事後密度にのみ基づいた経験 Bayes 法の利点

中央大学　　柳本 武美
九州大学　　大西 俊郎

1. 序：現状の経験 Bayes 法において奇妙な事実は、超母数の推定に周辺尤度が使われることである。潜在構造分析とか混合モデルならともかく、事前分布と事後分布が推論の骨子にあるはずの Bayes 推論で、周辺尤度の果たす役割が大きいはずがない。そこで、経験 Bayes 法を再編成して周辺尤度を用いない手法を構築する研究を進めている。しかし、この試みは大きなテーマである。そこで、我々の研究の展望を概観した上で、ごく最近に取り組んでいる Bayes 型交叉法から見たこの試みについて議論する。

2. 事前分布の仮定：標本密度 $p(\boldsymbol{x}; \theta)$ に対して経験 Bayes 法では $\pi(\theta; \delta)$ のように母数 θ に事前分布を仮定したい。しかし、lasso とか spline 平滑化のモデルを見るとそうではない。$\theta = (\theta_1, \theta_2)$ と分解して θ_1 のみに事前分布を仮定して、θ_2 には事前分布を仮定しないで超母数として扱う。技術的な問題で事前分布を仮定することが困難になることは殆どない。実際、reference prior あるいは Jeffreys' prior など無情報事前分布 $\pi_N(\theta_2)$ に対して、これを台として、$\pi(\theta_2; \delta_2) = \exp\{-\delta_2 D(\theta_2, c_0)\}\pi_N(\theta_2)k(\delta_2)$ として、proper な事前分布が仮定できるからである。$\pi_N(\theta_2)$ は $\delta_2 = 0$ での極限と理解できる。もし、$\pi_N(\theta_2)$ を仮定して事後密度が仮定できて事後密度が δ_2 に関して連続であれば何らの問題はない。

周辺尤度を用いる方法の欠陥は無情報事前分布は殆ど利用できないことにある。勿論、周辺密度が定義できないような事前密度は仮定すべきでないとの主張はそれなりの根拠もあり得る。しかし、上記のような事前密度においてどの程度の δ_2 まで許容するかと言う解決困難な問題が生じる。周辺尤度を用いる方法では、$\delta_2 = 0$ とできない事実は、δ_2 が小さいような“ 弱い ”事前情報の仮定に敏感であることを示している。このことは、二つの周辺尤度の比の積分表現などからも確認できる。

それでも、Bayes factor など周辺尤度を用いる方法が広く支持されているのは、平滑化法とか Lasso モデルで推定量がそれなりの振る舞いをするからである。演者の一人は以前には周辺尤度の最大化を適用した経験があるが (例えば [10])、明瞭な違和感はなかった。しかし、論理的に考察する必要がある。

Bayes factor が用いられるもう一つの理由としてモデル選択の一致性がある。しかし、モデル選択の一致性は上にも述べた”弱い”事前分布を利用するときの現象である (例えば [9])。しかも、事後分布から導かれる credible region ともかけ離れた現象である。また、検出力を無視した議論である。そもそも、正しいモデルを正しく選ぼうとする正当化の論理は、直感的な論理であって実証的推論の論理としては有用でない。

3. e-混合予測子：

この問題を予測子の観点から調べるために、通常の m-混合ではなく双対的な e-混合型の予測子を扱う。より一般的な予測子群の両端点に位置する [1]. 予測子 $p_e(\boldsymbol{y}|\,\boldsymbol{x})$ は

$$p_e(\boldsymbol{y}|\,\boldsymbol{x}) \;=\; c(\boldsymbol{x})\exp[E\{\log p(\boldsymbol{y};\,\theta)\,|\,\pi(\theta\,|\,\boldsymbol{x})\}]$$

但し $c(x)$ は正規化定数として与えられる [8]。この予測子は、損失 $\mathrm{D}(p(\boldsymbol{y}\,|\,\boldsymbol{x}),\ p(\boldsymbol{y};\,\theta))$ とし

