Lab Assignment 3
PB HLTH 250C: Advanced Epidemiologic Methods

Katherine Rose Wolf

March 10, 2020

Question One

Using the R code provided, complete Table 1 using the posterior samples of the odds ratios. ***(20 points)***

Table 1: Posterior median and 95% credible intervals for odds ratios from logistic regression model of overweight status on smoking, controlling for age, sex, and education level.

Variable	Vague prior	Informative Prior 1[a]	Informative Prior 2[b]
Current smoker (versus not)			
Age (per year increase)			
Male sex (versus female)			
High school education (versus < high school education)			
Some college (versus < high school education)			
College plus (versus < high school education)			

[a] Prior mean for OR of current smoking = 2, prior variance = 1000.

[b] Prior mean for OR of current smoking = 2, prior variance = 0.08.

Question Two

Using the parameterization for Informative Prior 1, calculate the prior 95% interval for the smoking OR. ***Hint: Calculate the interval on the scale of the log-OR (β) and transform the limits.*** **In** ***one or two sentences*** **describe how this compares to the prior interval for Informative Prior 2 stated in the instructions above.** ***(10 points)***

Let β_s denote the normal prior for the log odds ratio comparing the odds (risk) of overweight (body mass index > 25) in a smoker to that in a nonsmoker. The parameterization for β_s given for Informative Prior 1 states that β_s is normally distributed, i.e., $\beta_s \sim N(\mu_s, \sigma_s^2)$, such that the odds ratio e^{β_s}, or the natural exponentiation of the mean of the log-OR, is 2, i.e., $e^{E[\beta_s]} = e^{\mu_s} = 2$, and β_s has a variance of 1000, i.e., $\sigma_s^2 = 1000$.

To get the mean of the log-OR, then, we take the natural logarithm of the natural exponentiation of the mean of the log-OR, i.e., $\mu_s = \log e^{\mu_s} = \log(2)$.

To get the standard deviation of β_s, σ_s, we take the square root of the variance σ_s^2, i.e., $\sigma_s = \sqrt{\sigma_s^2} = \sqrt{1000}$.

Then we can calculate the prior 95% interval for the log-OR by taking 1.96 standard deviations above and below the mean:

Upper bound on prior 95% interval for β_s: $\mu_s + 1.96\sigma_s = \log(2) + 1.96\sqrt{1000} = 62.6737893$

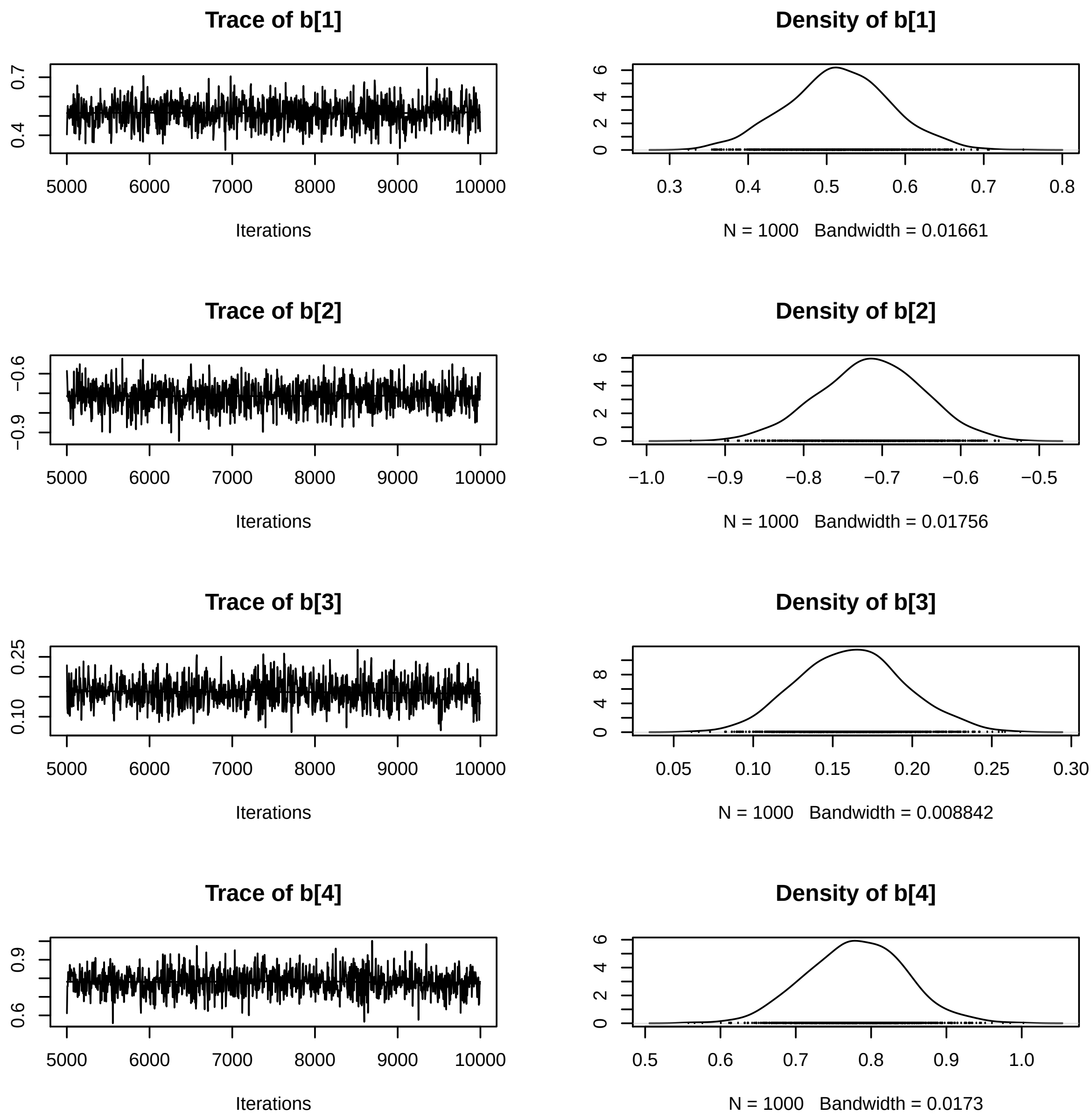
Lower bound on prior 95% interval for β_s: $\mu_s - 1.96\sigma_s = \log(2) - 1.96\sqrt{1000} = -61.287495$

To get the prior 95% interval for the OR, e^{β_s}, then, we exponentiate the prior 95% interval for β_s:

Upper bound on prior 95% interval for e^{β_s}: $e^{\mu_s + 1.96\sigma_s} = e^{\log(2)+1.96\sqrt{1000}} = 1.6553158 \times 10^{27}$

Upper bound on prior 95% interval for e^{β_s}: $e^{\mu_s - 1.96\sigma_s} = e^{\log(2)-1.96\sqrt{1000}} = 0$

Thus the prior estimate for e^{β_s} and 95% prior interval are $0.6931(1.6553 \times 10^{27}, 2.4165 \times 10^{-27})$.



R code

```
knitr::opts_chunk$set(echo = FALSE,
                      warning = FALSE,
                      message = FALSE)


library(R2jags)
library(coda)
require(foreign)
```

```r
load("frmgham_recoded_three.Rdata")


info1_mu <- signif(log(2), 4)

info1_prior_95_upper <- signif(exp(log(2) + 1.96*sqrt(1000)), 5)

info1_prior_95_lower <- signif(exp(log(2) - 1.96*sqrt(1000)), 5)



knitr::include_graphics(
  path = file.path("pdfs",
                   "TraceplotLogisticReg1.pdf"))
```