BAYESIAN PROJECT

Ana Mendoza

2025-04-08

load and clean the dataset

# Load your dataset  
df <- read.csv("/Users/anamdza/Downloads/bayesianproject.csv")  
  
# Keep only rows with both Sample Size and Number of Cases  
df\_clean <- df[!is.na(df$Sample.Size) & !is.na(df$Number.of.Cases), ]  
  
# Rename for simplicity  
df\_clean$y <- df\_clean$Number.of.Cases  
df\_clean$n <- df\_clean$Sample.Size

prepare data for jags

# Build the data list for JAGS  
data\_list <- list(  
 y = as.integer(df\_clean$y),  
 n = as.integer(df\_clean$n),  
 N = nrow(df\_clean)  
)

define and run jags model

library(rjags)

## Loading required package: coda

## Linked to JAGS 4.3.2

## Loaded modules: basemod,bugs

# Model: Binomial likelihood with uniform prior on theta  
model\_string <- "  
model {  
 for (i in 1:N) {  
 y[i] ~ dbin(theta[i], n[i])  
 theta[i] ~ dbeta(1, 1)  
 }  
}  
"  
  
# Compile and run the model  
jags\_model <- jags.model(textConnection(model\_string), data = data\_list, n.chains = 3)

## Compiling model graph  
## Resolving undeclared variables  
## Allocating nodes  
## Graph information:  
## Observed stochastic nodes: 173  
## Unobserved stochastic nodes: 173  
## Total graph size: 521  
##   
## Initializing model

update(jags\_model, 1000) # Burn-in  
samples <- coda.samples(jags\_model, variable.names = c("theta"), n.iter = 5000)

plot posterior for one study

posterior\_matrix <- as.matrix(samples)  
  
# Plot histogram for theta[1]  
hist(posterior\_matrix[, 1],  
 main = "Posterior Distribution for Study 1",  
 xlab = "Prevalence Rate",  
 col = "lightblue", border = "white")

