望ましい無情報事前分布の規準

統計数理研究所　　柳本 武美
Takemi Yanagimoto
Institute of Statistical Mathematics

1. 序

ベイス法は基本的に既存の事前情報を有効に取り込む手法である。しかし、解析者が想定する情報を取り込んだ解析をすると、解析結果は利用した事前情報に依存する。従って、異なる情報が正しいと信じる人とは解析結果を共有出来ない。解析結果を広く共有したい場合には致命的な欠点である。実際今日では重要性が高まっている evidence-based medicine あるいは evidence-based policy-making の分野での適用が求められる。

この問題を解決する試みが、無情報事前分布あるいは客観的事前分布の利用である。文字通り無情報事前分布であれば、それを利用しても解析結果は共有出来るはずである。多くの無情報事前分布が提案されている。それではどのような無情報事前分布が望ましいかが問題になる。次節で改めて紹介するように望ましい事前分布に関する議論がなされてきた。

ところが、望ましい無情報事前分布について、事前分布の枠組みの中だけで議論すると、内在的な欠陥がある。純粋な意味での無情報事前分布に拘ると議論が進まない。注意深く検討すると厳密な意味では、判断あるいは価値が想定されていることが分かる。

実用的な推定量あるいは信頼区間/信用区間では、その良さを評価する基準は定着している。頻度論の立場からはごく自然な規準に見える。実際に望ましいとされる事前分布はその性能の良さで受け入れられているように見える。このことは、厳密な意味での無情報性に拘らないで、事前分布を導出される推測法の性能で評価する方が実際的であることを示唆してい

る。更には通常の informative な事前分布の利用と無情報事前分布の利用との障壁を低くする。本稿では、無情報事前分布の制約を弱めて、その良さを導出される推測方式の良さで評価することの妥当性を議論する。

2. 無情報事前分布の理解

モデルを確率密度の族 $\mathcal{P} = \{p(\boldsymbol{y}|\theta); \boldsymbol{y} \in \mathcal{X}(\subset R^n), \theta \in \Theta(\subset R^p)\}$ とし、大きさ n の標本ベクトルを $\boldsymbol{x}$ と書く。母数空間 Θ 上の事前関数 $\pi(\theta)$ は improper であることを許容するが、その事後密度は存在すると仮定する。本稿で議論されるベイズ法の基本的な知識は Bernardo and Smith (2000) 及び Robert (2001) から得られる。

最も素朴な無情報事前分布は一様分布 $\pi_U(\theta) \propto 1$ である。見かけ上は事前分布を意識しないでよい。また、この事前関数を仮定したときの事後モードは最尤推定量 (MLE) に一致する。しかし、この定義では母数の変数変換に依存する。また、MLE との関連も MLE 自身が、母数の次元が高いとか標本サイズが小さいときには、望ましい推定量と見なされないことが多いので魅力的な性質ではない。

Jeffreys prior は $\pi_J(\theta) \propto \sqrt{|\mathrm{I}|}$, 但し I は Fisher information matrix, で定義される。大きい長所は一様分布とは異なり、母数を変数変換しても分布の同等性は保存される。また、母数が一次元の場合には後述する reference prior と一致する。しかし、母数の次元が高いときにはその性能が悪いことが知られている。また、left-invariant Haar measure になってしまうことも知られている。。

より洗練された無情報事前分布の提案が Bernardo (1979) によりなされた。事前関数 $\pi_R(\theta)$ を事後密度と事前関数の Kullback-Leibler divergence の期待値 $\mathrm{E}\{\mathrm{D}\big(\pi(\theta|\boldsymbol{x}), \pi(\theta)\big); m(\boldsymbol{x})\}$ を最大化するように選ぶ。この要請は「事後分布と事前分布の平均距離を大きくする」と解釈されるから直感的に無理がない。そのために多くの研究者に支持されている。また、right-invariant Haar measure になる。更に、代表的な分布では reference prior から導かれる推定量が良い性能を示すことが知られている。

