

氏 名 山 田 智 哉

学位（専攻分野） 博士(学術)

学 位 記 番 号 総研大甲第362号

学位授与の日付 平成11年3月24日

学位授与の要件 数物科学研究科 統計科学専攻

学位規則第4条第1項該当

学 位 論 文 題 目 近似主領域を考慮した修正情報量に基づく確率分布間の
定量的一様近似理論

論 文 審 査 委 員 主 査 教 授 江口 真透
教 授 平野 勝臣
教 授 松縄 規
助 教 授 栗木 哲
助 教 授 吉田 朋広（東京大学）

論文内容の要旨

本論文は Kullback-Leibler (K-L) 情報量が現代統計学の中で多用されているにも拘わらず、理論的に見てその使い方が必ずしも適切ではないとの認識から、その不適切さの克服を意図して行われた研究の成果である。研究の特徴は、比較の対象となる二つの確率分布(列)が正値を取る領域が必ずしも同一でない場合でも、通常 K-L 情報量の持つ発散の不都合を回避できる修正情報量を近似主領域の考えと共に導入することによって、これまで扱えなかった問題にも対処できる一様近似理論を展開していることにある。また近似誤差の定量的評価を行うために様々な有用な両側不等式も与えられている。

著者は K-L 情報量が現代統計学の中で多用されて来たにも拘わらず、理論的に見てその使い方に次のように検討を要する問題点が存在することを指摘している。第1の問題点は、従来原則として、比較の対象となる二つの確率分布が正値を取る領域が全く同一となる場合のみを考えていることである。このような場合に該当しない時は、解析的な都合からある種のコンベンションが置かれることが通常である。しかしこのような設定に納まらない状況は現実の近似問題で極く普通に起こり得る。そのような状況では K-L 情報量は発散し、近似の尺度としては全く意味をなさないことになる。第2の問題点として、これまでの多くの研究が、通常 K-L 情報量の定義に現われる尤度比 (or 情報密度比) の分母の取り方、従ってどちらの分布で平均を取るのかということの重要性に十分注意を払ってこなかったことを挙げている。多くの場合、計算の可能性や容易さを優先させるあまり、Boltzman(1887)のエントロピーや Kullback (1959) および大偏差確率の理論における情報量の意味と役割に整合しない議論展開がされている点も指摘されている。著者はこれらの問題点の克服のため、特に第1の点を中心にして研究に取り組み、分布間の一様近似の理論建設に貢献したいということが本研究の主な動機となっている。このような背景の下、著者は考察対象とする分布間の大域的近似問題に於て、特に興味を持つ領域を近似の主要な領域 (= 近似主領域) として設定し、そこに分布する確率測度とそこでの K-L 情報量に準じた隔たりの量 (= 修正情報量) の大きさを評価することにより、分布間の定量的な一様近似理論を展開することを目的として論文をまとめている。すなわち、本論文で展開している分布間の一様近似理論の骨子は、問題に適した近似主領域の構成とそこに分布する確率および修正情報量の大きさの評価である。そのため本論文では、これらの量の評価に様々な段階で必要となる有効な各種の不等式を新たに与えている。

なお本論文の具体的問題の設定の段階に於て、時には、確率分布列間の近似を考えている。この場合には、分布列間の隔たりには例えば標本サイズがパラメーターとして現われ、一様近似理論から近似誤差の定量的評価を含む漸近理論や極限理論がほぼ自動的に誘導できることにも注意している。

論文は四章から成っている。第1章では上でも触れた、研究の背景、従来の関連する研究の問題点が指摘され、それを基盤にした研究動機、論文全体の構成について紹介されている。

第2章では、修正 K-L 情報量と全変動の意味での分布間の一様距離に基づく近似の関係を明らかにし、それらと関連する良く用いられる分布間の隔たりの諸量とその修正した

ものを取り上げ、それらの間に成立するいくつかの有用な両側不等式を与えている。本章における議論は近似主領域のみならず一般の可測集合上での分布間の隔たりの評価に対しても成立するものとなっている。

第 3 章では、近似主領域の構成について議論されている。その結果に基づき、従来の方法では取り扱いが困難であった、パラメータが標本サイズの増加と共に変動する場合の、ガンマ分布やベータ分布の一樣正規近似、ordered Dirichlet 分布の多変量正規近似などについて、関連する確率分布列と具体的な近似主領域の列について、有限な標本サイズでも適用できる定量的な一樣近似が実行されている。

第 4 章では、近似主領域の概念を用いて多変量指数型分布族に関する一樣近似が研究されている。分布の正準パラメータが微小変化した時に修正情報量がどの程度揺動するかを評価している。その際、Barron and Sheu (1991) の複雑な接近法を避け、問題設定と情報量の定義の仕方も広く自然科学一般に共通する認識で捉え直し、第 2 章で得た一連の不等式を基に、修正 affinity, 修正 W-divergence を援用して、当該問題に対して見通しのよい系統的な方法で修正情報量の大きさを評価することに成功している。

上記四章に加え、付録として、本論文で提出するいくつかの理論不等式や修正情報量、近似主領域の確率の定量的評価に必要な基本的不等式とその証明を与えている。また本論文で考察した個別の近似問題についての各種計算の詳細も付録として与えられている。

論文の審査結果の要旨

- (1) 総合評価：本論文は Kullback-Leibler 情報量を含む分布間の隔たりを表す指標に関する研究を総合的に研究した。

特に、主領域の考えを導入して、上の指標の近似理論を提出した。

これは二つの確率分布が必ずしも同一の台を持たない場合においても一様な近似が可能になった。

また近似誤差の定量的評価を行うために様々な有用な不等式が考察されている。

総合的に判断して、本研究は出願者の統計基礎理論の分野における有意義な貢献を表しているものと評価できる。

- (2) 論文の概観

論文概要：提出された論文は四章から成っている。

第 1 章では、研究の背景、従来に関連する研究の問題点が指摘され、それを基盤にした研究動機、論文全体の構成について紹介されている。

第 2 章では、修正 Kullback-Leibler 情報量と全変動距離に基づく近似の関係を明らかにし、それらと関連する分布間の隔たりの諸量とその修正された量との関係を与えた。それに関するいくつかの有用な両側不等式を与えている。

第 3 章では、近似主領域の構成について議論されている。その結果に基づき、従来の方法では取り扱いが困難であった、パラメータが標本サイズの増加と共に変動する場合の、ガンマ分布やベータ分布の一樣正規近似、ordered Dirichlet 分布の多変量正規近似などについて、有限な標本サイズでも適用できる定量的な一樣近似が実行されている。

第 4 章では、近似主領域の概念を用いて多変量指数型分布族に関する一樣近似が研究されている。分布の正準パラメータが微小変化した時に修正情報量がどの程度変化するかを評価している。

第 2 章で得た一連の不等式を基に、修正 affinity, 修正 W-divergence を援用して、当問題に対して見通しのよい系統的な方法で修正情報量の大きさを評価することに成功している。

上記四章に加え、付録として、本論文で提出するいくつかの理論不等式や修正情報量、近似主領域の確率の定量的評価に必要な基本的不等式とその証明を与えている。また本論文で考察した個別の近似問題についての各種計算の詳細も付録として与えられている。

その他：

出願者は研究指導を受けた状況も良好で真面目に努力してきた。

本論文の一部分は統計の専門誌に刊行され、更に投稿中のものも内容的に positive な審査中間報告を得ており、修正原稿は accept の方向で進展している。