assignment 3

Beimnet Taye

2024-03-13

Q1

	Vague Prior	Informative Prior 1*	Informative Prior 2†
Current smoker (vs. non)	1.29 (1.09, 1.52)	1.29 (1.1, 1.51)	1.44 (1.25, 1.67)
Age (per year increase)	1.48 (1.37, 1.61)	1.48 (1.36, 1.61)	1.5 (1.38, 1.63)
Male sex (vs. female)	1.36 (1.16, 1.59)	1.36 (1.15, 1.59)	1.32 (1.13, 1.55)
High school education (vs. < HS)	0.76 (0.62, 0.92)	0.75 (0.62, 0.91)	0.75 (0.62, 0.91)
Some college (vs. < HS)	0.61 (0.48, 0.78)	0.61 (0.47, 0.77)	0.61 (0.48, 0.78)
College+ (vs. < HS)	0.59 (0.44, 0.78)	0.59 (0.44, 0.79)	0.59 (0.45, 0.78)

* Prior mean for OR of current smoking=2, prior variance of log-OR=1000.

† Prior mean for OR of current smoking=2, prior variance of log-OR= 0.02.

Q2

The prior 95% interval for Informative Prior 1 is: $log(2) \pm 1.96 * \sqrt{1000}$ which is equal to: $(-61.3, 62.7)$. This is a wide prior for the beta coefficient for smoking, especially when compared to the corresponding interval for Informative Prior 2, making it less informative than that of Informative Prior 2.

Q3

Informative Prior 2 had more of an effect of dragging the smoking OR closer to 2 since the precision is larger for informative prior 2 resulting in prior 2 to influence the posterior distribution more.

Q4

Based on the traceplots it seems to model converged nicely for all parameters measured. Every parameter has a hairy caterpillar like trace around a horizontal line indicating good convergence. The density plots also seem pretty uni-modal indicating good convergence. The autocorrelation plots for all parameters indicate that sample auto correlation drops off pretty quickly indicating good model mixing and convergence.

Q5

```
## [[1]]
##
## Fraction in 1st window = 0.1
## Fraction in 2nd window = 0.5
##
##     b[1]     b[2]     b[3]     b[4]     b[5]     b[6]     b[7]
## -0.49909 -0.08486 -0.93984  0.59362  2.18720  0.25918 -0.05694
##
##
## [[2]]
##
## Fraction in 1st window = 0.1
## Fraction in 2nd window = 0.5
##
##    b[1]    b[2]    b[3]    b[4]    b[5]    b[6]    b[7]
##  1.2364 -1.2067 -0.6073 -0.3816  1.4591 -0.5399 -0.1871
##
##
## [[3]]
##
## Fraction in 1st window = 0.1
## Fraction in 2nd window = 0.5
##
##    b[1]    b[2]    b[3]    b[4]    b[5]    b[6]    b[7]
##  1.0344 -1.9889 -1.1871  1.2676  0.6632  1.8664 -0.8262
```

From the results of the Geweke test it seems the 5th beta in the first chain and the second beta in the third chain failed to converge as their Z test statistics were greater than 1.96 in magnitude, indicating that the mean of the first 10% of the chain differed significantly from the mean of the final 50% of the chain. This means that the sampler varied/wandered too much when sampling.