しかし、上の要請が自然である一方で、定義が厳密である分適用範囲に制約がある。具体的なモデルに対する研究では漸近的性質を適用するしか

ない。また、母数が多次元の場合にはどの母数に興味があるかが議論される。しかし、どの母数に興味があるかは理論的な都合によるよりもデータを取り巻く諸々の条件による。この不都合は既存の条件付き MLE にも見られる。この理論では、邪魔者母数が理論的に定められている。その結果、正規分布 $\mathrm{N}(\mu, \sigma^2)$ の推測では平均 μ が邪魔者とされる。しかし殆どの実用例では双方の母数が重要である。しかも敢えて重要でない母数を特定するとすれば当然 σ^2 である。

無情報事前分布を理解するためには、通常の informative 事前分布と比較することも必要である。定数 $c \in \Theta$ と母数 θ との (擬) 距離を $d(\cdot,\cdot)$ として、事前密度

$$\pi(\theta) \;=\; C\exp\{-\delta_0 d(\theta.c)\}b(\theta) \tag{1}$$

但し C は規格化定数、δ_0 は所与の正数、を考える。この定式化では、$b(\theta)$ が無情報事前関数と見なされる。実際、δ_0 が非常に小さいと、推論上は $b(\theta)$ を仮定した場合と殆ど変わらない。端点 $\delta_0 \doteq 0$ であれば $b(\theta)$ と同等になり、別の端点 $\delta_0 \to \infty$ では一点分布 $\delta_D(\theta - c)$ に収束する。この事前関数は、母数が c の近くに分布するであろうとの事前の知識を取り込んでいる。その結果、$\delta_0 \neq 0$ であれば、ベイズ推定量 (事後平均/事後モード) は c に依存する。この事実は、ベイズ推定量としては大きい欠点ではないが、推定結果を他者と共有する面では欠点である。

3. 無情報事前分布への懐疑

本節では望ましい無情報事前分布について、導出される推定量と離れて議論することの限界について考察する。

先ず複数の一見尤もらしい複数の事前関数を比較する。二つの無情報事前関数 $\pi_1(\theta), \pi_2(\theta)$ の比 $\pi_1(\theta)/\pi_2(\theta)$ から相対的な性質が読み取れる。どちらがより無情報であるかを議論することは難しい。

例 1. 標本分布が二項分布 $\mathrm{Bi}(n;p)$ である場合、文献上の無情報事前関数は $Cp^{a-1}(1-p)^{a-1}$ と書ける。特に $a=0$ とすると Haldane prior $\pi_H(p)$, $a=1/2$ とおくと reference prior, $a=1$ とおくと p 上の uniform prior となる。べき乗指数 a が大きいほど相対的に中央値 $p=1/2$ の集中して

いること示している。(1) 式の informtative prior で $b(p) = \pi_H(p)$ とおいて記述することができる。その意味では Haldane prior が最も純粋に無情報事前分布であるようにも見える。また、p を estimand としたときには Haldane prior の下での事後平均は MLE に一致する。一方では、uniform prior は直感的に無情報に見える。その下での事後モードは MLE に一致する。。Reference prior は、この場合 Jeffreys prior でもあるが、ごく自然な無情報事前分布である。

これらすべての事前分布はごく自然な無情報事前分布である。どの事前分布が最も厳密な意味での無情報であるか否かを論じることは困難である。しかも、事後平均では estimand の選択に従い区々の推定量が導びかれる。実際、estimand として p 以外を選ぶと、Haldane prior の下での事後平均は MLE と一致しない。一方で、推定量の性能の良さはリスクを比較することにより議論することができる。

望ましい無情報事前分布を深く考察すると隘路に陥ることは偶然ではない。何故なら科学的推論に係わる基本的問題と関係するからである。今日の科学哲学の成果として、素朴な論理実証主義の限界の指摘がある。論理実証主義は、端的の述べると

