Exercise 5.3.1

Saurabh Steixner-Kumar
(social): - -

Contents

1 Question

Exercise 5.3.1 Influenza Clinical Trial Poehling, Griffin, and Dittus (2002) reported data evaluating a rapid bedside test for influenza using a sample of 233 children hospitalized with fever or respiratory symptoms. Of the 18 children known to have influenza, the surrogate method identified 14 and missed 4. Of the 215 children known not to have influenza, the surrogate method correctly rejected 210 but falsely identified 5. These data correspond to a = 14, b = 4, c = 5, and d = 210. Examine the posterior distributions of the interesting variables, and reach a scientific conclusion. That is, pretend you are a consultant for the clinical trial. What would your two- or three-sentence 'take home message' conclusion be to your customers?

2 Comments/Solution

The surrogate method does a better job detecting the absence of influenza than it does detecting the presence of influenza. The 95% Bayesian confidence interval (ETI) for kappa is (.52, .85), suggesting that the test is useful.

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.

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

3.2 Data

The data required for this particular stan model.

```r
# data  initialization
# Influenza
y <- c(14, 4, 5, 210)
# to be passed on to Stan
stan_data <- list(y=y)
```

3.3 Stan code

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

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

// Kappa Coefficient of Agreement
data {
  int<lower=0> y[4];
}
parameters {
  // Underlying Rates
  // Rate Objective Method Decides 'one'
  real<lower=0,upper=1> alpha;
  // Rate Surrogate Method Decides 'one' When Objective Method Decides 'one'
  real<lower=0,upper=1> beta;
  // Rate Surrogate Method Decides 'zero' When Objective Method Decides 'zero'
  real<lower=0,upper=1> gamma;
}
transformed parameters {
  simplex[4] pi;
  real xi;
  real psi;
  real kappa;
  // Probabilities For Each Count
  pi[1] = alpha * beta;
  pi[2] = alpha * (1 - beta);
  pi[3] = (1 - alpha) * (1 - gamma);
  pi[4] = (1 - alpha) * gamma;

  // Derived Measures
  // Rate Surrogate Method Agrees With the Objective Method
  xi = alpha * beta + (1 - alpha) * gamma ;
  // Rate of Chance Agreement
  psi = (pi[1] + pi[2]) * (pi[1] + pi[3]) + (pi[2] + pi[4]) * (pi[3] + pi[4]);
```

```
  // Chance-Corrected Agreement
  kappa = (xi - psi) / (1 - psi);
}
model {
  alpha ~ beta(1, 1);  // could be removed
  beta ~ beta(1, 1);  // could be removed
  gamma ~ beta(1, 1);  // could be removed
  // Count Data
  y ~ multinomial(pi);
} // ",

"5_3_1.stan")
```

3.4 code in R to run stan

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

```
myinits <- list(
  list(alpha=.5, beta=.5, gamma=.5),  # chain 1 starting value
  list(alpha=.5, beta=.5, gamma=.5))  # chain 2 starting value

# parameters to be monitored:
parameters <- c("kappa", "xi", "psi", "alpha", "beta", "gamma", "pi")

# The following command calls Stan with specific options.
# For a detailed description type "?stan".
mod_fit <- stan(file="5_3_1.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.

