MTH225 Final Problem 1 Instuctions

The data for this problem is in:

`MTH225_final_problem1.data.R`

The data contains a record of replications of an experiment. The variables in the dataset are:

- `N` the number replications
- `count` is the number of trials in this replication
- `success` is the number of successes in this replication
- `x` is a continuous predictor of success

The input data is in R source code format. Use the command:

```
source("MTH225_final_problem1.data.R")
```

Fit a Bayesian logistic simple regression model with logit parameterization (i.e., use the `binomial_logit` function in STAN).

The logit-transformed probability of success should have the form

$$\beta_0 + \beta_1 x$$

Use a normal (0,5) and a half-cauchy prior with parameter (0,10) for the standard deviation.

You can use the Stan model file `MTH225_final_problem1.stan`

- 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: Create a vector of 4,000 probability values using the mean $\overline{x}$ of the x values in the inverse logit transform:

$$p(x) = 1/(1 + \exp(-\beta_0 - \overline{x} \cdot \beta_1))$$

- 1 point: Use the `quantile` function to construct a 95% credible interval for p.

(10 points possible)