

Concepts of Bayesian Data Analysis: Project4

Charles Muiruri* Gianni Guarraci† Juan Vanegas Jadan‡ Ronald Makanga§

Report generated - May 23, 2024

Contents

Modeling using Open Bugs	2
Model	2
Density plots	2
Trace Plots	4
Correlation plot	5
Autocorrelation plot	6
Appendix	7

*Universitiet Hasselt, charles.muiruri@student.uhasselt.be

†Universitiet Hasselt, gianni.guarraci@student.uhasselt.be

‡Universitiet Hasselt, juan.vanegasjadan@student.uhasselt.be

§Universitiet Hasselt, ronald.makanga@student.uhasselt.be

Modeling using Open Bugs

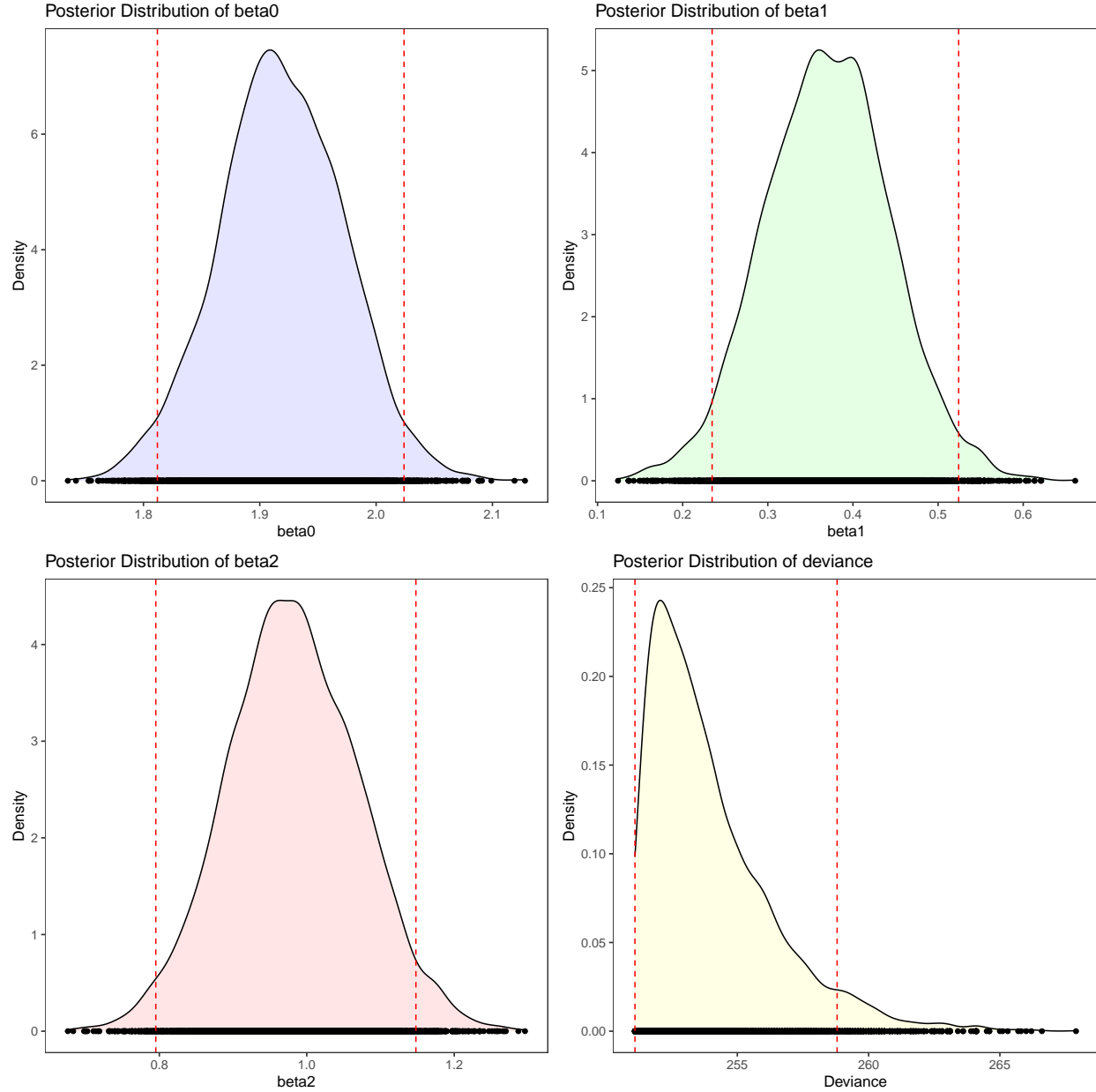
Model

```
## Abstracting beta0 ... 5000 valid values
## Abstracting beta1 ... 5000 valid values
## Abstracting beta2 ... 5000 valid values
## Abstracting deviance ... 5000 valid values

##
## Iterations = 5001:10000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 5000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean      SD Naive SE Time-series SE
## beta0      1.9187 0.05379 0.0007607      0.0008041
## beta1      0.3698 0.07410 0.0010480      0.0011437
## beta2      0.9811 0.08931 0.0012630      0.0014289
## deviance 253.9800 2.39063 0.0338086      0.0383106
##
## 2. Quantiles for each variable:
##
##           2.5%      25%      50%      75%      97.5%
## beta0      1.8120  1.8830  1.9170  1.9560  2.0240
## beta1      0.2239  0.3197  0.3700  0.4184  0.5164
## beta2      0.8058  0.9216  0.9797  1.0430  1.1620
## deviance 251.3000 252.2000 253.3000 255.1000 260.1000
##
## [[1]]
##           lower      upper
## beta0      1.8120  2.0240
## beta1      0.2346  0.5242
## beta2      0.7952  1.1480
## deviance 251.1000 258.8000
## attr(,"Probability")
## [1] 0.95
```

