

ivd: An R Package for Individual Variance Detection using Spike-and-Slab Priors

Summary

Research in fields such as psychology and education typically focus on the average performance of a unit (e.g., a student or a school), but the *consistency* or variability of that performance can also carry important information about the units of study [@raudenbush1987jes, @leckie_mixed-effects_2023, rast_modeling_2012]. Standard statistical approaches often treat within-cluster variability as a nuisance parameter or assume it is constant across groups. The `ivd` package implements the Spike-and-Slab Mixed-Effects Location Scale Model (SS-MELSM) as a Bayesian framework for explicitly modeling and detecting heterogeneous residual variances.

`ivd` enables users to estimate models for both means (location) and within-cluster variability (scale) simultaneously. It uses a spike-and-slab prior on the scale random effects as a probabilistic selection tool, distinguishing clusters with typical variability (shrunk to the population mean, the spike) from those with atypical variability (the slab) [@mitchell_bayesian_1988]. This approach identifies individual units that differ significantly from the norm, such as schools with unusually consistent or inconsistent test scores, or patients with highly stable or unstable symptoms [@leckie_modeling_2014].

Statement of Need

Standard mixed-effects models partition variance into between-group and within-group components, typically assuming the within-group residual variance (σ^2) is homogeneous across all groups [@raudenbush_hierarchical_2002]. While Mixed-Effects Location Scale Models (MELSM) relax this assumption by allowing σ^2 to vary as a function of covariates and random effects [@hedeker_application_2008], they do not inherently provide a decision rule for identifying *which* specific units deviate significantly from the average consistency.

In `ivd`, MELSMs are extended through the SS-MELSM approach, which computes a Posterior Inclusion Probability (PIP) for each cluster's scale random effect. The model simultaneously estimates fixed and random effects for both location and scale. For the scale, the residual standard deviation σ_{ij} is modeled as a log-linear function of covariates. The scale intercept's random effects are subject to a spike-and-slab prior, with a binary indicator δ_{jk} determining inclusion. A high PIP for δ_j provides strong evidence that the unit belongs to the slab distribution, corresponding to a larger Bayes factor for including the k th random effect for the j th school [@rodriguez_who_2022, @williams_putting_2021]. PIPs are calculated as the proportion of MCMC samples where $\delta_{jk} = 1$.

The package provides a user-friendly interface for NIMBLE [@de_valpine_programming_2017], enabling users to fit SS-MELSMs with standard R formula syntax, without custom BUGS or NIMBLE code. Since languages like Stan do not directly support discrete parameters, implementing these models can be challenging for applied researchers. The main function, `ivd()`, accepts two formulas: `location_formula` for the mean structure and `scale_formula` for within-cluster variance. `ivd` also uses the future package [@bengtsson_2021] to enable parallel processing of MCMC chains.

Analysis and Visualization

The `summary()` function returns the fixed effects for both location and scale, along with the PIPs for the random scale effects. Convergence of each estimate is summarised with computation of \hat{R} , and estimation efficiency by the effective sample size [@vehtariRankNormalization2021].

To help detect atypical units, the package offers several visualization methods, specified by the `type` argument in the `plot()` function. The default, "pip", highlights units exceeding a set probability threshold (default 0.75). "funnel" shows the relationship between PIP and estimated within-cluster standard deviation, while "outcome" displays the interaction between average performance and consistency. For MCMC convergence diagnostics, the `codaplot()` function provides trace and density plots for selected parameters.

Model Comparison

Model comparison is supported using the Widely Applicable Information Criterion (WAIC), which is computed during estimation by default (`WAIC = TRUE`). Additionally, `ivd` stores the pointwise log-likelihood matrix, allowing compatibility with the `loo` package [@loo] for predictive accuracy assessment using Pareto smoothed importance sampling Leave-one-out cross-validation (PSIS-LOO) [@vehtari_practical_2017].