「純粋な観察に基づく科学的知見の蓄積」

と要約できる。物理学・地質学など科学の勃興期を指導した考え方である。数学基礎論では、得られた命題が絶対的に正しいことを主張する直観/古典主義に相当する。

この考え方は実際の科学的観察と大きな齟齬がある。観察には既存の知識と多くの前提は避けられないからである。また、観察から単純に科学的知見が導かれる訳でもない。対峙する考え方は

1) 観察に伴う仮定を (観察の前に) 明示する

2) 予め設けた命題の否定を重視する

を基本とする、一種の相対主義である。何も仮定しない科学から、仮定を明示する科学への転換がある。今日では、実験・試験に伴う手順は予め書類に定めておく必要性が強調されるようになっている。結果を得た過程が検証できることになり、データの質を向上させる。数学基礎論では、

Hilbert-Bernays が提唱した公理/形式主義である。今日では広く受け入れられている。得られた命題が、予め設けた公理が受け入れら得ること前提として結果の正しさを主張する。

本節の補足として、無情報事前分布と公平な分配との関連を指摘する。公平な分配は効率的な管理・運営と並んで経済学の重要な概念である。また社会学的に見ても基本的である。母数空間が有限個の点集合 $\Theta = \{\theta_1, \cdots, \theta_K\}$ で K が小さければ、等分することがごく自然である。事前確率を等確率分布として表現できる。最も公平でない分配は一人の独占であるが、(1) で δ_0 を無限大として表現できて一点分布になる。独占の概念は、母数空間の K が大きくなっても同様に表現できるが、公平な分配を等確率分布で表現するのは現実的ではない。無情報事前分布の研究は、公平な分配の研究ほどに難しい (柳本, 2022)。

4. 統計量としてのベイズ推定量

無情報事前分布の良さを、事前分布が無情報であることを追究して定めることには根本的な問題があることを主張した。この節では、対峙する定義を議論する。そのために、推定量とは母数の推定に用いられる統計量であることを確認する。統計量の定義は 4.2 小節で少しきめ細かく議論する。次小節では、ごく常識的に「標本の可測関数である」とする。

4.1 望ましい無情報な事前分布

ここでは、「ある事前分布は望ましいとは、導出される推定量の性能が良いことである」とする。先ず事前分布に導出されるベイズ推定量が統計量になることを要件とする。その上で、推定量の性能の良さで事前分布を評価する。性能の良さは容易に定義されるから、統計量であるか否かの判定が議論の焦点になる。なお、性能の良さを規準に推定量を線形に順序づけることはできないが、事前分布の線形な順序づけが困難であることを示していると解釈できる。

この定義の長所は、論点を明解にすると共に、新しい無情報事前分布の導入する際の障壁を下げることである。ベイズ推定量が推定量であるとの要請は、細かい点を除くと当然の要請である。ある事前分布が無情報であるか否かと言う曖昧な議論をすることは建設的でない。また、導出された

推定量の性能が悪ければ、概念的には望ましい無情報事前分布であっても魅力的でない。むしろ、気がつかない欠点が内蔵されていると疑われる。

通常のベイズ推定量は統計量ではない。典型的な事前密度 (1) を仮定した場合には、ベイズ推定量は統計量ではない。観測値でない値 c に依存するからである。従って、informative prior は排除できる。その結果、統計量として得られるベイズ推定値は基本的にすべての人と共有できる。

推定量の性能は損失関数を仮定した下でリスクにより評価される。損失関数を $L(\hat{\theta}, \theta)$ としたとき、リスクは標本分布の下での期待値 $\mathrm{E}\{L(\hat{\theta}, \theta); p(\boldsymbol{x}|\theta)\}$ で定義される。従って推定量の性能は損失関数と真の母数の値に依存する。推定量を評価する規準については文献上も強い批判はない。

4.2 統計量の詳細な理解

統計学 (statistics) では統計量 (statistic) は基本的な用語であるが、その定義は必ずも明瞭ではない。本稿ではこの用語を用いて望ましい無情報推定量を定義した。その定義を具体的に拡張することを目的に、統計量の定義を更にきめ細かく議論する。

