集合を求められる可能性がある。変化点の組合せに対して信頼集合を提示する手法は、検討すべき時点の候補が多少増加することになるものの、真の異常原因を見逃す可能性を減らすことができるため、意思決定を補助する手法として有望であろう。

第3章で述べた通り、工程で生じる変化パターンには突発的な変化と持続的な変化が存在し、推定精度向上のためには適切な変化パターンを用いた推定手法を用いる必要がある。その一方で、データ解析前にはどのような変化パターンに基づく変化が工程で生じているか分からないため、データに対して適切な手法を事前に決められないという矛盾がある。与えられたデータに対してどちらの変化パターンが適切であるかは、実際に両方の変化パターンに基づく手法を適用したうえで、作業者の経験と勘に基づく判断に頼らざるを得ない。ここで与えられたデータに対して、その変化パターン自体も推定できる手法が望まれる。その恩恵は、変化パターンへの当てはめを作業者に頼らず客観的に判断できる、ということにとどまらない。現在の複数変化点検出は、ステップ変化だけが複数回、あるいは線形変化だけが複数回というように、単一の変化パターンにしか対応していない。データから変化パターンを推定できれば、突発的な変化と線形変化の両方を扱える複数変化点検出手法を実現できる。そのような手法が実現すれば、それぞれの変化点に対して適切な変化パターンを当てはめることができるため、より正確に変化点を推定できる可能性がある。

第4章で扱った Hoeffding の確率不等式の性能改善においては、計算に用いる定数の工夫によって、さらにタイトな確率不等式が得られる可能性が残されている。提案手法は Hoeffding の確率不等式を利用して定義されたものであり、特に計算に用いる実数 θ は Hoeffding の確率不等式に対して最適化された値をそのまま利用している。したがって、この θ の値を提案手法に合わせて最適化できれば、さらにタイトな上側確率の上界値を得られる可能性がある。上側確率の上界値を最小とする θ の値を数理解析的あるいは数値解析的に取得することは、今後の検討課題である。

参考文献

- [1] Akaike, H.: A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, Vol. 19 (1974), 716–723.
- [2] Arizono, I. and Takemoto, Y.: A proposal for setting electric power saving rate to avoid risk of electric power shortage occurrence, *Synthesiology English edition*, Vol. 6 (2013), 137–146.
- [3] Assareh, H., Noorossana, R., Mohammadi, M. and Mengersen, K.: Bayesian multiple change point estimation of Poisson rates in control charts, *Scientia Iranica*, Vol. 23 (2016), 316–329.
- [4] Bellman, R. and Roth, R.: Curve fitting by segmented straight lines, *Journal of the American Statistical Association*, Vol. 64 (1969), 1079–1084.
- [5] Birge, J. R. and Louveaux, F.: *Introduction to stochastic programming*, Springer New York, NY, 1997.
- [6] Hawkins, D. M.: Fitting multiple change-point models to data, *Computational Statistics & Data Analysis*, Vol. 37 (2001), 323–341.
- [7] Hoeffding, W.: Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association*, Vol. 58 (1963), 13–30.
- [8] Laghari, A. A., Jumani, A. K. and Laghari, R. A.: Review and state of art of fog computing, *Archives of Computational Methods in Engineering*, Vol. 28 (2021), 3631–3643.
- [9] Niaki, S. T. A. and Khedmati, M.: Change point estimation of high-yield processes with a linear trend disturbance, *The International Journal fo Advanced Manufacturing Technology*, Vol. 69 (2013), 491–497.

