A BAYESIAN APPROACH TO CONSTRUCTING MULTIPLE CONFIDENCE INTERVALS OF SELECTED PARAMETERS WITH SPARSE SIGNALS

Zhigen Zhao and Sanat K. Sarkar

Department of Statistics, Temple University, Philadelphia, PA 19122, USA

Supplementary Material

S1 Technical Proof

Proof of Theorem 2.1. It is easy to see that

$$\begin{aligned} & P\left(\beta_i \in CI_i|\boldsymbol{Y}\right) \\ = \;& P\left(\beta_i \in CI_i|\boldsymbol{Y}, \beta_i = 0\right) P\left(\beta_i = 0|\boldsymbol{Y}\right) + P\left(\beta_i \in CI_i|\boldsymbol{Y}, \beta_i \neq 0\right) P\left(\beta_i \neq 0|\boldsymbol{Y}\right). \end{aligned}$$

If $0 \notin CI_i$, then

$$\begin{aligned} P\left(\beta_i \in CI_i|\boldsymbol{Y}\right) &= P\left(\beta_i \in CI_i|\boldsymbol{Y}, \beta_i \neq 0\right) P\left(\beta_i \neq 0|\boldsymbol{Y}\right) \\ &= P\left(\beta_i \in CI_i|\boldsymbol{Y}, \beta_i \neq 0\right)\left(1 - fdr_i\left(\boldsymbol{Y}\right)\right) < 1-\alpha, \end{aligned}$$

which leads to a contradiction. Consequently, $0 \in CI_i$. □

Proof of Theorem 3.1. First,

$$\begin{aligned} & EL_i(\beta_i, CI_i|\boldsymbol{Y}) = k_1^i Len(CI_i) P(\beta_i \neq 0|\boldsymbol{Y}) \\ - \;& P(\beta_i \in CI_i, \beta_i \neq 0|\boldsymbol{Y}) + I(0 \in CI_i|\boldsymbol{Y})(k_2 - fdr_i(\boldsymbol{Y})). \end{aligned} \tag{S1.1}$$

Note that

$$P(\beta_i \in CI_i, \beta_i \neq 0|\boldsymbol{Y}) = P(\beta_i \in CI_i|\boldsymbol{Y}, \beta_i \neq 0)(1 - fdr_i(\boldsymbol{Y})),$$

where

$$P(\beta_i \in CI_i|\boldsymbol{Y}, \beta_i \neq 0) = \int_{CI_i} \psi(\beta_i|\boldsymbol{Y}, \beta_i \neq 0) d\beta_i.$$

Write $Len(CI_i)$ as $\int_{CI_i} 1 d\beta_i$. Then (S1.1) equals to

$$\left(1 - fdr_i\left(\boldsymbol{Y}\right)\right) \int_{CI_i} \left(k_1^i - \psi\left(\beta_i|\boldsymbol{Y}, \beta_i \neq 0\right)\right) d\beta_i + I\left(0 \in CI_i|\boldsymbol{Y}\right)\left(k_2 - fdr_i\left(\boldsymbol{Y}\right)\right). \tag{S1.2}$$

The minimizer of the first integration is given by $\left\{\beta_i : k_1^i < \psi(\beta_i|\boldsymbol{Y}, \beta_i \neq 0)\right\}$. Now consider two intervals CI_i^1 and CI_i^2 where $CI_i^1 = \left\{\beta_i : k_1^i < \psi(\beta_i|\boldsymbol{Y}, \beta_i \neq 0)\right\} \setminus \{0\}$ and $CI_i^2 = \left\{\beta_i : k_1^i < \psi(\beta_i|\boldsymbol{Y}, \beta_i \neq 0)\right\} \cup \{0\}$. Then both CI_i^1 and CI_i^2 minimize the first integration of (S1.2). Since $0 \in CI_i^2$ and $0 \notin CI_i^1$, then

$$EL_i(CI_i^2|\boldsymbol{Y}) = EL_i(CI_i^1|\boldsymbol{Y}) + (k_2 - fdr_i(\boldsymbol{Y})).$$

Consequently, the Bayes interval includes 0 if and only if $k_2 \leq fdr_i(\boldsymbol{Y})$, i.e. it is the one that is defined in (7). □