標準理論では、統計量は「標本 $\boldsymbol{x}$ の可測な関数である」と定義/説明される (例えば Lehmann and Casella (1998, p16), Kotz *et. al.* (2006, p8006))。この定義では、統計量が観測値のみに依存することを強調している。Informative prior (1) に含まれる c とか未知母数 θ が含まれないことに主眼がおかれている。更に「可測な」関数であることを要請している。しかし、可測であることは実際のデータ解析では当然の要請である。統計理論を測度論に基づいて数学的に定式化することに注力した数理統計学の時代を反映しているに過ぎない。

重要なのは「関数」についての制約/範囲についての議論である。この点についての説明が無い。先ず、モデル $\mathcal{P}$ に依存することが許容されると考えられる。もっとも、モーメント推定量とか最小二乗推定量のようにモデルに依存しない推定量の重視する記述統計学の立場もある。ベイズ推定量は必然的に $\mathcal{P}$ と観測値 $\boldsymbol{x}$ に依存する。以降では、モデルに依存することは許容する。そうすると、直ちに

1) 「事前関数をモデル $\mathcal{P}$ から構成すると、そのベイズ推定量 (事後平均・事後モード) は統計量になる」

が得られる。例えば Jeffreys prior は $\mathcal{P}$ のみに依存するから、それから導出されるベイズ推定量になる。また MLE も同様である。

最初の拡張として関数が観測値に依存する場合を許容する。1) の前半部分を

2) 「事前関数をモデル $\mathcal{P}$ と観測値 $\boldsymbol{x}$ とから構成する」

と変更することにより拡張される。一見奇妙な感じがするかも知れないが、統計量が観測値に依存するから格段の拡張ではない。もし違和感があれとすれば、データを得た後で解析法を変更するように見えるからである。逆に言えば、そもそも解析法をデータの入手前に定めない現状の方が異常である。重要でもないデータを適当に解析する場合には、事前に解析法を定める必要はないかも知れない。しかし、観察を行う前に観測値の解析方法を予め定めておくことが原則で、その方が信頼できる解析結果が得られる。従って、データに依存した事前分布を仮定することには何らの問題はなく、導出されるベイズ推定量は統計量になる。また一見データの二重使用に見えるかも知れない。しかし、二重使用が問題になるのは、単に損失を小さくする故の批判である。推定量の性能が良ければ何ら問題はない。本研究会でも、田畑ら (2022) は観測値のランクに依存した事前分布から導出されるベイズ推定量を議論している。ランクに依存した事前関数を仮定しても統計量になり得る。

ところが上の拡張では、一様事前分布に基づいた事前分布から導出されるベイズ推定量は統計量にならない。何故なら一つのモデルは異なる母数を用いて $\mathcal{P} = \{p(\boldsymbol{x}|\theta)|\theta \in \Theta\} = \{p(\boldsymbol{x}|\xi)|\xi \in \Xi\}$ と表現できる。言い換えると、異なる母数化により異なるベイズ推定量が得られる。しかし、異なる母数化の一つを選択することは、統計量が排除しようとした観測以外の値とは異なる。別の拡張は

3) 「事前関数をモデル $\mathcal{P}$、観測値 $\boldsymbol{x}$ 及び選んだ母数 ξ から構成する」

となる。一様事前分布を仮定した下での事後モードは MLE に一致する。事後モードは母数化に依存しない。しかし、3) の拡張を許容しないと、事後モードは統計量であるが事後平均は統計量とは言えないことになってしまう。

Reference prior は既に述べたように代表的な無情報事前分布として受

け入れられている。ところが上の拡張での範疇に含まれない。母数の分割 $\theta = (\lambda, \psi)$ に依存するからである。加えて、一方の成分 ψ が他方の成分より関心の程度が高いとする。議論の余地はあり得るが

4)「事前関数をモデル $\mathcal{P}$、観測値 $\boldsymbol{x}$ と母数の分解 $\theta = (\lambda, \psi)$　から構成する」