- [10] Noorossana, R. and Shadman, A.: Estimating the change point of a normal process mean with a monotonic change, *Quality and Reliability Engineering International*, Vol. 25 (2009), 79–90.
- [11] Ogunniran, A. J., Adekeye, K. S., Adewara, J. A. and Adamu, M.: A review of change point estimation methods for process monitoring, *Applied and Computational Mathematics*, Vol. 10 (2021), 69–75.
- [12] Perry, M. B. and Pignatiello Jr., J. J.: Estimation of the change point of a normal process mean with a linear trend disturbance in SPC, *Quality Technology & Quantitative Management*, Vol. 3 (2006), 325–334.
- [13] Perry, M. B., Pignatiello Jr, J. J. and Simpson, J. R.: Estimating the change point of a Poisson rate parameter with a linear trend disturbance, *Quality and Reliability Engineering International*, Vol. 22 (2006), 371–384.
- [14] Perry, M. B., Pignatiello Jr., J. J. and Simpson, J. R.: Change point estimation for monotonically changing Poisson rates in SPC, *International Journal of Production Research*, Vol. 45 (2007), 1791–1813.
- [15] Samuel, T. R. and Pignatiello Jr., J. J.: Identifying the time of a change in a Poisson rate parameter, *Quality Engineering*, Vol. 10 (1998), 673–681.
- [16] Samuel, T. R., Pignatiello Jr., J. J. and Calvin, J. A.: Identifying the time of a step change with $\bar{X}$ control charts, *Quality Engineering*, Vol. 10 (1998), 521–527.
- [17] Shuto, T., Yasuhiko, T. and Ikuo, A.: Change point detection for multiple linear trend changes, in *The 22nd Asia Pacific Industrial Engineering and Management Society (APIEMS2022)*, 2022.
- [18] Takemoto, Y. and Arizono, I.: Study on state transition tracking method in process fraction nonconforming, in *Proceedings of 9th IFAC Conference on Manufacturing Modelling, Management and Control (MIM'19)*, 2019.
- [19] Tanabashi, S., Torisaka, Y., Takemoto, Y. and Arizono, I.: A state tracing method for the system data obeying Poisson distribution, in Huang, C.-Y., Dekkers, R., Chiu, S. F., Popescu, D. and Quezada, L. eds., *Intelligent and Transformative Production in Pandemic Times*, Cham, 2023, Springer International Publishing.

- [20] 有蘭育生, 竹本康彦: 電力不足発生リスク回避のための節電率設定方法への一提言, *Synthesiology*, Vol. 6 (2013), 140–151.
- [21] 有蘭育生, 棚橋秀斗, 竹本康彦: Hoeffding の確率不等式の性能向上, 日本経営工学会論文誌, Vol. 73 (2023), 189–209.
- [22] 有蘭育生, 友廣亮介, 浅原昂範, 竹本康彦: 異なる母集団からのデータの平均に関する Hoeffding の確率不等式の性能向上に関する考察, 日本経営工学会論文誌, Vol. 65 (2014), 61–69.
- [23] 英寿下平: モデルの信頼集合と地図によるモデル探索, 統計数理, Vol. 41 (1993), 131–147.
- [24] 本位田光重, 皆川健多郎: 生産工学—ものづくりマネジメント工学—, コロナ社, 2012.
- [25] 竹本康彦, 有蘭育生: 状態変化追跡方法, 工程管理支援装置, 工程管理支援方法, 及び状態変化追跡方法及び工程管理支援方法を実行させるためのプログラム, 特開 2013-029901, 2013.
- [26] 竹本康彦, 有蘭育生: 状態変化追跡方法の提案とその利活用に関する一考察, 日本経営工学会論文誌, Vol. 66 (2015), 240–248.
- [27] 竹本康彦, 有蘭育生, 棚橋秀斗: モノづくりのスマート化に資する品質管理の方法: ベイズ理論・情報理論・統計科学を通じて, オペレーションズ・リサーチ: 経営の科学, Vol. 68 (2023), 232–240.
- [28] 竹本康彦, 岩本史恵, 有蘭育生: 限定された需要情報のもとで許容欠品率を保証する発注点の決定方法に関する一考察, 日本経営工学会論文誌, Vol. 62 (2011), 21–24.
- [29] 棚橋秀斗, 有蘭育生, 竹本康彦: Hoeffding の確率不等式の根本的な性能向上のための基礎的考察, 日本経営工学会論文誌, Vol. 73 (2022), 70–73.
- [30] 棚橋秀斗, 竹本康彦, 有蘭育生: ポアソン過程における状態追跡法, 日本経営工学会論文誌, Vol. 72 (2021), 159–168.
- [31] 棚橋秀斗, 竹本康彦, 有蘭育生: 変化点検出モデルの考察と展開: 製造工程での的確かつ迅速な異常原因究明のために, オペレーションズ・リサーチ: 経営の科学, Vol. 68 (2023), 241–249.