Proof of Theorem 3.2.
Consider the posterior non-coverage probability $P(\beta_i \notin CI_i^{BD}|\boldsymbol{Y})$,

$$\begin{aligned}
& P(\beta_i \notin CI_i^{BD}|\boldsymbol{Y}) \\
= \ & P(\beta_i \notin CI_i^{BD}|\boldsymbol{Y}, \beta_i = 0)P(\beta_i = 0|\boldsymbol{Y}) + P(\beta_i \notin CI_i^{BD}|\boldsymbol{Y}, \beta_i \neq 0)P(\beta_i \neq 0|\boldsymbol{Y}) \\
= \ & fdr_i(\boldsymbol{Y})P(\beta_i \notin CI_i^{BD}|\boldsymbol{Y}, \beta_i = 0) + (1 - fdr_i(\boldsymbol{Y}))P(\beta_i \notin CI_i^{BD}|\boldsymbol{Y}, \beta_i \neq 0) \\
\leq \ & fdr_i(\boldsymbol{Y})I(fdr_i(\boldsymbol{Y}) < k_2) + \alpha(1 - fdr_i(\boldsymbol{Y})) \\
= \ & \alpha + fdr_i(\boldsymbol{Y})(I(fdr_i(\boldsymbol{Y}) < k_2) - \alpha).
\end{aligned}$$

Consequently,

$$\begin{aligned}
PFCR &= \frac{1}{R \vee 1} \sum_{i \in \mathcal{R}(\boldsymbol{Y})} P\left(\beta_i \notin CI_i^M|\boldsymbol{Y}\right) \\
&\leq \alpha + \frac{1}{R \vee 1} \sum_{i \in \mathcal{R}(\boldsymbol{Y})} fdr_i(\boldsymbol{Y})\left(I(fdr_i(\boldsymbol{Y}) < k_2) - \alpha\right) \leq \alpha,
\end{aligned}$$

and so $BFCR = E(PFCR) \leq \alpha$. □

S2 Gibbs Sampling

Define the parenthesis operator (i) as

$$\boldsymbol{\beta}_{(i)} = (\beta_1, \cdots, \beta_{i-1}, \beta_{i+1}, \cdots, \beta_p),$$

and

$$\boldsymbol{X}_{(i)} = (\boldsymbol{X}_1, \cdots, \boldsymbol{X}_{i-1}, \boldsymbol{X}_{i+1}, \cdots, \boldsymbol{X}_p)$$

where $\boldsymbol{X}_i$ is the ith column of the matrix $\boldsymbol{X}$. Let $\boldsymbol{D}_\tau = diag(\tau_1^2, \tau_2^2, \cdots, \tau_p^2)$.

1.

$$\begin{aligned}
& l_i = P(\beta_i = 0|rest) \\
= \ & \frac{\pi_0}{\pi_0 + \pi_1 \sqrt{\frac{1}{1+\tau_i^2 \boldsymbol{X}_i^T \boldsymbol{X}_i}} \exp\left(\frac{\left[(\boldsymbol{Y} - \boldsymbol{X}_{(i)}\beta_{(i)})^T \boldsymbol{X}_i\right]^2}{2(\boldsymbol{X}_i^T \boldsymbol{X}_i + \frac{1}{\tau_i^2})\sigma^2}\right)};
\end{aligned}$$

2.

$$\beta_i|Rest \sim l_i 1(\beta_i = 0) + (1 - l_i) N \left(\frac{\left(\boldsymbol{Y} - \boldsymbol{X}_{(i)}\beta_{(i)}\right)^T \boldsymbol{X}_i}{\boldsymbol{X}_i^T \boldsymbol{X}_i + \frac{1}{\tau_i^2}}, \frac{\sigma^2}{\boldsymbol{X}_i^T \boldsymbol{X}_i + \frac{1}{\tau_i^2}} \right);$$

3. Let $Z_i = 1\,(\beta_i = 0)$, then

$$\pi_0|rest \sim Beta\left(k\eta + \sum_i Z_i, k(1-\eta) + p - \sum_i Z_i\right);$$

4.

$$\tau_i^2|rest \sim \begin{cases} Exp(\lambda^2/2), & \text{if} \quad \beta_i = 0 \\ \left(INGaussian\left(\sqrt{\frac{\lambda^2\sigma^2}{\beta_i^2}}, \lambda^2\right)\right)^{-1}, & \text{if} \quad \beta_i \neq 0. \end{cases}$$

Here, the $inverseGaussian(\mu', \lambda')$ has the density function of

$$f(x) = \sqrt{\frac{\lambda'}{2\pi}} x^{-3/2} \exp\left\{-\frac{\lambda'(x-\mu')^2}{2(\mu')^2 x}\right\}, x > 0;$$

5. $\lambda^2|rest \sim Gamma\left(shape = p + r, rate = \sum_i \tau_i^2/2 + \delta\right)$;

6. $\sigma^2|rest \sim INGamma\left(shape = \frac{n-1+p-\sum_i Z_i}{2}, scale = \frac{(\boldsymbol{Y}-\boldsymbol{X\beta})^T(\boldsymbol{Y}-\boldsymbol{X\beta})+\boldsymbol{\beta}^T\boldsymbol{D}_\tau^{-1}\boldsymbol{\beta}}{2}\right)$.