MTH225 Spring2017 Final Problem 20

In this exercise we estimate the mean and standard deviation of a normal population from a sample.

The CDC guidelines for lead levels in the blood of children ages 1-5 state that concentrations above 5 micrograms per deciliter of blood. Suppose the data in `MTH225_Spring2017_Final_Problem20.csv` represents blood lead levels in a sample of children.

The variable names are:

- `lead` lead concentration in micrograms per deciliter

The model in this exercise can use the following STAN file listed on the `example_models.html` web page:

- `normal.stan` Model to estimate the mean and standard deviation of a normal population

- 2 points: Write R code to read the data and convert it to an R data frame.
- 1 point: Write the data block of a STAN model file that extracts the data from the R workspace.
- 1 point: Write the parameters block of a STAN model file that declares the paramter(s) of your model.
- 2 points: Write the model block of a STAN model file that specifies the priors and likelihood for your model.
- 1 point: Write R code to apply the `extract` function to the data structure output from the `stan` function.
- 1 point: Use the `extract()` function of the RSTAN package to obtain the values for the parameters from the posterior draw.
- 1 point: Use the vector of values of `mu` and `sigma` produced by the `extract` function to generate a sample of 4,000 predicted lead levels with the statement `rnorm(4000,mu,sigma)`.
- 1 point: Use the vector of predicted lead levels to estimate the proportion of children in this population who have a lead level higher than 5 micrograms per deciliter.

(10 points possible)