- [32] 山口俊之：図解入門ビジネス 最新 生産工場の DX がよ〜くわかる本, 秀和システム, 2021.
- [33] 山下信雄：非線形計画法, 朝倉書店, 2015.
- [34] 光岡拓夢, 有蘭育生, 竹本康彦：Hoeffding の確率不等式の性能向上のための数値解析的手法に関する研究, 日本経営工学会論文誌, Vol. 71 (2020), 76–79.
- [35] 藤原秀樹：工場におけるデータ駆動型マネジメントを実現する Digital Factory への取り組み, 横河技報, Vol. 1 (2021), 19–25.
- [36] 日本規格協会（編）：品質管理, JIS ハンドブック, No. 57, 日本規格協会, 2022.
- [37] 稲本稔, 細野泰彦：わかりやすい 品質管理, オーム社, 第 4 版, 2016.
- [38] 海津聰：不等式の工学への応用, 森北出版, 2004.

付 録 A 初期値を既知とした場合の ポアソン分布を想定した 複数変化点検出手法

ポアソン分布に従う工程について，管理状態におけるポアソン分布の強度が λ_0^* で既知とする．その場合，式 (2.2) に示した統計モデルにおいてパラメータの初期値は $\lambda_0 = \lambda_0^*$ で与えられる．さらに対数尤度関数について

$$\begin{aligned} \log L(\boldsymbol{\tau}^{(K)}, \boldsymbol{\lambda}^{(K)} \mid \mathbf{X}, T, \lambda_0^*) \\ = \sum_{k=1}^K \sum_{t=\tau_{k-1}+1}^{\tau_k} \left\{ x_t \log \left(\frac{\lambda_k}{\lambda_0^*} \right) - (\lambda_k - \lambda_0^*) \right\} + \sum_{t=1}^T \{ x_t \log(\lambda_0^*) - \lambda_0^* - \log(x_t!) \} \end{aligned} \quad (\text{A.1})$$

と整理できる．各変化点に対応する強度 $\lambda_k, k = 1, \dots, K$ は式 (2.6) で計算できる．式 (A.1) からモデルに依存しない項を除けば，対数尤度関数のうち 2 個の連続する変化点 τ_{k-1}, τ_k に挟まれた区間における最大値は

$$\Lambda^*(\tau_{k-1}, \tau_k, \hat{\lambda}_k \mid \mathbf{X}, \lambda_0^*) = \sum_{t=\tau_{k-1}}^{\tau_k} \left\{ x_t \log \left(\frac{\hat{\lambda}_k}{\lambda_0^*} \right) - (\hat{\lambda}_k - \lambda_0^*) \right\} \quad (\text{A.2})$$

で計算される．式 (A.2) を用いれば，ある時点 t までに k 回の変化が生じた場合の対数尤度関数の最大化問題は

$$D^*(k, t \mid \lambda_0^*) = \max_{k \leq \tau_{k-1} < t} \left\{ D^*(k-1, \tau_{k-1} \mid \lambda_0^*) + \Lambda^*(\tau_{k-1}, t, \hat{\lambda}_k \mid \mathbf{X}, \lambda_0^*) \right\} \quad (\text{A.3})$$

となる．式 (A.3) の境界条件は

$$D^*(0, \tau_0 \mid \lambda_0^*) \equiv 0 \quad (\text{A.4})$$

である．式 (A.3) を用いることで，所与の変化点数 K における変化点の組合せの最尤推定量は

$$\hat{\boldsymbol{\tau}}^{(K)} = \arg \max_{\boldsymbol{\tau}^{(K)}} D^*(K, T \mid \lambda_0^*) \quad (\text{A.5})$$

で与えられる．ここで式 (A.3) の計算には動的計画法が適用できる．

$\lambda_0 = \lambda_0^*$ と既知の値で与えられるから，式 (2.2) の統計モデルで推定されるパラメータは $\boldsymbol{\tau}^{(K)} = (\tau_0, \dots, \tau_{K-1})$ と $\lambda_k, k = 1, \dots, K$ である．したがって AIC のパラメータ数によるペナルティ項は

$$B^*(K) = 2K \quad (\text{A.6})$$

と定義される．AIC を援用することで，変化点数は

$$\hat{K} = \arg \min_{0 \leq K < T} \{-D^*(K, T \mid \lambda_0^*) + 2K\} \quad (\text{A.7})$$

で推定される．

付 録 B Hoeffding の確率不等式の 改良に用いる 定理 1 の証明

$Z^\dagger > Z' \geq 0$ の関係を満足する任意の $Z^\dagger, Z'$ について考察を行う．まず $\Phi(Z^\dagger)$ を与える Z_i の組合せを $Z_i^\dagger, i = 1, \dots, n$ と定義する． $Z_i^\dagger$ の算術平均は

$$\frac{1}{n} \sum_{i=1}^n Z_i^\dagger = Z^\dagger > Z' \quad (\text{B.1})$$

である．さらに， $Z_i^\dagger$ のうち値が 0 以上となるインデックスの集合

$$\mathbb{M} = \left\{ i \in \{1, \dots, n\} \mid Z_i^\dagger > 0 \right\} \quad (\text{B.2})$$

