

## 科学研究費助成事業 研究成果報告書

平成 26 年 6 月 16 日現在

機関番号：32641

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500357

研究課題名(和文)事後密度に基づいた経験ベイズ法への転換

研究課題名(英文)Reconstructing the empirical Bayes method through the use of the posterior density

研究代表者

柳本 武美 (YANAGIMOTO, Takemi)

中央大学・理工学部・その他

研究者番号：40000195

交付決定額(研究期間全体)：(直接経費) 3,200,000円、(間接経費) 960,000円

研究成果の概要(和文)：経験ベイズ法の現状を越えて、事後密度から導かれる統計量により事前分布を評価する研究を行った。これ迄に発展させてきたe-混合ベイズ予測子の性質を利用して、弱い情報しか含まない事前分布を含めて多様な事前分布の利用を企図した。情報量の大きさの程度は特定の位置に分布が集中していると表現される。そこで関数の凸性を厳密に定義して集中度の新しい半順序を厳密に定義した。

更にはベイズモデルの尤度の概念が新しい視点を与えた。尤度は基本量であるので、予測密度の定義において事後密度の代わりに事前密度を用いることにより自然に尤度が定義できることを指摘した。また、証拠の統合についても適応できることが分かってきた。

研究成果の概要(英文)：An unknown hyperparameter is contained in a prior density in the empirical Bayes method. Our primary aim is placed on attempting to evaluate a prior density in terms of a posterior density. We developed the empirical Bayes methods applicable to a wide variety of prior densities. Emphasis is placed on the flexible use of a prior density containing only limited amount of information. Amount of information contained in a prior density is represented through the heaviness of concentration of the density about a fixed point. For this purpose we rigidly define a novel notion of the heaviness of concentration.

Further, a likelihood of a Bayesian model based on a mixture of sampling density is introduced. It may be surprising that any formal definition of such a likelihood is not found at all in existing literature. At the final stage of the present research it becomes apparent that our approach is applicable to an important problem of combining evidences from different sources.

研究分野：情報学

科研費の分科・細目：統計科学

キーワード：Bayes 法 事後密度 退化型事前分布 凸関数 予測子 DIC

## 1. 研究開始当初の背景

事前密度仮定する Bayes 法はその仮定の妥当性が基本的な問題であった。その問題は決着がつかずもないうちに大きく展開したのは、1980 前後である。需要のサイドからは高次元母数の推測が喫緊の問題になったことである。一方で供給サイドでは高次元数値積分についての技法の発展があった。高次元モデルの必要性は、空間統計学と時系列解析における平滑化がその嚆矢であった。その有用性は瞬く間に広く認識され、頻度論者にも受け入れられた (Breslow, 1990 JASA; Efron 1990, Stat. Sci)。高次元積分のシミュレーションによる近似計算は尤も華やかな研究テーマとなった。

多様なモデルへの適用は古典的なベイズモデルの拡張として、超母数を含む経験ベイズモデルの利用を促した。その結果超母数の推定が必要となった。実際的な問題として無情事前密度の評価がある。元来は役に立つ程度の情報を含むことが前提になっていた体系であった。急速に複雑化するモデルへの対応に伴い生じた問題である。深い議論もなく直感的に周辺尤度が用いられていたが、その破綻は明瞭に見える。この方向は Kass ら (1995, JASA) が代表的である。その関心は統計的検定との対比に夢中でベイズ法の発展に向いていない。しかも残念なことにモデル選択と検定とは全く異なるものであって比較は意味がないのである。この中で予測子における Corcuera ら (1999, SJS) の指摘は問題点を明らかにしている。Spiegelhalter ら (2002, JRSS) は一つの試みとして DIC を提案しているが、目先の一般性を追いついて様々な問題点が生じている。

我々の視点は推定量ではなく、基本的に予測子を中心に据える。そして予測子を導出する標本密度の混合について、様々な意味での連続性に重点を置く。最大の問題はモデルに含まれる全情報量が推論の基盤であるから、得

られる情報から導かれる情報は連続的に変化するはずだからである。

## 2. 研究の目的

急速に拡大する経験ベイズ法を実際に即して理論的に支えることが目的である。事前密度に含まれる未知の超母数を評価することが主要な問題である。事前密度の評価は従来では主として周辺尤度が用いられてきた。この現状を超えて、事後密度から導かれる統計量により事前分布を評価する研究を行う。先に申請者が発展させてきた e-混合ベイズ予測子の性質が幅広く利用できることが見えてきている。そこで、Spiegelhalter ら (2002) による DIC を参考に、現状から適用領域の広い経験ベイズ法への転換を試みる。周辺尤度の欠陥が弱い事前分布の利用にあることから、ベイズ型交差検証法など関連した問題での展開を通じ、平滑化を含む回帰型モデルを中心に実際問題への適用の拡大を図る。