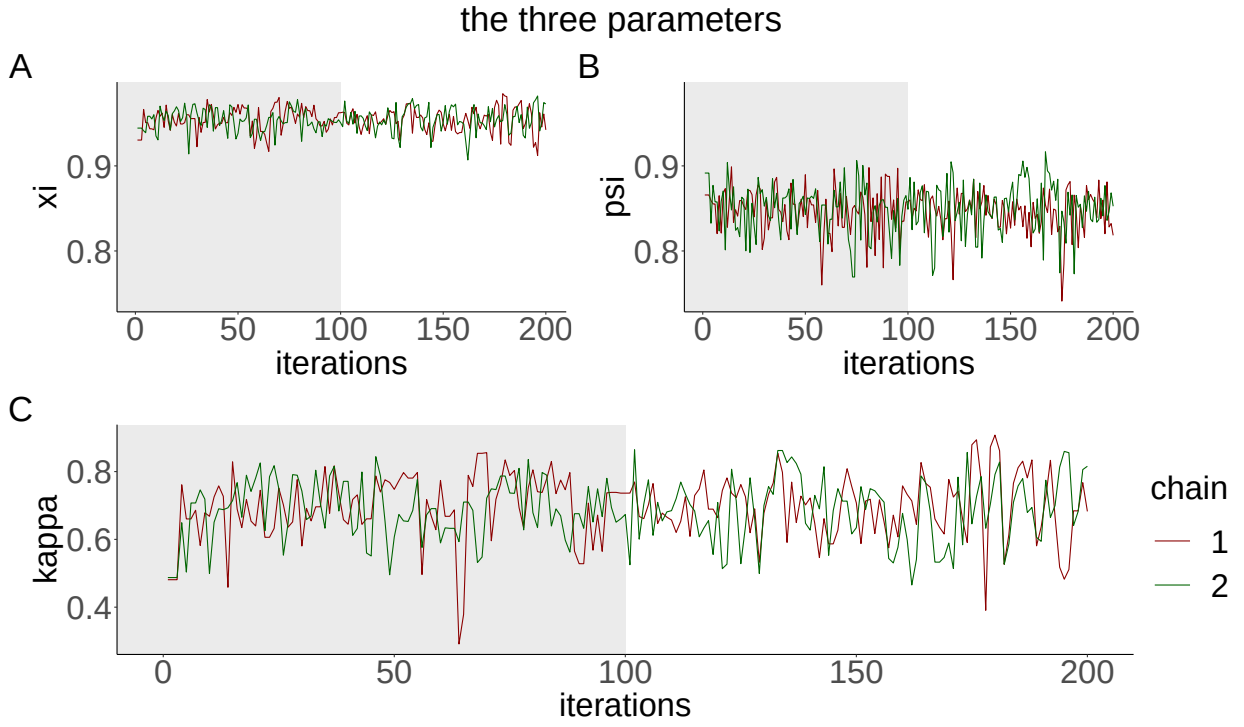
```
## Inference for Stan model: 5_3_1.
## 2 chains, each with iter=2000; warmup=100; thin=1;
## post-warmup draws per chain=1900, total post-warmup draws=3800.
##
##          mean se_mean   sd    2.5%     25%     50%     75%   97.5% n_eff
## kappa    0.70    0.00 0.08    0.52    0.64    0.71    0.76    0.85  2485
## xi       0.95    0.00 0.01    0.93    0.95    0.96    0.96    0.98  2661
## psi      0.85    0.00 0.03    0.79    0.83    0.85    0.87    0.90  4368
## alpha    0.08    0.00 0.02    0.05    0.07    0.08    0.09    0.12  5567
## beta     0.75    0.00 0.09    0.54    0.69    0.76    0.82    0.91  2383
## gamma    0.97    0.00 0.01    0.95    0.97    0.97    0.98    0.99  2786
```

```
## pi[1]    0.06    0.00 0.02    0.03    0.05    0.06    0.07    0.10  3860
## pi[2]    0.02    0.00 0.01    0.01    0.01    0.02    0.03    0.04  2815
## pi[3]    0.03    0.00 0.01    0.01    0.02    0.02    0.03    0.05  2794
## pi[4]    0.89    0.00 0.02    0.85    0.88    0.89    0.91    0.93  4124