を定義する． $Z^\dagger > 0$ より $\mathbb{M} \neq \emptyset$ である．これを用いて変数列

$$Z'_i = \begin{cases} Z_i^\dagger \left(1 - \frac{n(Z^\dagger - Z')}{\sum_{i \in \mathbb{M}} Z_i^\dagger} \right), & i \in \mathbb{M} \\ Z_i^\dagger & i \notin \mathbb{M} \end{cases} \quad (\text{B.3})$$

を定義する．明らかに $\frac{n(Z^\dagger - Z')}{\sum_{i \in \mathbb{M}} Z_i^\dagger} > 0$ である． Z'_i の算術平均は

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n Z'_i &= \frac{1}{n} \sum_{i=1}^n Z_i^\dagger - \frac{(Z^\dagger - Z')}{\sum_{i \in \mathbb{M}} Z_i^\dagger} \sum_{i \in \mathbb{M}} Z_i^\dagger \\ &= Z' \end{aligned} \quad (\text{B.4})$$

である．

ここで

$$\sum_{i \notin \mathbb{M}} Z_i^\dagger = \sum_{i \notin \mathbb{M}} Z'_i \leq 0 \quad (\text{B.5})$$

より,

$$\begin{aligned}
n(Z^\dagger - Z') &= \sum_{i=1}^n Z_i^\dagger - \sum_{i=1}^n Z'_i \\
&= \sum_{i \in \mathbb{M}} Z_i^\dagger + \sum_{i \notin \mathbb{M}} Z_i^\dagger - \sum_{i \in \mathbb{M}} Z'_i - \sum_{i \notin \mathbb{M}} Z'_i \\
&= \sum_{i \in \mathbb{M}} Z_i^\dagger - \sum_{i \in \mathbb{M}} Z'_i
\end{aligned} \tag{B.6}$$

が成り立つ. $Z' \geq 0$ と式 (B.5) より

$$\sum_{i \in \mathbb{M}} Z'_i = nZ' - \sum_{i \notin \mathbb{M}} Z'_i \geq 0 \tag{B.7}$$

ある. よって

$$n(Z^\dagger - Z') \leq \sum_{i \in \mathbb{M}} Z_i^\dagger \tag{B.8}$$

である.

以上のことから $0 < \frac{n(Z^\dagger - Z')}{\sum_{i \in \mathbb{M}} Z_i^\dagger} \leq 1$ が成り立つ. したがって $0 \leq Z'_i < Z_i^\dagger, i \in \mathbb{M}$ である. $\eta(Z_i)$ は $Z_i \geq 0$ で狭義単調増加であるから

$$\eta(Z'_i) < \eta(Z_i^\dagger), \quad i \in \mathbb{M} \tag{B.9}$$

$$\eta(Z'_i) = \eta(Z_i^\dagger), \quad i \notin \mathbb{M} \tag{B.10}$$

である. $\eta(Z_i) \geq e^{\frac{Z_i}{n}\theta} > 0$ を踏まえれば

$$\prod_{i=1}^n \eta(Z'_i) < \prod_{i=1}^n \eta(Z_i^\dagger) = \Phi(Z^\dagger) \tag{B.11}$$

が成り立つ. さらに $\Phi(Z') \leq \prod_{i=1}^n \eta(Z'_i)$ の関係より

$$\Phi(Z') < \Phi(Z^\dagger) \tag{B.12}$$

である. ここで Z' は $Z' \geq 0$ を満足する任意の値であった. したがって, $\bar{Z} \geq 0$ において $\Phi(\bar{Z})$ は $\bar{Z}$ に対して狭義単調増加である.

