MR.SPI-Robust.CI

ivanxu259

September 2023

1 Introduction

2 Algorithm

Here is an example of a simple algorithm:

Algorithm 1 Constructing A Robust CI via Searching and Sampling with Trimmed Initial Interval

Input: GWAS summary statistics of independent SNPs: $\{\gamma_{bj}, \sigma_{b\gamma j}, \Gamma_{bj}, \sigma_{b\Gamma j}\}_{1\leq j\leq p}$; Sample sizes n_1 for the exposure and n_2 for the outcome; Threshold α^* for selecting relevant IVs; Significance level $\alpha \in (0,1)$; Sampling number M.

Output: The robust confidence interval CI_{robust}.

1: Estimate the set of relevant IVs $\hat{S}$ and valid IVs $\hat{V}$ as in Algorithm 1 with the resulting Voting Matrix: $Vote.Mat$.

2: Find L_{trim} and U_{trim} using -

1. **majority rule:**

$$L_m = \min_{\substack{\hat{V}'\subset\hat{V} \\ |\hat{V}'|=\left\lceil\frac{|\hat{V}|}{2}\right\rceil}} \hat{\beta}_{\hat{V}'} - Z_{1-\frac{\alpha}{2}}\sqrt{\hat{Var}(\hat{\beta}_{\hat{V}'})}$$

$$U_m = \max_{\substack{\hat{V}'\subset\hat{V} \\ |\hat{V}'|=\left\lceil\frac{|\hat{V}|}{2}\right\rceil}} \hat{\beta}_{\hat{V}'} + Z_{1-\frac{\alpha}{2}}\sqrt{\hat{Var}(\hat{\beta}_{\hat{V}'})}$$

2. **plural rule:**

$$L_p = \min_{\substack{\hat{V}'\subset\hat{V} \\ |\hat{V}'|=\max_{l\in\hat{s}/\hat{V}}\mathbf{VM}_l}} \hat{\beta}_{\hat{V}'} - Z_{1-\frac{\alpha}{2}}\sqrt{\hat{Var}(\hat{\beta}_{\hat{V}'})}$$

$$U_p = \max_{\substack{\hat{V}'\subset\hat{V} \\ |\hat{V}'|=\max_{l\in\hat{s}/\hat{V}}\mathbf{VM}_l}} \hat{\beta}_{\hat{V}'} + Z_{1-\frac{\alpha}{2}}\sqrt{\hat{Var}(\hat{\beta}_{\hat{V}'})}$$

$\hat{\beta}_{\hat{V}'}$ and $\hat{Var}(\hat{\beta}_{\hat{V}'})$ are calculated using equation (10) and equation (11) respectively and $VM_i = \sum_{i=1}^{|V|} Vote.Mat_i$

3: Use the initial interval $[L_m, U_m]$ or $[L_p, U_p]$ to obtain the corresponding grid set B.

4: Follow the steps 3-8 in Algorithm 2 to Construct the robust confidence interval CI_{robust}

$$CI_{trim} = [max(L_{robust}, L_{trim}), min(U_{robust}, U_{trim})]$$

3 Results

We compared the length and coverage of the CIs generated by the new methods with the current results generated by MR.SPI in 14 scenarios:

(**S1**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_5, -\mathbf{1}_5)^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_6, \mathbf{1}_4)^T$.

(**S2**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_5, -\mathbf{1}_5)^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_4, \mathbf{1}_3, -\mathbf{1}_3)^T$.

(**S3**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_5, -\mathbf{1}_5)^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_6, \mathbf{1}_2, 0.25, 0.25)^T$.

(**S4**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_5, -\mathbf{1}_5)^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_4, \mathbf{1}_2, 0.25, \mathbf{1}_2, -0.25)^T$.

(**S5**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_5, -\mathbf{1}_5)^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_6, \mathbf{0.25}_4)^T$.

(**S6**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_5, -\mathbf{1}_5)^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_6, \mathbf{0.25}_2, -\mathbf{0.25}_2)^T$.

(**S7**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_5, -\mathbf{1}_5)^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_4, \mathbf{0.25}_3, -\mathbf{0.25}_3)^T$.

(**C1**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_{50}, -\mathbf{1}_{50})^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_{60}, \mathbf{1}_{40})^T$.

(**C2**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_{50}, -\mathbf{1}_{50})^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_{40}, \mathbf{1}_{30}, -\mathbf{1}_{30})^T$.

(**C3**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_{50}, -\mathbf{1}_{50})^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_{60}, \mathbf{1}_{20}, -\mathbf{0.25}_{20})^T$.

(**C4**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_{50}, -\mathbf{1}_{50})^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_{40}, \mathbf{1}_{20}, \mathbf{0.25}_{10}, -\mathbf{1}_{20}, -\mathbf{0.25}_{10})^T$.