とすることで許容することが実際的である。

更には条件付き MLE ではもう少し複雑になる。分解を要請すると共に ancillary 統計量 t、通常は λ の MLE、を用いて確率密度の分解条件

$$p(\boldsymbol{x}|\theta) \;=\; pm(t|,\lambda)\cdot pc(\boldsymbol{x}|t,\psi) \tag{2}$$

を要請する。次節では条件付き MLE に基づいた新しい事前分布の導入を議論する。その事前分布に合わせて条件を記述すると

5)「事前関数をモデル $\mathcal{P}$ と $pc(\boldsymbol{x}|t,\psi)$ から構成する」

となる。見かけ上 $\boldsymbol{x}$ は消えるが、$pc(\boldsymbol{x}|t,\psi)$ に現れる。

表中の PML, MPML priors は次節で議論する。更にモーメント推定量と informative prior を加えて表にすると次のようになる。

Table 1. Sources of information necessary to derive selected Bayesian and non-Bayesian estimators

	Bayesian est.	Sources	Non-Bayesian est.
1 -1)		$\boldsymbol{x}$	Moment E.
2 -1)	Jeffreys prior	$\boldsymbol{x}$, $\mathcal{P}$	MLE
2 -2)	Data-dependent prior	$\boldsymbol{x}$, $\mathcal{P}$	
2 -3)	Uniform prior	$\boldsymbol{x}$, $\mathcal{P}$, ξ	
2 -4)	Reference prior	$\boldsymbol{x}$, $\mathcal{P}$, (λ,ψ)	
2 -5)	MPML, PML priors	$\mathcal{P}$, $pc(\boldsymbol{x}|t,\psi)$	Conditional MLE
3 -1)	Informative prior	$\boldsymbol{x}$, $\mathcal{P}$, $\pi(\theta)$	(Penalized MLE)

表 1 では推定量を大まかに三分割している。1) は狭まい意味での統計量であり、モデル $\mathcal{P}$ には全く依存しない。逆に、3) は統計量ではない。その中間である 2) に挙げた 5 種類の推定方式が本稿での考察の対象であ

る。その内枝番 -1), -2) も殆ど議論の余地はない。MLE は推定量と考えられるからである。また、枝番 -3) も格別の主観性は認められない。

最も厳密な考察が求められるのは、reference prior の下でのベイズ推定量と条件付き MLE である。これらの推定量の導出では、母数ベクトルの成分を関心のある成分と邪魔者の成分との区別する研究者が殆どである。その用語には明らかに価値あるいは主観性が関与している。実際、条件付き MLE では邪魔者成分を忌避して推定しないとの習慣がある。しかし、そうした習慣に拘らなければ格別の主観性はない。良い推定量を導出するための方策と見なせばよい。従って、条件付き MLE も reference prior の下でのベイズ推定量も統計量と見なされる。

この表から観察される二点に注意する。一点は、データ依存事前分布は一様事前分布より必要な情報が少ない事実である。別の事実は広く受け入れられている reference prior が Jeffreys prior より必要な情報が多い事実である。後者については注意点がある。即ち、reference prior の本来の定義では母数ベクトルの成分の分解は必要ではない。実際の導出の過程で、漸近的評価と共に、成分の分解が必要になる。

頻度論の枠組みでは、より広義の意味で推定量の用語が用いられている。順序統計量の最適な重み付けなどである。更には明らかに推定量ではない例として、罰金項付き MLE がある。罰金は恣意的に設けられるから、得られる maximizer は統計量ではない。しかし、この場合にも文献上では MLE の名が付されている。

5. 動機になった事前分布

本稿での考察の動機は、条件付き MLE を改良するベイズ推定量の研究である。条件付き MLE は分解条件 (2) が成り立つときに、興味のある成分 ψ を $\hat{\psi}_{CML} = \mathrm{Argmax}\, pc(\boldsymbol{x}|t,\psi)$ で推定する。別の成分は $\hat{\lambda}_{ML} = \mathrm{Argmax}_{\delta}\, pm(t|\hat{\psi}_{CML},\lambda)$ で推定する。問題は $(\lambda_{ML},\hat{\psi}_{CML})$ より優越する事後平均を導出する事前分布を求めることである。