Usage Example

The following example demonstrates the workflow with the `saeb` dataset included in the package. Here, students' mathematics proficiency is modeled within schools, predicting both mean achievement and residual variability as functions of student and school socioeconomic status (SES).

```
library(ivd)
library(data.table)

## Prepare data: separate within- and between-school SES effects
data(saeb)
school_ses <- saeb[, .(school_ses = mean(student_ses, na.rm = TRUE)), by = school_id]
saeb <- saeb[school_ses, on = "school_id"]
saeb$student_ses <- saeb$student_ses - saeb$school_ses
saeb$school_ses <- c(scale(saeb$school_ses, scale = FALSE))

## Fit the Spike-and-Slab MELSM
model <- ivd(
  location_formula = math_proficiency ~ student_ses * school_ses + (1|school_id),
  scale_formula     = ~ student_ses * school_ses + (1|school_id),
  data              = saeb,
  niter             = 2000,
  nburnin          = 1000,
  WAIC              = TRUE,
  workers           = 2
)

## Summarize results
summary(model)

## Visualizations
## Identify schools with unusual variability (PIP > 0.75)
plot(model, type = "pip")

## Inspect the relationship between PIP and within-cluster SD
plot(model, type = "funnel")

## Check convergence for the intercept
codaplot(model, parameters = "Intc")
```

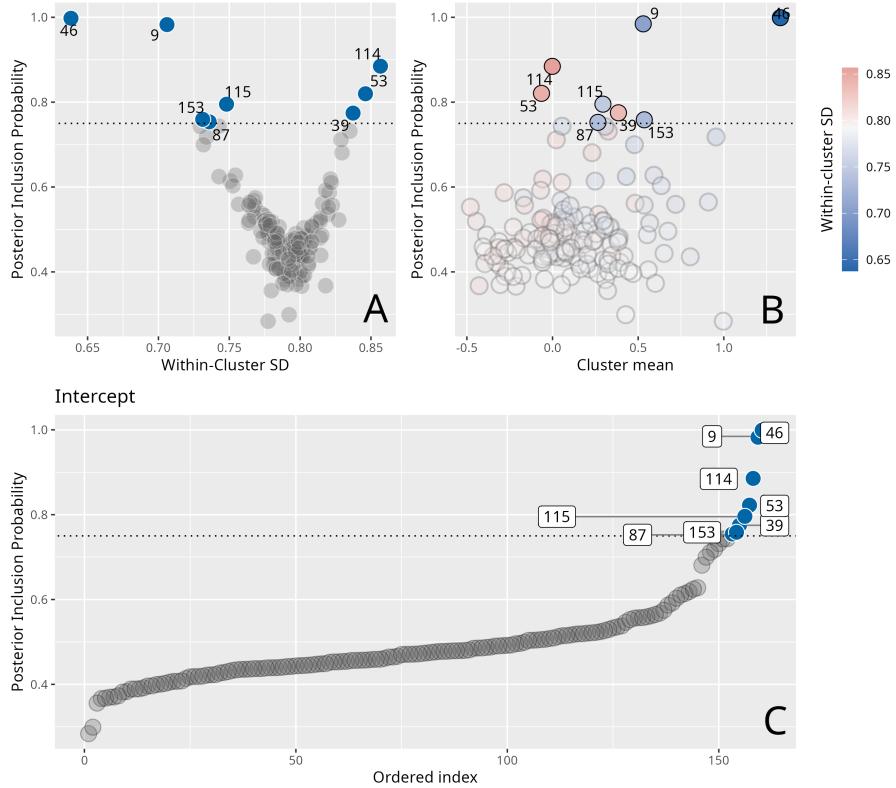


Figure 1: Scatter plots show posterior inclusion probability (PIP) for the scale random intercept. Panel A plots PIP against the estimated within-cluster standard deviation (SD), with the y-axis as PIP and the x-axis as within-cluster SD. Panel B plots PIP against estimated math achievement scores, with the x-axis showing math achievement scores. Panel C shows PIPs for each school, sorted on the horizontal axis. The dotted horizontal line denotes the PIP threshold of 0.75.

Acknowledgements

This work was supported by the Tools Competition catalyst award for the project consistentlyBetter to PR.