Density plots

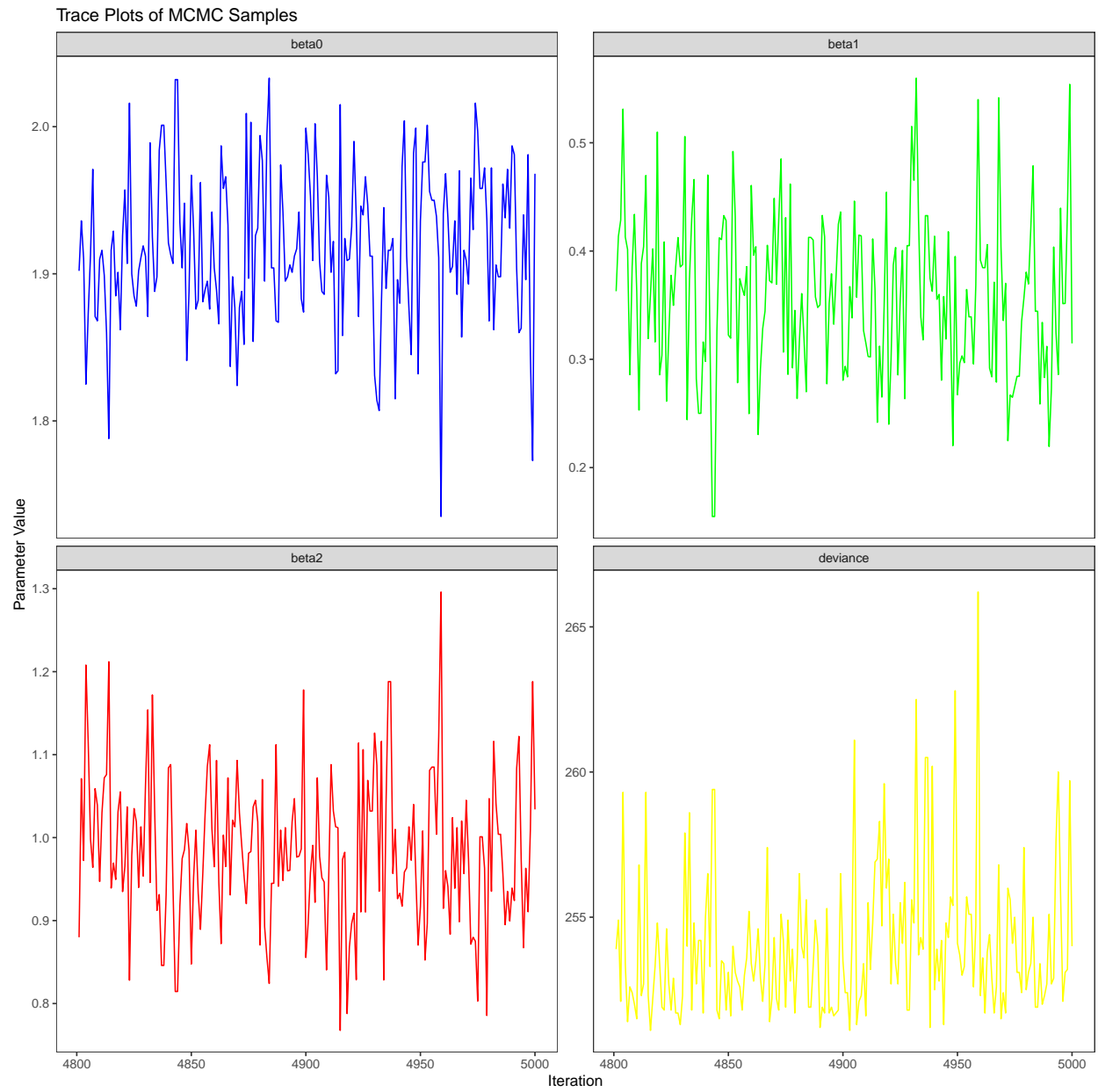
```
## [1] 0.2346
```



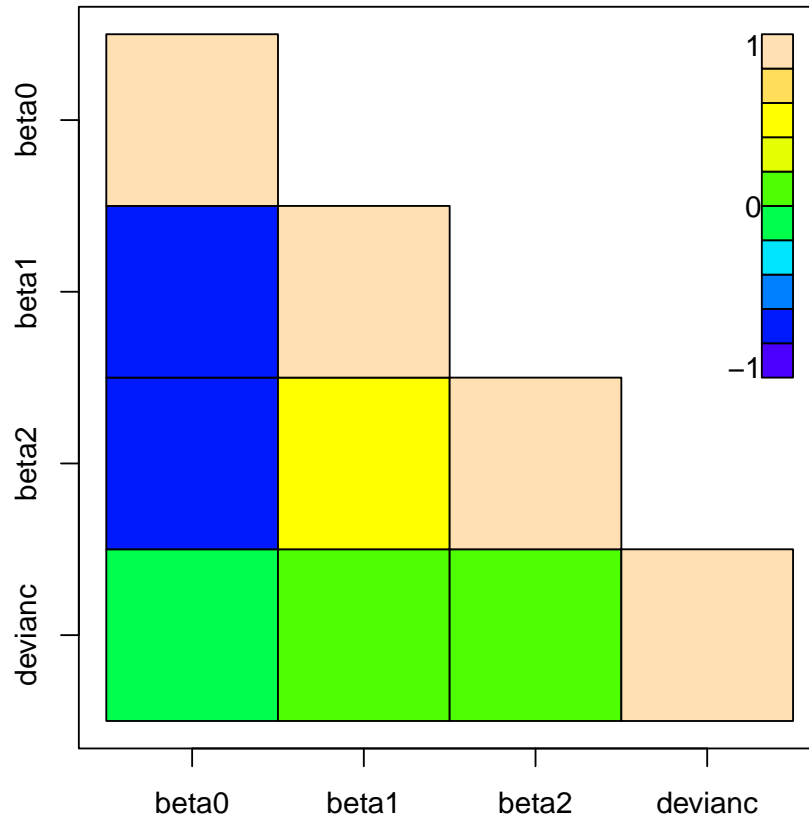
Summary of posterior distribution

- The Posterior of β_0 has a mean of **1.9186524** which is the central tendency of the intercept term, with a 95% credible HPD interval of **(1.812, 2.024)** in which the intercept term lies with 95% probability.
- The Posterior of β_1 has a mean of **0.3697874** which is the central tendency of the Poverty = 133% to <400% FPL term, with a 95% HPD interval of **(0.2346, 0.5242)** in which this Poverty term lies with 95% probability.
- The Posterior of β_2 has a mean of **0.9811075** which is the central tendency of the Poverty = >400% FPL term, with a 95% HPD interval of **(0.7952, 1.148)** in which this Poverty term lies with 95% probability.
- The Posterior of the **Deviance** has a mean of **253.98** which is the central tendency of the Deviance, with a 95% HPD interval of **(251.1, 258.8)**.

