

Priors for second-order unbiased Bayes estimators

Mana Sakai^{1,2} Takeru Matsuda^{1,3} Tatsuya Kubokawa¹

¹The University of Tokyo ²RIKEN AIP ³RIKEN CBS



Background: The Challenge of Bias in Bayes Estimation

- Bayes estimators generally have $O(n^{-1})$ bias with n sample size
- Asymptotically Unbiased Priors** (Hartigan, 1965) yield **second-order unbiased** Bayes estimators, i.e. **bias** = $o(n^{-1})$
- Limitation**: Hartigan's work **assumes i.i.d.** data, which excludes many models (e.g. Bayesian regression analysis)

Hartigan. (1965) The asymptotically unbiased prior distribution.

Purpose: Extend the theory of second-order unbiased Bayes estimators to **non-i.i.d. settings**

Setup & Assumptions

- $(X_1, \dots, X_n) \sim f_n(x_1, \dots, x_n | \theta)$
- $\theta = (\theta_1, \dots, \theta_p) \in \Theta \subset \mathbb{R}^p$
- $\ell_n(\theta) = \log f_n(x_1, \dots, x_n | \theta)$
- $\pi(\theta)$: prior density of θ
- $\hat{\theta}^B = \frac{\int_{\Theta} \theta \pi(\theta) \exp(\ell_n(\theta)) d\theta}{\int_{\Theta} \pi(\theta) \exp(\ell_n(\theta)) d\theta}$: Bayes est.

Regularity conditions

- $\hat{\theta}^{ML} - \theta = O_p(n^{-1/2})$, where $\hat{\theta}^{ML}$ is the MLE
- $\ell_n(\theta) = O_p(n)$
- $\ell_n(\theta)$ is three times continuously differentiable
- $\pi(\theta)$ is differentiable
- $H_n(\theta) = -\frac{1}{n} \frac{\partial^2}{\partial \theta \partial \theta^T} \ell_n(\theta)$ is invertible
- $H_n(\hat{\theta}^{ML})$ is positive definite

Assumptions on the log likelihood

- $H_n(\theta) = H(\theta) + O_p(n^{-1/2})$, and $H(\theta)$ is invertible
- $\frac{1}{n} \frac{\partial^3}{\partial \theta_r \partial \theta_s \partial \theta_t} \ell_n(\theta) = K_{rst}(\theta) + o_p(1)$
- $\mathbb{E} \left[\frac{1}{n} \left(\frac{\partial \ell_n(\theta)}{\partial \theta_r} \right) \left(\frac{\partial \ell_n(\theta)}{\partial \theta_s} \right) \right] = I(\theta) + o(1)$
- $\mathbb{E} \left[\frac{1}{n} \left(\frac{\partial^2}{\partial \theta_r \partial \theta_s} \ell_n(\theta) \right) \left(\frac{\partial}{\partial \theta_t} \ell_n(\theta) \right) \right] = J_{rs,t}(\theta) + o(1)$

Contribution 1: Characterization of asymptotically unbiased priors for non-i.i.d. models

The bias of the Bayes estimator

$$\mathbb{E}(\hat{\theta}^B - \theta) = \frac{\frac{1}{n} H^{-1}(\theta) \left(\frac{\partial}{\partial \theta} \log \pi(\theta) - \phi(\theta) \right)}{O(n^{-1}) \text{ bias}} + o\left(\frac{1}{n}\right),$$

$$\phi(\theta) := -\frac{1}{2} \sum_{r=1}^p \sum_{s=1}^p H^{rs}(\theta) A_{rs}(\theta), \quad A_{rs}(\theta) = \begin{bmatrix} K_{1rs}(\theta) + 2J_{1r,s}(\theta) \\ \vdots \\ K_{prs}(\theta) + 2J_{pr,s}(\theta) \end{bmatrix} + \sum_{t=1}^p \sum_{u=1}^p H^{tu}(\theta) I_{su}(\theta) \begin{bmatrix} K_{rt1}(\theta) \\ \vdots \\ K_{rtp}(\theta) \end{bmatrix}$$