周辺尤度による事前分布の評価による重大な欠陥を明確に示せると見込んでいる。既にいくらかの結果を得ているからである。従って、比較的目標の達成が見込める研究課題である。より実際的には事後密度に基づいた推測が適した分野を確認することになる。この中で、e-混合 Bayes 予測子の視点からはごく自然にベイズ型交叉検証法が提案できる。しかも、DIC とはかなり性能が異なることが分かる。この検証法を具体的なモデルの下に理論的及びシミュレーション研究を行うことで一つの纏まった研究を達成することが目標である。

## 3. 研究の方法

研究の焦点として、弱いけれども確かに情報を含む事前分布の評価の問題に集中する。周辺尤度の利用は到底望めないが、実際の適用にとって重要であり理論的な困難を伴うか

らである。まず初めに e-混合ベイズ 予測子から導かれるベイズ型交叉検証統計量を参照規準として、DIC の改善を計る。より具体的には、平滑化を含む回帰型のモデルの下で周辺尤度に基づく規準などとの関連を調べる。手法の比較研究を実施するにはより深い理解が得られる解析的方法に加え、より実際的な状況の下での確実な知識を得るためのシミュレーション実験をも併用することにより見通しの良い結果を得る。基礎的な研究の一方で貢献が確実なテーマがベイズ型交叉検証法である。最近の Plummer (2008) の貢献は、頻度論とは異なる枠組みがあることを示した。例えば Wahba の一般化交叉検証法と大きく異なる。しかも我々の枠組みで見ると対数予測子が理論の骨子に据えられているので、その拡張と整理は容易であると見込んだ。

連携研究者の大西は、指数分布族を仮定しないで、ある意味でより実際的な位置母数族を仮定して指数分布族において成り立つ命題の適用拡大を計った。その際に二つのモデルに潜むはずの解析的な双対性を探る。その双対性はある意味で 2 重構造の双対性をなしていることが確認できた。その 2 重構造が本質的である。

本研究の基本的な枠組みは 初年度中に提示することを見込んだ。この枠組みを確実にした上で、関連手法との比較及び特定モデルでの性能評価と結果が示唆する評価量の修正に取り組む。枠組みの詳細で問題になるのは、弱い情報をもつ事前分布の利用の場合である。Lunn ら (2009) も認めるように研究の主戦場である。周辺尤度の利用は全く接近不能である。交叉検証法の特性を活かした DIC の修正を図る。提示した手法の性能を、平滑化法における階差次数の選択のような特定の問題に適用する。この場合の比較対照は階層的ベイズ法である。階層的ベイズ法は超母数に対する事前分布の研究が進んでい

ないが、概念的は本来の姿であるはずだからである。

研究の進展には海外の研究者との討論は必須である。幸い本研究の該当年度には、ベイズ統計と生物統計に関する国際研究集會が開催される。この機会を捉えて若い世代の強力研究者と共に積極的な交流を図る。また、こうした会合は我が国の研究者の関心が向けられる面でも好都合である。本研究課題に関しては我が国には近年強力な研究者集団が育ちつつあるので、関連研究者と相互に刺激し合うことにより一層の研究の進展を計るための環境が生まれる。

#### 4 . 研究成果

従前の研究の延長線上として、無情報事前分布を用いた手法でも鞍点性質が成り立つことを示した (文献 1)。鞍点性質は予測子が良い性質を示すための基本的であるから、幅広い事前密度の利用が指示される。この性質を満たす十分条件が必要になるのでこの条件を調べる必要がある。これをまとめた結果を文献 6) である。理論的でなく直感的に導入された無情報事前密度には望ましくないものがあることを指摘した。分布の極限として形式的に扱うだけではなくて、対応するベイズ法の連続性を論じた。多くの場合に情報量の大きさの程度は、ある特定の位置に分布が集中していることと表現される。そこで分布の集中度の新しい半順序を厳密に定義した。このためには関数の凸性を厳密に定義する必要があった。更にはベイズモデルの尤度の概念が新しい視点を与えることが分かった。ベイズモデルの尤度が文献上で定義されていないことはむしろ意外である。頻度論では尤度は基本量である。我々は、予測密度の定義において事後密度の代わりに事前密度を用いることにより自然に尤度が定義できることを指摘した。つまり、尤度の定義には標本密度の混合に依存することの指摘で

ある。この視点の生物統計学特に臨床試験計画の評価についても一定の結果を得た。再試験については文献 3) において、また比較試験の検定については文献 7) において纏めることができた。また、本研究の最終盤になって証拠の統合についての問題にここで用いた方法が適応できることが分かってきた。広範な事前密度の利用について着実な成果を示すことができた。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 7 件)

