博士学位請求論文審査報告書

式見 拓仙

Large Deviations for Smoothed Empirical Distributions

「滑らかな経験分布関数に関する大偏差原理」

（１）論文の概要

経験分布関数に関する大偏差原理は多くの研究者によりこれまで研究されてきている。しかしこの分野に於ける大部分の先行研究が “経験分布関数” に関する大偏差原理を扱ってきたのに対し、本論文の特筆すべき点は “カーネル型経験密度関数” の場合について考えたことにある。この問題の背景には、統計学では多くの場合密度関数の存在が仮定されている事を上げることが出来る。本論文で扱っているカーネル型密度関数推定は現在、時系列分析等で広く用いられているものではあるが、大偏差原理等を証明する際には幾つかの技術的な困難を伴う。その中でも、標本数毎に Band 幅が異なることから独立・同一分布の系列に対する大偏差原理ではなく、独立な Triangular Array に対する大偏差原理を求めることとなり、旧来の方法をそのままの形で適応出来ないところに、この問題の確率論的な難しさがある。

式見氏は、まずはじめに、この問題をスタンダードな方法を精密化することにより解いたが、その後、全く別の方法で解決している。特に第２の方法は新しいもので、それ自体としても確率論的・数学的に興味深いものである。そして式見氏は本論文の中で、大偏差原理は用いられるカーネルに依存しないことを証明した。この結果はカーネル型密度関数推定に於ける一次の有効性は全てのカーネルに対し一定ということを意味しており、カーネル型推定が一般には合理的なものであることを示したといえよう。著者は、更に、Dvoretzky-Kiefer-Wolfowitz の不等式がカーネル型推定量に対しても成立することや、カーネル型ブーツトラップ密度関数推定量も大偏差原理に従うことを示している。これらの結果はブーツトラップ検定の正当性に対する理論的裏付けを与える。

さて、本論文は５章より成り、その構成は以下の通りである。

第１章　Introduction.
第２章　Analysis and topological preliminaries.
第３章　Properities of the Kullback-Leibler distance.
第４章　Large deviations for smoothed empirical distributions.
第５章　An alternative proof and several applications.

以下、本論の内容を章ごとに紹介する.

（２）本論文の内容

第１章（Introduction）では、はじめに本論文で取り扱うカーネル型経験分布関数 (smoothed empirical distribution function) を定義している：

$$\hat{F}^n(x) = \frac{1}{n}\sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right), \quad x \in \mathbb{R}^d$$

但し、$X_1, X_2, \ldots$ は $\mathbb{R}^d$ の確率ベクトル列、K は $\mathbb{R}^d$ 上の分布関数（カーネル）、また、h_n はバンド幅である（$h_n > 0,\ h_n \to 0,$ as $n \to \infty$ を仮定する）。第４章で旧来の方法による $\{\hat{F}_n\}$ の大偏差原理を証明するが、そこでは K はルベーグ測度に関して絶対連続で密度 κ を持つことを仮定している。従って、当面 、$\hat{F}^n$ はカーネル密度関数

$$f^n(x) = \frac{1}{nh_n^d}\sum_{j=1}^{n} \kappa\left(\frac{x - X_j}{h_n}\right)$$

をもつものとする。

第１．１節では通常の経験分布に関する大偏差原理（Sanov の定理）について概観している。ついで、第１．２節では一般の距離空間上での Donsker-Varadhan による大偏差原理およびそれと同等なラプラス原理の定式化を述べている。ここで、大偏差原理とは

「 距離空間 $\mathcal{X}$ に値を取る確率変数列 $\{X^n\}$ が次の性質を満たすとき、$\{X^n\}$ は $\mathcal{X}$ で rate 関数 I の大偏差原理を満たすという。

(a) 任意の閉集合 $F \subset \mathcal{X}$ に対し、

$$\limsup_{n\to\infty} \frac{1}{n}\log P(X^n \in F) \leqslant -\inf_{x\in F} I(x). \tag{1}$$

(b) 任意の開集合 $G \subset \mathcal{X}$ に対し、

$$\liminf_{n\to\infty} \frac{1}{n}\log P(X^n \in G) \geqslant -\inf_{x\in G} I(x). \tag{2}$$

但し、$I : \mathcal{X} \to [0, \infty]$ が各 $t < \infty$ に対して、$\{x \in \mathcal{X} : I(x) \leqslant t\}$ がコンパクトであるとき、I を rate 関数と呼ぶ」

のことである。一方、$\{X^n\}$ がラプラス原理を満たすとは

「任意の有界連続関数 $\varphi : \mathcal{X} \to \mathbb{R}$ に対し、

$$\lim_{n\to\infty} \frac{1}{n} \log E \exp(-n\varphi(X^n)) = -\inf_{x\in\mathcal{X}} [\varphi(x) + I(x)] \tag{3}$$

