Exercise 3.3.2

Saurabh Steixner-Kumar
(social): - -

Contents

1 Question

Exercise 3.3.2 Try the data k1 = 0, n1 = 10, k2 = 10, n2 = 10. What does this analysis infer the common rate θ to be? Do you believe the inference?

2 Comments/Solution

In this example of the dataset we see that the individually the k1 and k2 are drastically different and show the two extreme probabilities of 0 and 1 respectively. But, since we have the model set up that a common rate (latent probability) controls both we see that the posterior estimate θ to be in the middle around 0.5. Check out the plots section.

The model used to calculate the required values and the plots is scripted below. Copy/pasting the given code will generate the same result on your own machine.

3 Code

3.1 libraries

The libraries required for the script and the plots.

```
# clears workspace
rm(list=ls())
#load libraries
library(rstan)
library(ggplot2)
library(patchwork)
```

3.2 Data

The data required for this particular stan model.

```
# data  initialization
k1 <- 0;n1 <- 10;k2 <- 10;n2 <- 10
# to be passed on to Stan
stan_data <- list(k1 = k1, n1 = n1, k2 = k2, n2 = n2)
```

3.3 Stan code

Stan code, that can be written in R as such or in a seperate new file with stan extension.

```
write("// Stan code here in this section

// Inferring delta through theta1 and theta2
data {
  int<lower=1> n1;
  int<lower=1> n2;
  int<lower=0> k1;
  int<lower=0> k2;
}
parameters {
  real<lower=0,upper=1> theta;
}
model {
  // Prior on Single Rate Theta
  theta ~ beta(1, 1);
  // Observed Counts
  k1 ~ binomial(n1, theta);
  k2 ~ binomial(n2, theta);
} // ",

"3_3_2.stan")
```

3.4 code in R to run stan

Running stan through R (with the required input parameters).

```
myinits <- list(
  list(theta=.1),  # chain 1 starting value
  list(theta=.9))  # chain 2 starting value

# parameters to be monitored:
parameters <- c("theta")

# The following command calls Stan with specific options.
# For a detailed description type "?stan".
mod_fit <- stan(file="3_3_2.stan",
                data=stan_data,
                init=myinits,  # If not specified, gives random inits
                pars=parameters,
                iter=2000,
                chains=2,
                thin=1,
                warmup=100,  # Stands for burn-in; Default = iter/2
```

```
                seed=123  # Setting seed; Default is random seed
)
```

4 Outputs

4.1 Model summary

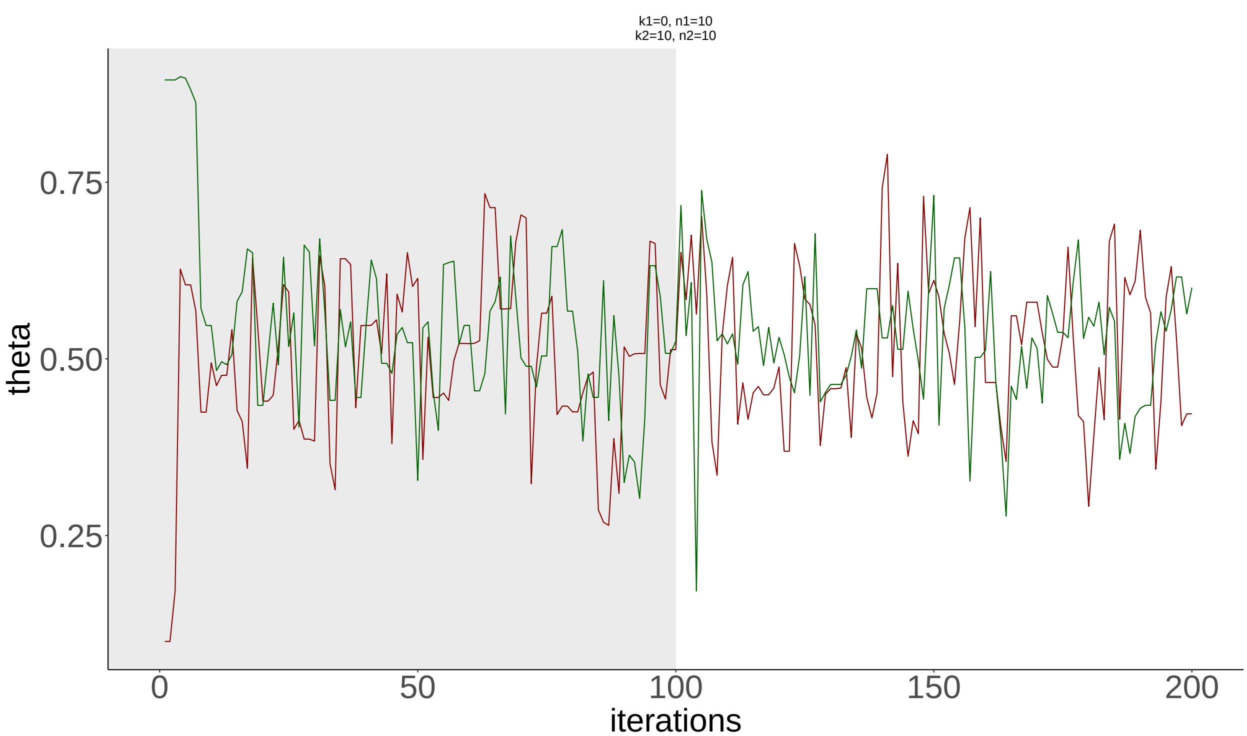
In order of definition.

```
## Inference for Stan model: 3_3_2.
## 2 chains, each with iter=2000; warmup=100; thin=1;
## post-warmup draws per chain=1900, total post-warmup draws=3800.
##
##          mean se_mean  sd   2.5%    25%    50%    75%  97.5% n_eff Rhat
## theta   0.50    0.00 0.1   0.30   0.43   0.50   0.57   0.70  1513    1
## lp__  -15.76    0.02 0.7 -17.78 -15.93 -15.49 -15.30 -15.25  2169    1
##
## Samples were drawn using NUTS(diag_e) at Mon Oct 26 18:55:11 2020.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

4.2 Plots

4.2.1 Plot (chains)

The initial movement of the chains are shown here (including the warmup phase). The two chains begin from the initial starting points of as defined in the input parameters of the stan model.



4.2.2 Plot (posterior)

The plot of the θ values per chain superimposed on each other.