この目的のために先ず profile marginal likelihood $pm(t|\psi,\hat{\lambda}_{ML})$ に着目して

1) PML prior, $\pi_m(\psi) \propto 1/pm(t|\psi,\hat{\lambda}_{ML})$

を導入する。多くの例では、この事前関数の下での事後モードは $(\lambda_{ML}, \hat{\hat{\psi}}_{CML})$ となる。この事前関数に対応して

2) MPML prior, $\pi_M(\psi, \lambda) = \pi_m(\psi)\pi_J(\psi, \lambda)$

を導入する。

この場合のように新しい事前関数を導入すると無情報であるかを議論することは難しい。望ましい無情報事前分布かを議論することは尚更難しい。これら二つの事前関数は条件付き MLE の導出に用いられた量とか Jeffreys orior である。何らかの主観的情報が含まれるとは考えにくい。更に、下の例 2 で議論する正規分布の場合にも見られるように、MPML prior は reference prior と一致する例が多い。一致する場合には、MPML prior の良い性質の一つとして reference prior を特徴付ける性質を満たすことが挙げられる。言い換えると、無情報事前分布を特徴付ける性質から、良い性質の一つとして理論的に整理される。MPML prior の下での事後平均は Miyata (2011) 及び Yanagimoto and Ohnishi (2009) の結果を適用して条件付き MLE が改良できることが示唆される。

例 2. 正規分布 $\mathrm{N}(\mu, 1/\tau)$ の場合: 計算すると $\pi_m(\mu, \tau) \propto 1/\sqrt{\tau}, \pi_M(\mu, \tau) \propto 1/\tau$ が求められる。従って、$\pi_m(\mu, \tau) = \pi_J(\mu, \tau)$ であり、また $\pi_M(\mu, \tau) = \pi_R(\mu, \tau)$ が成り立つ。共に代表的な無情報事前分布と一致する。

事前関数 $\pi_m(\mu, \tau)$ は事後モードが $(\hat{\mu}_{Mod}, \hat{\tau}_{Mod}) = (\bar{x}, \hat{\tau}_{ML})$ が成り立つように設計した。右辺は MLE である。事後平均を求めるために estimand を自然母数 $(\mu\tau, \tau)$ とすると。$(\hat{\mu}_M, \hat{\tau}_M) = (\bar{x}, \hat{\tau}_{CML})$ が得られる。この例では二つの推定量が一致する。一致した推定量が MLE より性能が良いことは Yanagimoto and Anraku (1989) が論じている。

例 3. 下端が未知の指数分布 $\mathrm{E}(\lambda, 1/\psi)$ の場合: 指数分布の下端が未知の値 λ であり、平均が $\lambda + \psi$ である。大きさ n の標本ベクトルを $\boldsymbol{x}$ とする。λ の MLE は最小順序統計量 $x_{(1)}(= t)$ となる。これから $\pi_m(\psi, \lambda) \propto 1/\psi$ が導かれる。形式的 Jeffreys prior 計算すると $\pi_J(\psi, \lambda) \propto 1$ となるから $\pi_m(\lambda, \psi) = \pi_M(\lambda, \psi)$ とおける。

PML prior $\pi_m(\lambda, \psi)$ の下での事後モード $(\hat{\lambda}_m, \hat{\psi}_m)$ は $(t, \hat{\psi}_{CML})$ となる。ここで $\hat{\psi}_{CML}$ は t を与えたときの ψ の条件付き MLE である。

Estimand を $(\lambda\psi, \psi)$ として MPML prior $\pi_M(\lambda, \psi)$ の下での事後平均 $(\hat{\lambda}_M, \hat{\psi}_M)$ を求めると $(nt/(n-1) - \bar{x}/(n-1), \hat{\psi}_{CML})$ となる。この推定量は Cohen and Helm (1972) に見える。母数 $(\lambda, 1/\psi)$ の不偏推定量として導入された。我々は魅力的な事前分布と自然母数に着想して形式的に同じ推定量を導いた。