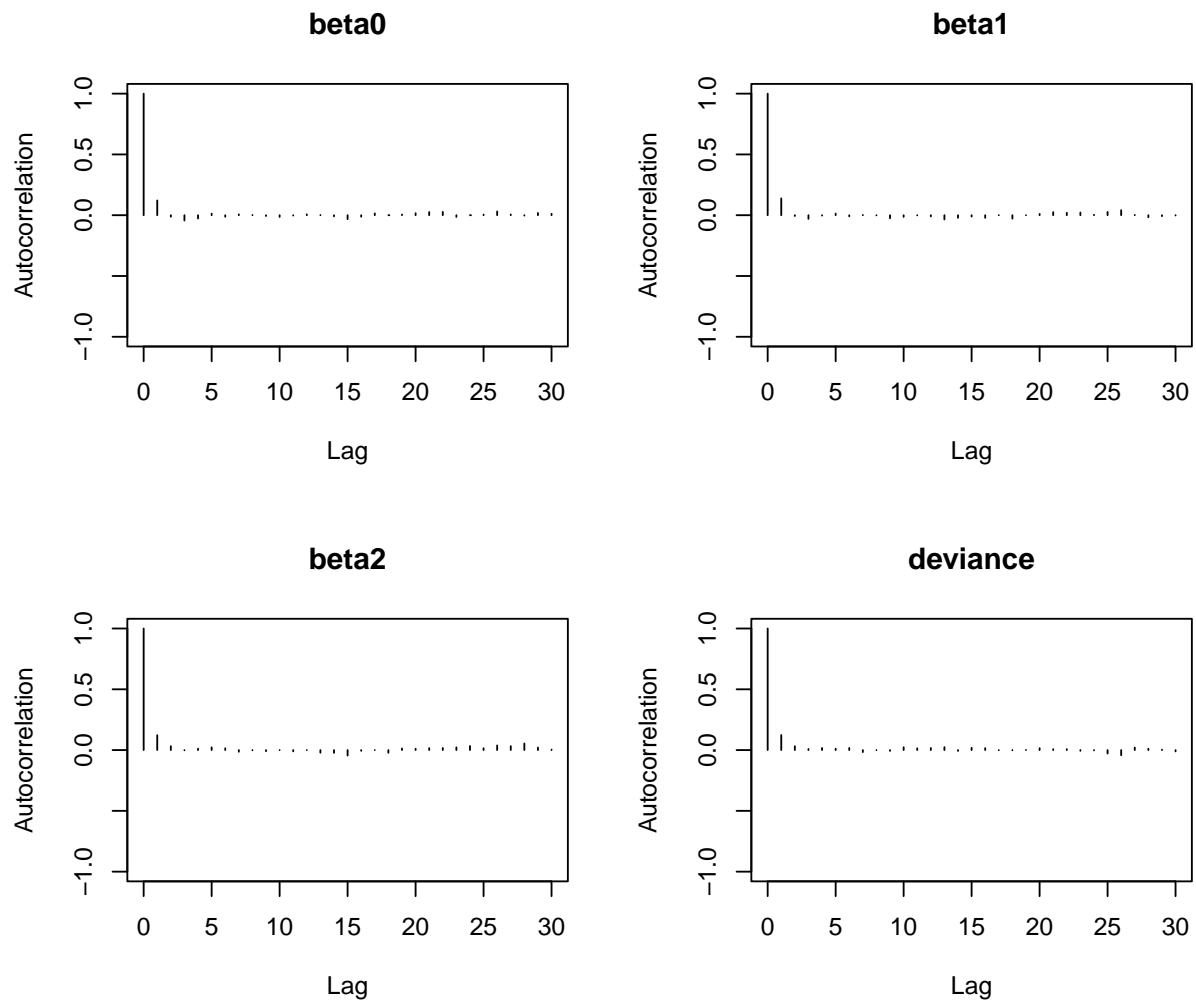
Trace Plots



Correlation plot



Autocorrelation plot



Intepretation of the results is done in the main text.