$$\frac{\partial}{\partial \theta} \log \pi(\theta) = \phi(\theta)$$

$\Rightarrow \hat{\theta}^B$ is second-order unbiased

Contribution 2: Systematic construction of an asymptotically unbiased prior

Check the existence

$\exists \pi(\theta)$ s.t. $\frac{\partial}{\partial \theta} \log \pi(\theta) = \phi(\theta)$

$$\Leftrightarrow \frac{\partial}{\partial \theta_u} \phi_t(\theta) = \frac{\partial}{\partial \theta_t} \phi_u(\theta)$$

Define functions

Fix arbitrary $(c_1, \dots, c_p) \in \Theta$ and define

$$\psi_t(\theta_t, \dots, \theta_p) := \phi_t(c_1, \dots, c_{t-1}, \theta_t, \dots, \theta_p)$$

Integration

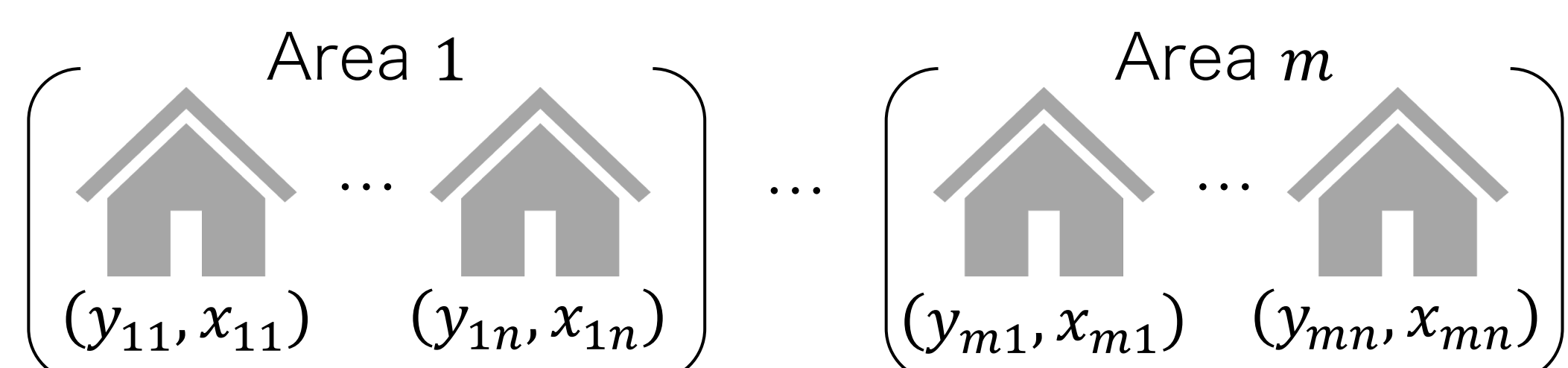
Take the prior as

$$\pi(\theta) \propto \exp \left(\sum_{t=1}^p \int_{c_t}^{\theta_t} \psi_t(z, \theta_{t+1}, \dots, \theta_p) dz \right)$$

Contribution 3: New prior for the random effects model

Model

$$y_{ij} = x_{ij}^T \beta + v_i + \epsilon_{ij} \quad (i = 1, \dots, m, j = 1, \dots, n)$$



- $v_i \sim N(0, \tau^2)$, $\epsilon_{i1}, \dots, \epsilon_{in} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$, $v_i \perp \epsilon_{ij}$
- $\theta = (\theta_1, \dots, \theta_p, \theta_{p+1}, \theta_{p+2}) = (\beta, \tau^2, \sigma^2) \in \mathbb{R}^{p+2}$
- m : number of areas, $m \rightarrow \infty$
- n : number of units in an area, fixed
- Observe $y_{ij} \in \mathbb{R}$, $x_{ij} \in \mathbb{R}^p$, independently for each i