付 録 C KKT 条件を用いたエッジ解の導出

$\Phi(z)$ の値を求める問題は、非線形計画問題

$$\begin{aligned} & \text{minimize} && \prod_{i=1}^n \eta(Z_i) \\ & \text{subject to} && nz - \sum_{i=1}^n Z_i = 0 \\ & && Z_i - \delta \leq 0, \quad i = 1, \dots, n \end{aligned} \tag{C.1}$$

として定義できる．ここで制約条件は全て一次式であるから，式 (C.1) には，その局所的最小解 $\mathbf{Z}^* = (Z_1^*, \dots, Z_n^*)^\top$ に対して KKT 条件

$$\nabla \prod_{i=1}^n \eta(Z_i) \Big|_{\mathbf{Z}=\mathbf{Z}^*} + l^* \nabla \left(nz - \sum_{i=1}^n Z_i \right) \Big|_{\mathbf{Z}=\mathbf{Z}^*} + \sum_{i=1}^n \left\{ \rho_i^* \nabla (Z_i - \delta) \Big|_{Z_i=Z_i^*} \right\} = \mathbf{0} \tag{C.2}$$

$$nz - \sum_{i=1}^n Z_i^* = 0 \tag{C.3}$$

$$Z_i^* - \delta \leq 0 \text{ for } i = 1, \dots, n \tag{C.4}$$

$$\rho_i^* \geq 0 \text{ for } i = 1, \dots, n \tag{C.5}$$

$$\rho_i^* (Z_i^* - \delta) = 0 \text{ for } i = 1, \dots, n \tag{C.6}$$

を満足する l と $\boldsymbol{\rho}^* = (\rho_1^*, \dots, \rho_n^*)$ ，ひいては KKT 点 $(\mathbf{Z}^*, l^*, \boldsymbol{\rho}^*)$ が存在する．

式 (C.1) の局所的最小解を求めるために， $\rho_i^* = 0, \forall i$ の場合と $\rho_i^* > 0, \exists i$ の場合に分けて KKT 点を導出する．

まず $\rho_i^* = 0, \forall i$ の場合における KKT 点について考察する．この場合，式 (C.2) の目的関数は

$$\nabla \prod_{i=1}^n \eta(Z_i) \Big|_{\mathbf{Z}=\mathbf{Z}^*} + l^* \nabla \left(nz - \sum_{i=1}^n Z_i \right) \Big|_{\mathbf{Z}=\mathbf{Z}^*} = \mathbf{0} \tag{C.7}$$

となる．ここで式 (C.7) の各要素は式 (4.48) と一致する．したがって KKT 点にて $\rho_i^* = 0, \forall i$ であれば，これを満足する局所的最小解は重根解もしくは $Z_i \leq \delta, \forall i$ の複号根解となる．

次に， $\rho_i^* > 0, \exists i$ の場合における KKT 点について考察する．式 (C.1) は変数 $Z_i, i = 1, \dots, n$ の並び順に依存しない．これを踏まえて， $\rho_i^* > 0$ となる ρ_i^* が H 個存在する場合を代表して， $\rho_i^* = 0, i = 1, \dots, n-H$ かつ $\rho_i^* > 0, i = n-H+1, \dots, n$ の場合を扱う．

式 (C.6) より， $\rho_i^* > 0$ ならば

$$Z_i^* = \delta, \quad i = n-H+1, \dots, n \quad (\text{C.8})$$

となる．この場合，式 (C.2) の i 番目の要素は

$$(2\alpha Z_i^* + \beta) (\eta(\delta))^H \prod_{\substack{j=1 \\ j \neq i}}^{n-H} \eta(Z_j^*) - l^* = 0, \quad i = 1, \dots, n-H \quad (\text{C.9})$$

$$(2\alpha\delta + \beta) (\eta(\delta))^{H-1} \prod_{j=1}^{n-H} \eta(Z_j^*) - l^* + \rho_i^* = 0, \quad i = n-H+1, \dots, n \quad (\text{C.10})$$

である．式 (C.10) 左辺の第 1 項と第 2 項は $i = n-H+1, \dots, n$ で共通であるから

$$\rho_{n-H+1}^* = \dots = \rho_n^* \equiv \rho_{(\delta)}^* > 0 \quad (\text{C.11})$$

である．

式 (C.10) に式 (C.9) を代入することで， $i = 1, \dots, n-H$ に対して

$$\rho_{(\delta)}^* = (2\alpha Z_i^* + \beta) (\eta(\delta))^H \prod_{\substack{j=1 \\ j \neq i}}^{n-H} \eta(Z_j^*) - (2\alpha\delta + \beta) (\eta(\delta))^{H-1} \prod_{j=1}^{n-H} \eta(Z_j^*) \quad (\text{C.12})$$