## lp__  -106.23    0.03 1.25 -109.55 -106.80 -105.92 -105.31 -104.81  1772
##       Rhat
## kappa    1
## xi       1
## psi      1
## alpha    1
## beta     1
## gamma    1
## pi[1]    1
## pi[2]    1
## pi[3]    1
## pi[4]    1
## lp__     1
##
## Samples were drawn using NUTS(diag_e) at Thu Nov 05 21:25:46 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

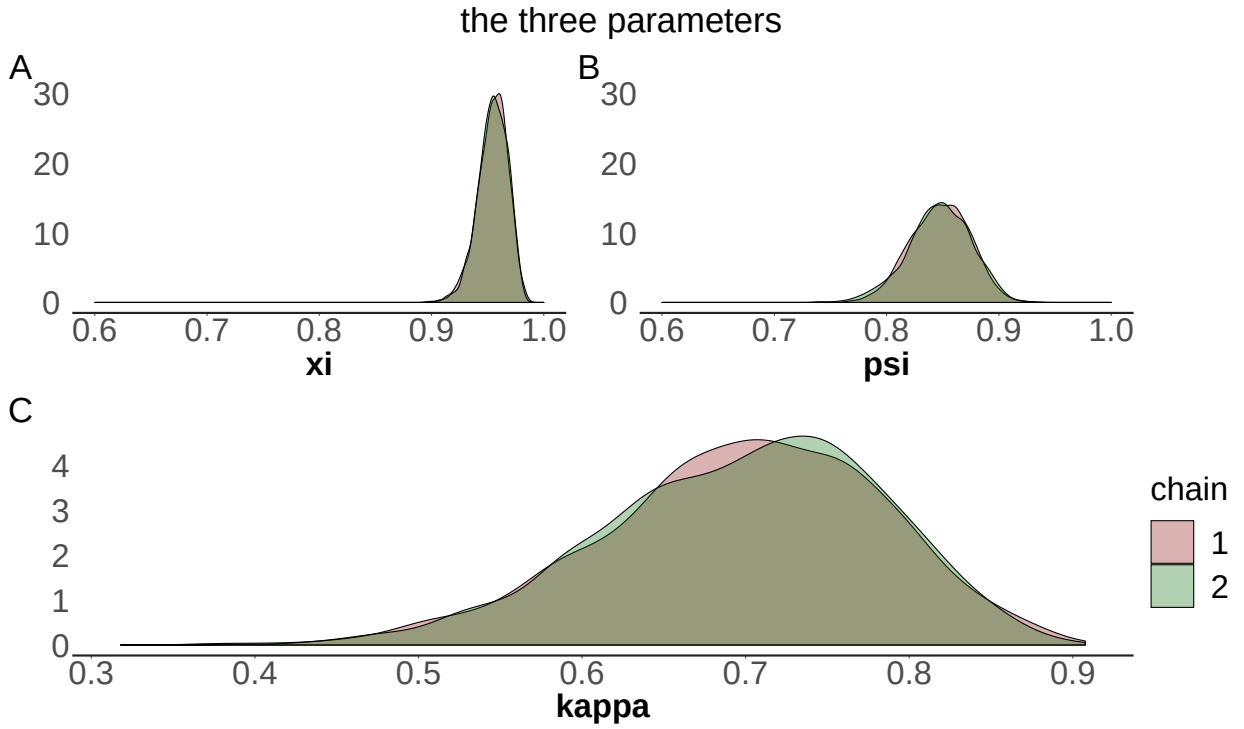
4.2.1 Plot (chains)

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.



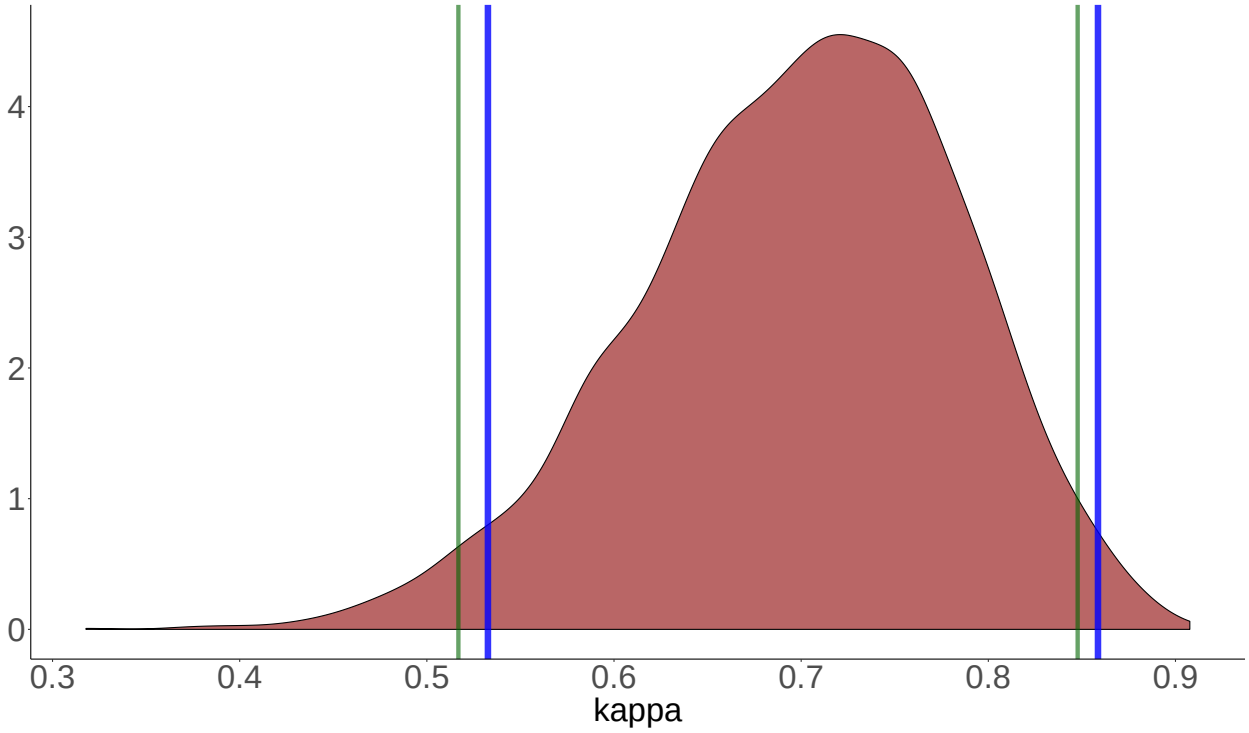
4.2.2 Plot (posterior)

The plot of the rate of agreement (ξ), rate of agreement by chance (ψ) and the chance-corrected agreement rate (κ) values per chain superimposed on each other.



4.2.3 Plot (posterior with CI)

The plot of the combined κ value with confidence intervals. (HDI CI : blue; ETI CI : green)



4.3 Confidence interval value

The confidence interval values.

```
## # Highest Density Interval
##
##       95% HDI
##  [0.53, 0.86]
```

```
## # Equal-Tailed Interval
##
##       95% ETI
##  [0.52, 0.85]
```