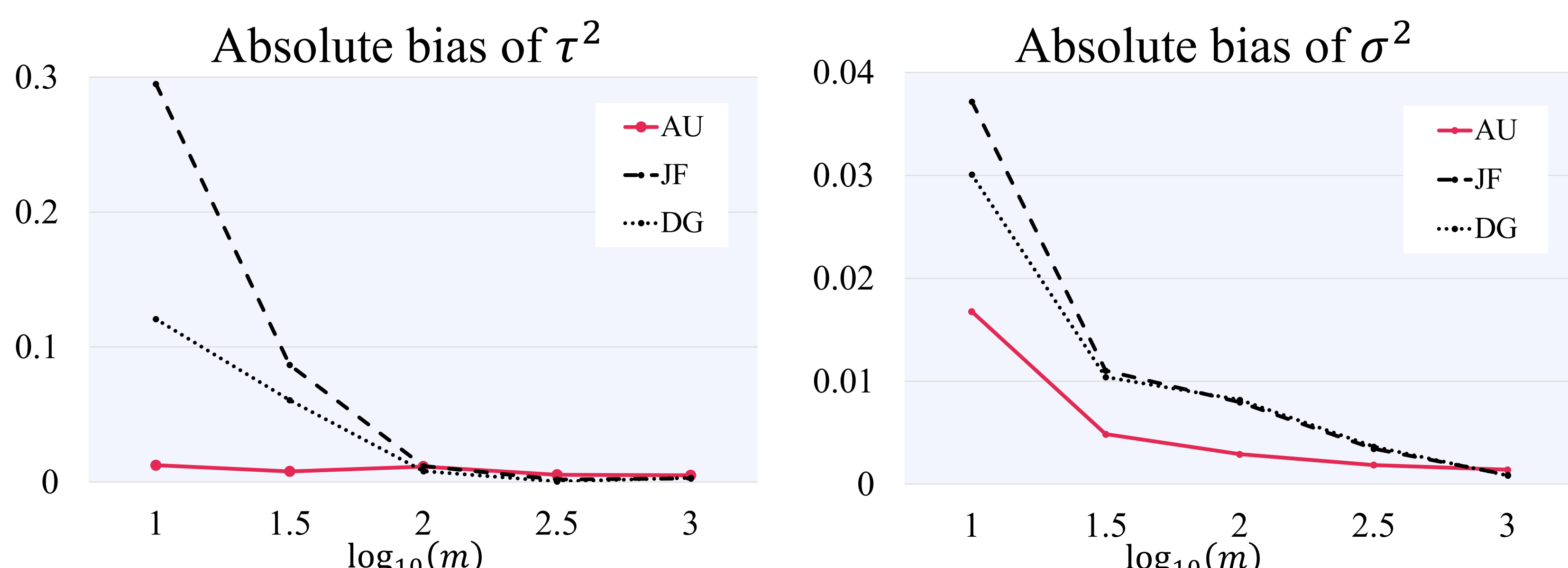
Asymptotically unbiased prior

$$\pi(\theta) \propto [\sigma^2(\sigma^2 + n\tau^2)]^{-2}$$

Simulation

AU	Asymptotically unbiased prior	$\pi(\theta) \propto [\sigma^2(\sigma^2 + n\tau^2)]^{-2}$
JF	Jeffreys' prior for (τ^2, σ^2)	$\pi(\theta) \propto [\sigma^2(\sigma^2 + n\tau^2)]^{-1}$
DG	Prior of Datta & Ghosh (1991)	$\pi(\beta) \propto 1$, $\tau^2 \sim IG(a_\tau, b_\tau)$, $\sigma^2 \sim IG(a_\sigma, b_\sigma)$

Datta & Ghosh. (1991). Bayesian prediction in linear models: Applications to small area estimation.



* The bias of β remains small and stable across priors and sample sizes