直感的には ψ の不偏推定量が導かれるように感じるが、実際は $1/\psi$ の不偏推定量が得られた。偶然のように感じるかも知れないが、正規分布の場合も同じように $1/\sigma^2$ の事後平均を求めることにより σ^2 の不偏推定量が導かれる。上の導出を見ると estimand として (λ, ψ) とする方が母数の形も簡単に見える。しかし、$n = 2$ の場合に の事後平均が存在しない。あるいは特殊な病理例と思うかも知れない。しかし、同じ現象は他の分布でも起こる。

性能の良さを解析的に示すことは難しい。しかし、数値計算あるいはシミュレーションで性能の良さを調べることは可能である。

6. 議論

本稿の目的は、望ましい無情報事前分布を事前分布の枠内で議論することが生産的ないことを示すことであった。統計量というよりより厳密に定義できる枠組みでの長所を主張した。この主張は、informative prior の利用にもインパクトを与えることができる。ベイズ統計量は元来解析者が信じるあるいは信用する情報を有効にに利用する手法である。そうした informative prior の利用を阻む最も大きい要因が、解析結果をその情報を信用しない第三者との共有ができないことである。ところが informative prior の下でも Bayes 推定量が MLE に優越することがある。正規分布 N($\boldsymbol{\mu}$, I) の平均ベクトルの推定がその一例である。Informative prior を利用すると MLE に対する改善量は主観的情報に依存する。それでも母数の真の値に拘わらずに一様に改善する。従って、事前分布が informative であることは、結果を共有することを阻まない。

更に踏み込んで考察すれば、例え informative な事前関数であっても、多くの人が共有できる場合がある。想定される母数の範囲が特定できる場合である。その例は平滑化問題が挙げられる。平滑化 prior を利用する

と MLE を含めて多くの推定量より、真の傾向が滑らかであれば、良い性能を示す。経験的にも多くの系列データが滑らかに変動しているとの仮定は、共有できると思われる。

付記： 5 節の議論は宮田庸一博士 (高崎経済大学) 及び小椋透博士 (三重大学病院) との共同研究に負う。

文献

Bernardo, J.M. (1979). Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society. Series B*, **41**, 113-147.

Bernardo, J.M., Smith, A.F.M., (2000). *Bayesian Theory.* Wiley: Chichester.

Cohen, A. C., Helm, F. R. (1973). Estimation in the exponential distribution. *Technometrics*, **15**, 415-418.

Garvan, C.W. and Ghosh, M. (1997). Noninformative priors for dispersion models. *Biometrika* **84**. 976-982.

Kotz, S., Balakrishnan, N., Read, C. B. and Vidakovic, B. (2005). *Encyclopedia of Statistical Sciences*, John Wiley Sons.New York.

Lehmann and Casella , (1998) "Theory of Point Estimation", 2nd ed. Wiley, New York.

Miyata, Y. (2011). Fully exponential Laplace approximations using asymptotic modes. J. Am. Statist. Assoc., 99:468, 1037-1049,

Robert, C.P. (2001). *The Bayesian Choice (2nd ed.)* Springer, New York.

Yanagimoto, T. and Anraku, K.: Possible superiority of the conditional MLE over the unconditional MLE. *Ann. Inst. Statist. Math.*, 41 (2), 269-278, 1989.

Yanagimoto, T. and Ohnishi (2009). Bayesian prediction of a density function in terms of e-mixture." *J. Stat. Plann. Inf.*, **139**, 3064-3075.

田畑耕治, 岸村　遼, 柳本武美 (2022) 順位情報に基づく事前分布を用いた多項確率の推定、本研究会講究録在中

柳本武美 (2022).　無情報事前分布の理解：公平な分配と未知母数から 科研費シンポジウム ”Applications of Data Science in Social Science”で報告