た Bayes リスクを最小にする意味で最適である [8]。更に、予測子として鞍点予測子に限ると最適性が成り立つ。

この予測子は θ を指数分布族の自然母数とすると plug-in predictor になる。また、Spiegelhalter ら [5] の有力な特定版が $-2\log p_e(\boldsymbol{x}|\,\boldsymbol{x}) + 4\log c(\boldsymbol{x})$ と表現される。この表現は補正項が、罰金項と言うよりもバイアス補正項であることを明瞭に示している。DIC を用いて入門統計学に現れる問題を、無情報事前密度と適当な事前密度と比較のための評価基準として用いると良好な結果が得られる。

4. 交叉検証法：

交叉検証の考え方はごく直感的に理解できる。そのために歴史的に見ても古くから試みられている。最近になって本格的に研究された例として、Stone [6], Geisser and Eddy [2], Vehtari and Lampinen [7] が挙げられる。元来は効用関数などを用いて議論されている。きめの細かい議論は少なく、Bayes 予測子が適用されるときは m-混合予測子が用いられる。予測子の視点から見ると、予測子 $p(\boldsymbol{y}|\,\boldsymbol{x})$ を評価するために、$\prod p_e(x_i|\,\boldsymbol{x}_{-i})$ を用いる。ここで $\boldsymbol{x}_{-i}$ は $\boldsymbol{x}$ の第 i 要素を除いた小ベクトルである。交叉検証法を適用するためには、標本密度に制約 $p(\boldsymbol{x};\theta) = \prod p(x_i; g_i(\theta))$ をおく。e-混合予測子の定義のように

$$p_e(y_i|\,\boldsymbol{x}_{-i}) = c(\boldsymbol{x}_{-i})\exp[\mathrm{E}\{\log p(y_i; g_i(\theta))|\,\pi(\theta|\,\boldsymbol{x}_{-i})\}]$$

とおいて、$p_c(\boldsymbol{y}|\,\boldsymbol{x}) = \prod p_e(y_i|\,\boldsymbol{x}_{-i})$ と定義できる。この定義はごく自然に感じられるが、e-混合予測子が研究されていないので文献上は見あたらない。

評価基準は $p_c(\boldsymbol{x}|\,\boldsymbol{x})$ から構成されるが、この構成は観察予測子 $p_e(\boldsymbol{x}|\,\boldsymbol{x})$ のバイアスを補正するという点で DIC の設計指針と共通している。実際多くのモデルで同じような振る舞いをする。しかし、母数の次元が高くなる場合には大きく異なる。Plummer [4] は DIC の欠陥を指摘して、plug-in 予測子による罰金項の補正を提案している。その提案は肯定的に評価されている [3]。

困難であるが興味深い一つの問題として、平滑化法に用いられる事前密度の階差の次数を評価するための基準の構築がある。この問題に対して形式的に適用できる既存の基準は、何れも旨く働かない。Bayes factor は適用が到底無理であり、DIC の性能も良くない。いわゆる飽和モデルの扱いが問題になる。その場合分散の推定値が 0 になるので、その困難を除く必要がある。交叉検証法では少なくともこの困難は排除できる。渡部ら [11] は一つの提案を行っているが、やや特殊な工夫がなされていおり、改善の余地がある。彼らが扱ったデータに CVC を適用した。

文献：[1] Corcuera, J.M. and Giummole F. (1999). *Scand. J. Statist.* 26, 265-279. [2] Geisser, S. and Eddy, W.F. (1979). *J. Am. Statist. Assoc.*, **74**, 153-160 [3] Lunn, D. *et. al.* (2009). *Statist. Med.*, **28**, 3049-3067. [4] Plummer, M. (2008). *Biostatisitcs*, **9**, 523-539. [5] Spiegelhalter, D.J. *et al.* (2002). *J. Roy. Statist. Soc.* 64, 583 - 639. [6] Stone, M. (1974). *J. Roy. Statist. Soc., B*, **36**, 111-147 [7] Vehtari, A. and Lampinen, J. (2002). *Neural Computation*, **14**, 2439-2468 [8] Yanagimoto, T. and Ohnishi, T. (2009). *J. Statist. Plann. Inf.* 139, 3064-3075. [9] Yanagimoto, T. and Ohnishi, T. (2009). *Japan. J. Statist. Assoc.* 39, 111-131. [10] Yanagimoto, T. and Yanagimoto, M. (1987). *Technometrics*, 29, 95-101. [11] 渡部大志ら (2010). 2010 年度統計関連学会連合大会報告集, 133.