が成り立つ．式 (C.12) の両辺に $\eta(Z_i^*)$ を乗じて合計すれば

$$\begin{aligned} \rho_{(\delta)}^* \sum_{j=1}^{n-H} \eta(Z_j^*) \sum_{j=1}^{n-H} \eta(Z_j^*) &= (\eta(\delta))^H \sum_{j=1}^{n-H} (2\alpha Z_j^* + \beta) \prod_{j=1}^{n-H} \eta(Z_j^*) \\ &\quad - (2\alpha\delta + \beta) (\eta(\delta))^{H-1} \prod_{j=1}^{n-H} \eta(Z_j^*) \sum_{j=1}^{n-H} \eta(Z_j^*) \end{aligned} \quad (\text{C.13})$$

となる．ここで

$$\sum_{j=1}^{n-H} (2\alpha Z_j^* + \beta) = (n-H) \left(2\alpha \frac{nz - H\delta}{n-H} + \beta \right) \quad (\text{C.14})$$

であるから, $\rho_{(\delta)}^*$ の値を

$$\begin{aligned} \rho_{(\delta)}^* = & \frac{(n-H)}{\sum_{j=1}^{n-H} \eta(Z_j^*)} \left(2\alpha \frac{nz - H\delta}{n-H} + \beta \right) (\eta(\delta))^H \prod_{j=1}^{n-H} \eta(Z_j^*) \\ & - (2\alpha\delta + \beta) (\eta(\delta))^{H-1} \prod_{j=1}^{n-H} \eta(Z_j^*) \end{aligned} \quad (\text{C.15})$$

と表記できる.

式 (C.15) を式 (C.10) に代入することで

$$l^* = \frac{(n-H)}{\sum_{j=1}^{n-H} \eta(Z_j^*)} \left(2\alpha \frac{nz - H\delta}{n-H} + \beta \right) (\eta(\delta))^H \prod_{j=1}^{n-H} \eta(Z_j^*) \quad (\text{C.16})$$

が導出される. さらに式 (C.16) を式 (C.9) に代入することで

$$\frac{(2\alpha \frac{nz-H\delta}{n-H} + \beta) (n-H)}{\sum_{j=1}^{n-H} \eta(Z_j^*)} = \frac{(2\alpha Z_i^* + \beta)}{\eta(Z_i^*)}, \quad i = 1, \dots, n-H \quad (\text{C.17})$$

が導出される. 式 (C.8) は, 確率変数の個数 n を $n-H$, 確率変数の算術平均 z を $\frac{nz-H\delta}{n-H}$ で置き換えた式 (4.53) に他ならない. すなわち, $Z_i^*, i = 1, \dots, n-H$ は $\prod_{i=1}^{n-H} \eta(Z_i)$ に対する重根解または複号根解で与えられる.

以上のことから, $\Phi(z)$ を与える可能性のある Z_i の組合せには, 重根解と複号根解に加えて, 定義域の上限 $Z_i = \delta$ を含んだうえで, 残る Z_i を重根解もしくは複号根解で与えた組合せが含まれる.

謝辞

学部から博士後期課程に至るまでの研究活動にあたり，岡山大学自然科学研究科 有菌育生教授には幾度となく手厚いご指導を受け賜りました．深く感謝申し上げます．学位論文をまとめるにあたり，岡山大学環境生命自然科学研究科 田中俊二教授より多くの助言をいただきましたこと深く感謝申し上げます．研究遂行にあたり近畿大学理工学部 竹本康彦教授には多くの助言をいただきましたこと深く感謝申し上げます．岡山大学環境生命自然科学研究科 柳川佳也准教授ならびに Pukyong National University の Prof. Lee Jieun には研究に取り組むにあたりたくさんのお力添えを頂きましたこと，深く感謝いたします．また共に研究に勤しんだ生産知能学研究室の皆様にも心より感謝いたします．

本研究は JST 科学技術イノベーション創出に向けた大学フェローシップ創設事業 JPMJFS2128 の支援を受けたものです．博士後期課程在学中は OU-SPRING（旧 OU フェローシップ）に採択していただき，研究費や面談，スキルアップのための講習会など様々なサポートをいただきました．博士前期課程 2 回生，博士後期課程 1 回生および 3 回生の計 3 年間は公益財団法人大本育英会様から給与奨学金によるご支援をいただきました．また博士後期課程 1 回生と 2 回生の計 2 年間は岡山大学大学院様から大学院博士後期就学支援金のご支援をいただきました．これらのご支援のおかげで博士後期課程での活動を遂行できました．深く感謝申し上げます．

最後に，これまで私を温かく見守ってくれた家族に心から感謝いたします．