Appendix

R Code

```
if (!require(pacman)) install.packages("pacman")
p_load(rjags, coda, nimble, R2OpenBUGS, ggplot2, here, dplyr, ggpubr, tidyr)

projdata <- as.data.frame(read.csv(here("data/projectdata.txt")))|>
  mutate(less_133pc = ifelse(Poverty == "<133% FPL", 1,0),
         btn133_400_pc = ifelse(Poverty == "133% to <400% FPL", 1,0),
         great_400pc = ifelse(Poverty == ">400% FPL", 1,0))
# Data prep for bugs
model_data <- list(
  Y = projdata$Vaccinated,
  N = projdata$Sample.Size,
  btn133_400_pc = projdata$btn133_400_pc,
  great_400pc = projdata$great_400pc,
  J = nrow(projdata)
)

model_inits <- list(
  list(beta0 = 0, beta1 = 0, beta2 = 0)
)

parameters <- c("beta0", "beta1", "beta2")

model1 <- function(){
  for (i in 1:J){
    Y[i] ~ dbin(p[i], N[i])
    logit(p[i]) <- beta0 + beta1*btn133_400_pc[i] + beta2*great_400pc[i]
  }
  #priors
  beta0 ~ dnorm(0, 0.001)
  beta1 ~ dnorm(0, 0.001)
  beta2 ~ dnorm(0, 0.001)
}

# Write model to file
write.model(model1, here("models/model1code.txt"))
# View file
file.show(here("models/model1code.txt"))

model.out <- bugs(model_data, model_inits,
  parameters = parameters, model.file = here("models/model1code.txt"),
  n.chains = 1, n.iter = 10000, n.burnin = 5000, codaPkg = TRUE,
  debug = FALSE)

# debug=TRUE opens openBug and displays traceplots and summaries

# Model output
out <- read.bugs(model.out)
summary(out)

# Prep data for density and trace plots
```

```

mcmc_samples <- as.mcmc(out)
mcmc_df <- as.data.frame(mcmc_samples)
mcmc_df$iteration <- 1:nrow(mcmc_df)
mcmc_long <- pivot_longer(mcmc_df, cols = -iteration, names_to = "Parameter",
                          values_to = "Value")

points_data <- data.frame(x = mcmc_df$beta0, x1 = mcmc_df$beta1, x2 = mcmc_df$beta2,
                          x3 = mcmc_df$deviance, y = rep(0, nrow(mcmc_df)))

# Density plots
ggarrange(p_beta0 <- ggplot(mcmc_df, aes(x = beta0)) +
  geom_density(fill = "blue", alpha = 0.1) +
  geom_point(data = points_data, aes(x = x, y = y)) +
  labs(title = "Posterior Distribution of beta0", x = "beta0", y = "Density")+
  theme_bw(base_size = 12)+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()),

p_beta1 <- ggplot(mcmc_df, aes(x = beta1)) +
  geom_density(fill = "green", alpha = 0.1) +
  geom_point(data = points_data, aes(x = x1, y = y)) +
  labs(title = "Posterior Distribution of beta1", x = "beta1", y = "Density") +
  theme_bw(base_size = 12)+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()),

p_beta2 <- ggplot(mcmc_df, aes(x = beta2)) +
  geom_density(fill = "red", alpha = 0.1) +
  geom_point(data = points_data, aes(x = x2, y = y)) +
  labs(title = "Posterior Distribution of beta2", x = "beta2", y = "Density")+
  theme_bw(base_size = 12)+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()),

p_dev <- ggplot(mcmc_df, aes(x = deviance)) +
  geom_density(fill = "yellow", alpha = 0.1) +
  geom_point(data = points_data, aes(x = x3, y = y)) +
  labs(title = "Posterior Distribution of deviance", x = "Deviance", y = "Density")+
  theme_bw(base_size = 12)+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()),
nrow = 2, ncol = 2)

# Trace plots
ggplot(mcmc_long, aes(x = iteration, y = Value, color = Parameter)) +
  geom_line() +
  scale_color_manual(values = c("blue", "green", "red", "yellow")) +
  facet_wrap(~ Parameter, scales = "free_y") +
  labs(title = "Trace Plots of MCMC Samples", x = "Iteration", y = "Parameter Value") +
  theme_bw(base_size = 12)+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        legend.position = "none")

# Autocorrelation and crosscorrelation plots
crosscorr.plot(out)
autocorr.plot(out)

```