(**C5**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_{50}, -\mathbf{1}_{50})^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_{60}, \mathbf{0.25}_{40})^T$.

(**C6**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_{50}, -\mathbf{1}_{50})^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_{60}, \mathbf{0.25}_{20}, -\mathbf{0.25}_{20})^T$.

(**C7**): $\boldsymbol{\gamma} = 0.25 \cdot (\mathbf{1}_{50}, -\mathbf{1}_{50})^T$ and $\boldsymbol{\pi} = 0.25 \cdot (\mathbf{0}_{40}, \mathbf{0.25}_{30}, -\mathbf{0.25}_{30})^T$.

Table 1: Performance Metrics (1000 replications for scenarios S1-S7)

Scenario	Metrics	MR.SPI	S&S	Bound(M)	Bound(P)	Combined(M)	Combined(P)
S1	Coverage	0.938	0.999	1.000	1.000	0.999	0.999
	Length	0.081	0.164	0.144	0.142	0.141	**0.139**
S2	Coverage	0.942	0.991	0.985	0.987	0.980	0.981
	Length	0.106	1.263	0.157	0.215	**0.150**	0.193
S3	Coverage	0.950	0.999	0.999	0.996	0.998	0.995
	Length	0.080	0.164	0.145	0.184	**0.141**	0.149
S4	Coverage	0.715	0.998	0.967	0.967	0.966	0.966
	Length	0.102	0.209	0.167	0.166	0.159	**0.158**
S5	Coverage	0.903	0.979	0.960	0.965	0.953	0.957
	Length	0.080	0.172	0.146	0.133	0.129	**0.121**
S6	Coverage	0.893	0.994	0.996	0.991	0.992	0.986
	Length	0.081	0.175	0.148	0.154	0.141	**0.138**
S7	Coverage	0.485	0.431	**0.698***	0.642	0.317	0.286
	Length	0.096	0.209	0.179	0.179	0.102	0.099

Note: In **S7**, none of the methods achieve 95% coverage. However, the new initial bound selected using Majority Rule was able to achieve the highest coverage(70%) with a shorter average CI length compared with **Sampling&Searching**.

Table 2: Performance Metrics (50 replications for scenarios C1-C7, NA represents intractable within 5000 runs using Genetic Algorithm)

Scenario	Metrics	MR.SPI	S&S	Bound(M)	Bound(P)	Combined(M)	Combined(P)
C1	Coverage	0.84	1.00	NA	1.00	NA	1.00
	Length	0.044	0.173	NA	0.154	NA	**0.149**
C2	Coverage	0.68	0.24	NA	**0.78**	NA	0.24
	Length	0.079	0.603	NA	0.263	NA	-0.142
C3	Coverage	0.78	1.00	NA	1.00	NA	1.00
	Length	0.0432	0.164	NA	**0.154**	NA	0.220
C4	Coverage	0.62	1.00	NA	1.00	NA	1.00
	Length	0.057	0.361	NA	0.228	NA	**0.158**
C5	Coverage	0.50	1.00	NA	0.98	NA	0.98
	Length	0.042	0.199	NA	0.155	NA	**0.147**
C6	Coverage	0.58	1.00	NA	1.00	NA	1.00
	Length	0.045	0.255	NA	0.170	NA	**0.167**
C7	Coverage	0.02	0.76	NA	0.76	NA	0.32
	Length	0.054	0.313	NA	**0.223**	NA	0.218

Scenario	Metrics	MR.SPI	Robust	Robust.vote	Robust.gamma	Robust.combine
$(0_8, 1_7)$	Coverage	0.89	1.00	1.00	1.00	1.00
	Length	0.074	0.163	0.163	0.190	0.189
$(0_6, 1_5, -1_4)$	Coverage	0.940	1.00	1.00	1.00	1.00
	Length	0.090	0.185	0.184	0.232	0.233
$(0_8, 0.25_7)$	Coverage	0.82	0.98	0.97	0.98	0.98
	Length	0.072	0.239	0.238	0.236	0.226
$(0_6, 0.25_5, -0.25_4)$	Coverage	0.53	0.79	0.96	0.90	0.98
	Length	0.082	0.282	0.280	0.242	0.235
$(0_{13}, 1_{12})$	Coverage	0.87	1.00	1.00	1.00	1.00
	Length	0.063	0.163	0.164	0.165	0.165
$(0_9, 1_8, -1_8)$	Coverage	0.01	0	0	0	0
	Length	0.078	NA	NA	NA	NA
$(0_{13}, 0.25_{12})$	Coverage	0.77	0.99	0.98	0.95	0.99
	Length	0.062	0.278	0.268	0.278	0.257
$(0_9, 0.25_8, -0.25_8)$	Coverage	0.02	0.34	0.36	0.68	0.61
	Length	0.070	0.352	0.341	0.345	0.333