ベイズ事後分布に含まれる頻度論的な情報の利用

伊庭 幸人 モデリング研究系 教授

ベイズ統計におけるばらつきと頻度論的な統計でのばらつきは定義上は全く異なるものだが，モデルが正則で母集団をよく表現すれば漸近的に一致することがvon-Mises Bernsteinの定理として古くから知られている．最近，これとはまったく別に、ベイズ事後分布からの MCMCサンプルに豊富な頻度論的な情報が含まれていて，事後共分散の形で取り出せることが認識されてきた．この関係は経験影響関数の事後共分散表現に基づくもので，モデルが正しくなくても成立する．

ベイズ感受率公式（local case sensitivity formula): Millar and Stewart (2007), Perez et al (2006)

$$\mathrm{IF}(\mathbf{E}_{\mathrm{pos}}[A],i) = \frac{\partial}{\partial \epsilon_i}\mathbf{E}^{\epsilon_i}_{\mathrm{pos}}[A(\theta)]\Big|_{\epsilon_i=0}$$
$$= \mathbf{Cov}_{\mathrm{pos}}[A(\theta),\log p(y_i|\theta)]$$

- 観測値を摂動したときの影響の表現として提案された
- 頻度論とベイズを結ぶ橋として機能（下で解説）
- 統計物理でいう線形応答公式の一種（観測値の重み=外場）

$\mathbf{E}_{\mathrm{pos}}[\ ]$, $\mathbf{Var}_{\mathrm{pos}}[\ ]$, $\mathbf{Cov}_{\mathrm{pos}}[\ ]$: 事後期待値，事後分散 / 共分散

$$\log p(y|\theta) = \sum_{i=1}^{N}\log p(y_i|\theta)$$

公式の証明は下の式を積分記号のもとで微分するだけ

$$\mathbf{E}^{\epsilon_m}_{\mathrm{pos}}[A(\theta)] = \frac{\int A(\theta)\exp(\epsilon_m \log p(y_m|\theta) + \log p(y|\theta) + \log p(\theta))\,d\theta}{\int \exp(\epsilon_m \log p(y_m|\theta) + \log p(y|\theta) + \log p(\theta))\,d\theta}$$

WAIC，PCICのGibbs損失版のベイズ感受率公式による解釈

$$\mathrm{WAIC}_2 = \left\{\sum_i \mathbf{E}_{\mathrm{pos}}[\log(y_i|\theta)] - \sum_i \mathbf{Var}_{\mathrm{pos}}[\log p(y_i|\theta)]\right\}$$

Watanabe (2010, 2018)　(-2)倍は省略

$$\sum_{i=1}^{N}\mathbf{E}^{-i}_{\mathrm{pos}}[\log p(y_i;\theta)]$$

交差検証（CV）から出発： $\mathbf{E}^{-i}_{\mathrm{pos}}[A] = \int A(\theta)p(\theta|y_1,\cdots,y_{i-1},y_{i+1},\cdots,y_N)d\theta$

$$\simeq \mathbf{E}_{\mathrm{pos}}[\log p(y;\theta)] - \frac{1}{N}\sum_{i=1}^{N}\mathrm{IF}(\mathbf{E}_{\mathrm{pos}}[\log p(y_i;\theta)],i)$$

まるまる1個抜くので$\epsilon_i = -1$（大胆に線形近似）

$$= \mathbf{E}_{\mathrm{pos}}[\log p(y;\theta)] - \sum_{i=1}^{N}\mathbf{Cov}_{\mathrm{pos}}[\log p(y_i;\theta),\log p(y_i;\theta)]$$

local sensitivity formulaを $A(\theta) = \log p(y_i|\theta)$として適用
$\mathbf{Cov}_{\mathrm{pos}}[A,A] = \mathbf{Var}_{\mathrm{pos}}[A]$

- 渡辺によるキュムラント展開を用いた導出と数学的に等価だが，補正項の直観的な意味がより明らかに

$$\mathrm{PCIC}_G = \left\{\sum_i \mathbf{E}_{\mathrm{pos}}[\nu(y_i|\theta)] - \sum_i \mathbf{Cov}_{\mathrm{pos}}[\nu(y_i|\theta),\log p(y_i|\theta)]\right\}$$

Iba and Yano (in preparation)

- 任意の損失関数に一般化すると事後分散が事後共分散になる

損失関数$\nu(y_{new}|\theta)$，とりあえず正則モデル限定
平均ギブス損失$\mathbf{E}_{y_{\mathrm{new}}}\left[\mathbf{E}_{\mathrm{pos}}[\nu(y_{new}|\theta)]\right]$を推定

頻度論的共分散を事後共分散で表示する公式 Giordano(2020): blog post, StanCon2020

$$\mathbf{Cov}_y[\mathbf{E}_{pos}[A],\mathbf{E}_{pos}[B]] = \frac{1}{N^2}\sum_i \mathrm{IF}(\mathbf{E}_{\mathrm{pos}}[A];i)\mathrm{IF}(\mathbf{E}_{\mathrm{pos}}[B];i) = \sum_i \mathbf{Cov}_{\mathrm{pos}}[A,\log p(y_i|\theta)]\mathbf{Cov}_{\mathrm{pos}}[B,\log p(y_i|\theta)]$$

- 統計量の頻度論的分布が多変量正規分布で漸近的に表現できるとき
  頻度論的分布の共分散$\mathbf{Cov}_y$が事後共分散で表示できる
  → 1次漸近論における頻度論的情報はすべて
  ベイズ事後分布から事後共分散の形で取り出せる

いま研究中の関連した内容
- 動的な推定 (early stopping) ：久保公式
- 高次のブートストラップ近似への応用

共同研究者・競争的資金

- 本研究は矢野恵佑氏との共同研究である
- 科研費 基盤(C) 多様な予測に対応した情報量規準の開発：計算統計的アプローチ　伊庭（代表）/矢野（分担）