- 1) Ogura, T. and Yanagimoto, T., Powerful test of two proportions by assuming a registered prior density. Comm. In Statist., Comput. Simul. (To appear) (査読有).
- 2) Yanagimoto, T. and Ohnishi, T., Permissible boundary prior function as a virtually proper prior density Ann. Inst. Statist. Math. (To appear) (査読有)
- 3) 柳本武美, 事実の認定を支える証拠と公的な判断. 「法廷のための統計リテラシー」, 近代科学社(印刷中). (査読無)
- 4) Ohnishi, T. and Yanagimoto, T. Twofold structure of duality in Bayesian model averaging. J. Japan. Statist. Soc., 43, 29-55, (2013) (査読有).
- 5) Yanagimoto, T. and Kobayashi, C., Performance of a Bayesian predictor-based procedure of an additional trial

after a non- significant result. Lecture Notes in Engin. and Comput. Sci.

WCE2013, vol. 1 (S.I. Ao et al. Eds) 291-294, (2013) (査読有).

- 7) Yanagimoto, T. and Ohnishi, T., Saddlepoint condition on a predictor to reconfirm the need for the assumption of a prior distribution. J. Statist. Plann. Inf., 41, 1990-2000. (2011) (査読有)
  - 6) Yanagimoto, T. and Ohnishi, T. Partial order of concentration about a position for comparing Bayesian prior densities. Far East J. Theoret. Statist., 45, 111-132. (2013) (査読有).
- [学会発表](計 14 件)
- 1) 柳本武美・大西俊郎: ある混合の下でのベイスモデルの尤度. 第 13 回西東京統計研究会, March 2014, 中央大学.
  - 2) 柳本武美: 事前分布の選択とその受容: ビッグデータ・疎なデータへの対応. 科研費研究集会、2013年12月6日, 東北大学
  - 3) 小椋透・柳本武美: 信用水準検定の McNemar検定への適用. 日本計算機統計学会. 2013年11月16日, 崇城大学ホール・熊本市
  - 4) 大西俊郎・柳本武美: Bayes予測における尤度最大化とシャノン・エントロピー最大化の双対性. 2013年度統計関連学会連合大会 2013年9月8日, 大阪大学
  - 5) 柳本武美・大西俊郎: 弱い事前情報が繋ぐ情報なし/情報あり事前分布. 2013年度

統計関連学会連合大会 2013年9月8日, 大阪大学	saddlepoint equalities in model averaging and their implication, ims-APRM2012 July 3, 2012, Tsukuba.
6) Ohnishi, T. and Yanagimoto, T. : Dual roles of maximizing likelihood and Shannon entropy in Bayesian prediction. The 59 <sup>th</sup> WSC, August 29, 2013, HKCEC, Hong Kong.	13) Yanagimoto, T.: Permissive boundary prior function as a virtually proper prior density, ISBA, June 2012, Kyoto.
7) Yanagimoto, T. and Ohnishi, T. : Examining the Role of a Non- informative Prior Function Through Weakly Informative Prior Densities, The 59 <sup>th</sup> WSC, August 29, 2013, HKCEC, Hong Kong.	14) 柳本武美, 小椋透, 小林千鶴 臨床試験 登録が担保する事前分布の柔軟な仮定 日 本計量生物学会年会 平成 24 年 5 月 26 日 統計数理研究所
8) Yanagimoto, T. and Kobayashi, C. : Performance of a Bayesian Predictor- Based Procedure of an Additional Trial After a Non- Significant Result, WCE 2013, ICCSDE, July 05, 2013, ICL, London, UK	[ 図書 ] ( 計 0 件 )
9) 柳本武美: 検定統計量としての対立仮説 モデルのベイズ信用度 統計関連学会連 合大会,平成 24 年 9 月, 北海道大学	[ 産業財産権 ] 出願状況 ( 計 0 件 )
10) Yanagimoto, T.: Designing an additional trial after a non- significant result: a Bayesian predictor-based procedure, IBC, 28 August, 2012, Kobe KICC.	名称 : 発明者 : 権利者 : 種類 : 番号 : 出願年月日 : 国内外の別 :
11) Ogura, T. and Yanagimoto, T.: Flexible and Powerful Exact Test for Hypotheses of Two Proportions in the Era of Clinical Trial Registration, IBC, 27 August, 2012, Kobe KICC.	取得状況 ( 計 0 件 )
12) Ohnishi, T. and Yanagimoto, T. : Dual	名称 : 発明者 : 権利者 : 種類 : 番号 : 取得年月日 : 国内外の別 :  [ その他 ] ホームページ等 なし 6 . 研究組織

(1)研究代表者

柳本 武美 (Yanagimoto, Takemi)

中央大学・理工学部・客員教授

研究者番号： 40000195

(2)研究分担者

( )

研究者番号：

(3)連携研究者

大西 俊郎 (Ohnishi, Toshio)

九州大学・経済学研究科・准教授

研究者番号： 60353413