が成り立つ事をいう」。

第１．３節、第１．４節ではこれら２つの原理の同等性（同等性定理、縮約原理）が説明されている。但し、同等性に関しては、第５章にて smoothed bootstrap empirical distribution の大偏差原理を得るために、より一般化された両原理の同等性が証明されている。この部分は著者のオリジナルな貢献である。

第２章（Analysis and topological preliminaries）では主に位相解析でこれまで良く知られた結果のなかで本論文で必要なものをまとめている。特に第２．２節では後に必要となるごく初歩的な位相解析と確率論の結果をほぼ証明なしにあげている。一方、第２．３節では可分距離空間（$\mathcal{Y}$）上で定義された、$\sigma-$ 有限測度に関する密度関数全体からなる空間（Δ）の位相的・実解析性質を調べている。ここでは分布の弱位相に関して Δ が可分であることが証明されている。しかし、一般にはたとえ分布間の距離として Lévy-Prokhorov の距離を採用し、$\mathcal{Y}$ に完備性を仮定しても、Δ が完備であるとはいえない。この点は、確率測度全体のつくる空間とは異なっている。

第３章（Properties of Kullback-Leibler distance）では Kullback-Leibler 情報量の種々の基本的性質についてまとめてある。すなわち、$g \in \Delta$ の $f \in \Delta$ に関する情報量、

$$R(g|f) = \begin{cases} \int_{\mathcal{Y}} g \left(\log \frac{g}{f}\right) d\lambda & \text{if } g \ll f, \\ \infty & \text{otherwise.} \end{cases} \tag{4}$$

に関する変分公式、凸性、下半連続性、およびレベル集合 $\{g \in \Delta : R(g|f) \leqslant t\}$ のコンパクト性があげられている。これらはよく知られた結果であるが、最後のコンパクト性の証明に独自の工夫がみられる。

第４章（Large deviations for smoothed empirical distributions）と次章が本論文の主要部である。まず、「カーネルタイプの経験分布関数 $\hat{F}^n$ のラプラス原理（＝大偏差原理）は、カーネル密度推定量に対するラプラス原理と縮約原理より証明できる」ということより、この章では，カーネル密度推定量に対するラプラス原理の証明に多くの紙幅が割かれている。$\{\hat{F}^n\}$ に対するラプラス原理はその後に示されている。この章では、$X_0, X_1, \ldots$ は密度関数 f を持つ $\mathbb{R}^d$ 値確率ベクトル列であると仮定し、若干の

技術的理由より第４．１ 節であらためて、多変量の ($\mathbb{R}^d$ 上の) カーネル密度推定量を

$$f^n(x) = \frac{1}{nh_n^d}\sum_{j=0}^{n-1}\kappa\left(\frac{x-X_j}{h_n}\right),\ n \geqslant 1,\ x \in \mathbb{R}^d,$$

と定義している。f^n は $\mathbb{R}^d$ 上のルベーグ測度 λ に関する密度関数全体の空間 $\mathit{\Delta}_\lambda(\mathbb{R}^d)$ に値をとる確率変数として可測である。第４．２節では 大偏差確率に関する Hoeffding の不等式について言及している．第４．３ 節はごく初等的な stochastic control の解説である。そして、以下のカーネル密度関数のラプラス原理が第４．４節で証明されている。

[定理] カーネル密度推定量 f^n は $\mathit{\Delta}(\mathbb{R}^d)$ 上で大偏差（ラプラス）原理を満たし、その rate 関数は $R(\cdot\,|\,f)$ で与えられる。

証明は基本的には Dupuis and Ellis (1997) の方法の精密化といえる。さらに、第４．５節ではカーネル型経験分布関数 $\hat{F}^n(x) = \int_{-\infty}^{x} f^n(t)dt$ のラプラス原理（= 大偏差原理）が証明されている。ここで rate 関数は言うまでもなく Kullback-Leibler 情報量で表され、証明は縮約原理を適用し完了している。

この結果から派生する大偏差確率に関する若干の応用結果は、次の章でカーネル型経験分布関数に対する大偏差原理の別証明を考察した後に、述べられている。

第５章（An alternative proof and several applications）ではこれまでとは全く異なる新しい方法でカーネル型経験分布関数の大偏差原理が証明されている。前章の弱収束アプローチは大偏差問題一般に対して有効な方法であるが、本章では当該問題の特殊性を十分に考慮し $\mathbb{R}^d$ 上の分布全体からなる空間 $\mathcal{F}$ に以下で定義する距離を導入し、Sanov の定理を援用することによって証明している。

まず、第５．２節で、$\mathcal{F}$ の弱位相を距離付ける距離を、

$$\rho(G,H) \triangleq \sum_{k=1}^{\infty}\frac{1}{2^k}\sup_{t\in[-k,\,k]}|\,\phi_G(t)-\phi_H(t)\,|, \tag{5}$$

