

Readme for simulation files:

What is in this folder?

Scripts to run and summarize simulations for the paper under strong IV assumptions, with Gamma distribution for the errors on outcome.

Codes scripts:

Script to run the simulations with 3 different sample sizes

1. Simulation_strongIV_n_100_example.R
2. Simulation_strongIV_n_500_example.R
3. Simulation_strongIV_n_2000_example.R

Script to summarize the output from each file

1. SummarizeSimulations_n_100.R
2. SummarizeSimulations_n_500.R
3. SummarizeSimulation_n_2000.R

Structure of these files for different sample sizes are very similar. We created different versions to automate data generation, model fitting and summarizing the model fits.

Script to run simulation and summarizing scripts all at once:

runSimulations.R

Run time:

Data:

Example dataset is included in the package

It contains an example of simulated dataset to fit the model and summarize fits.

To see what is in the dataset type following in R console after loading the package 'BayesIV'

```
Data(package='BayesIV')
```

What is in this data file?

- i. Yobs = observed outcome vector
- ii. D = vector of treatment assignment
- iii. X = matrix of confounders
- iv. Z = vector of instrumental variable

How to analyze this dataset:

1. Fitting the NP-Bayes IV model with default initialization and priors to draw 100 posterior samples. Refer to BayesIV-manual for more details.

```
Model = mcmcRun_Normal_DPM(Yobs = Yobs, Tr = D, X = X, Z = Z, niter = 100, priors = NULL,  
initialVals = NULL)
```

2. Summarizing the fitted model by specifying burnin and thin values for the MCMC chain

```
ATE_chain = getATE_posterior_NDPM(fittedModel = Model, X = X, niter = 100, burnin = 0, thin = 1)  
ATE = mean(ATE_chain)
```

Instruction on how to install the package in R:

```
install.packages('devtools')  
library(devtools)  
install_github('SamAdhikari/BayesIV')
```