と定義している。但し、ϕ_G 　は $G \in \mathcal{F}$ の特性関数、 $G,\,H \in \mathcal{F}$ $[-k,\,k] = [-k,k]\times\cdots\times[-k,k] \subset \mathbb{R}^d$ である。さて、この距離によって、$\mathcal{F}$ はポーランド空間になり更に、$\{H_n\}$ を $0 \in \mathbb{R}^d$ に退化した分布 δ_0 に弱収束する分布列とすると、$\mathcal{F}$ 上の距離 ρ に関して

$$\sup_{G\in\mathrm{F}}\rho(G*H_n,\,G) \to 0 \tag{6}$$

が成り立つことが証明されている。但し $*$ は畳み込みを表す。これらの結果をもとに、第５．３節では Sanov の定理を含む次の結果が導かれている：

[定理]　$X_1, X_2, \ldots$ を分布 $F \in \mathcal{F}$ をもつ i.i.d 確率ベクトル、F^n を通常の経験分布関数とし、H_n を δ_0 に弱収束する分布列とすると、$\{F^n * H_n\}$ は $\mathcal{F}$ 上で rate 関数が $R(\cdot|F)$ の大偏差原理を満たす。

この定理は、(6)、大偏差原理とラプラス原理の同等性および Sanov の定理から導かれている。又、$H_n = \delta_0$ とおけば、Sanov の定理に帰着する。そして、任意のカーネル $K \in \mathcal{F}$ を持つカーネル型経験分布関数

$$\hat{F}^n(x) \triangleq \frac{1}{n}\sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right)$$

の大偏差原理は上記の結果に含まれている。

第５.４節では若干の応用を与えている。それらの項目を以下に列挙する。

A1. 通常の経験分布関数に対する Dvoretzky-Kiefer-Wolfowitz の不等式と類似の結果。

A2. カーネル型のブーツトラップ経験分布の大偏差原理。

A3. 密度推定論への応用。

（３）評価

冒頭にも述べたように大偏差原理は統計学・確率論に於いて古くから論じられている問題でこれまでにも数多くの論文が発表されてきている。最近では、無限次元抽象空間上の大偏差原理が主に確率論の研究者の間で議論されているが、式見氏の研究はより統計学的な問題意識からのもので彼らの研究とは一線を画しているといえる。

本論文の大きな功績は以下の２点にあるといえよう。第１は第４章の議論にみられるように、旧来の方法の精密化による大偏差原理の導出である。この方法による証明は高度な数学的スキルを必要とするが、残念ながら余り見通しが良い方法とはいいがたい。反面、非常に一般的であり、本論文で論じられた問題よりもより一般的な問題に対しても適用可能である。そして大偏差原理は用いられるカーネルとは独立に成立するという結果は、後で述べるように統計学的には若干問題を残すが、確率論的な見地からみれば十分価値のある結果である。氏の第２の貢献は第５章で、これまでとは全く異なる、氏独自の方法を用い大偏差原理を証明したことにある。ここでは、特性関数を用いた新しい距離を使い、大偏差原理を証明している。第４章の方法に比べると一般性には欠けるものの、本問題に関しては簡潔で非常に見通しの良い解決法を与えている。しかし、氏が提唱した新しい距離については、他の問題に対しても応用の可能性があるだろう。これからの氏のこの方面での研究に期待したいところである。

最後に、本論文における疑問点と今後の課題について触れておきたい。まず疑問点としては次の点があげられる。本論文の主要結果である、連続型経験分布関数の大偏

差原理は、用いられるカーネルに依存しないという結果は統計学的には新たな問題の出発点ということになる。例えば、氏が第５章で論じている幾つかの応用例について、統計学的にはどの様なカーネルを用いれば “最も効率的” な推定が可能かということが興味の的であろう。残念ながら本論文の結果からは、カーネルの違いによる推定量の効率性の差という所までは出てこない。実際、カーネル関数の選択問題を議論するためには大偏差原理の収束のスピードや、所謂漸近展開が必要になることが予想される。その意味で本研究はこの分野の研究に対する一つの出発点としての価値はあるが、決して最終的なものではない。収束スピードについての研究は式見氏のこれからの研究課題として、十分な時間を投入して行ってもらいたいというのが我々審査委員の一致した意見である。そして、大偏差原理への収束のスピードを評価することは、数学的にみても、これまでの方法で行うことは非常に難しく、それを議論しないといって本論文の価値が減ずるものではない。

（４）結論

審査員は、以上の評価と口述試験の結果に基づいて、式見拓仙氏に一橋大学博士（経済学）の学位を授与することが適当であると判断する。

２００１年2月１４日

論文審査員：　早川毅　田中勝人　桑名陽一　石村直之　高橋　一