STAT 532 Assignment 3, Fall 15
Due: Monday, Sept. 21st (by 4:15 pm)

Show all work **neatly** and **in order** for full credit and only include computer code and output that is necessary to completely answer a question. Other well organized code and output can be included in the appendix so that I can check your work and provide comments if needed. (i.e. I do not want to have to search through code and output to find your answer or proof you did it). Any plots should appear with the corresponding exercise solution so that I do not have to search through your homework to find them.

I expect you to *not* refer to any solutions you might find or homework's from previous students who have taken the class.

1. (8 pts) Explicitly write out and discuss the prior distributions used in the paper you chose. Do they attempt to incorporate prior information or expert beliefs? If so, to what degree? If they are choosing priors to be "non-informative," what are their reasons for doing this? Do you think their decisions are well justified (based on what you know)? This should be typed and written in complete sentences. If there are multiple parts to the paper, use the part that you have the most questions about the prior distributions.

2. (3 pts) Re-read the *Difficulties with non-informative prior distributions* section on page 54 in the text. Write one sentence providing your main take-away message from the section.

3. (5 pts) Re-read Section 2.9 **Weakly Informative Prior Distributions**
   (a) Do you agree with Gelman's stance on weakly informative priors? Why or why not?
   (b) Use the setting described in Kenny's D2L post under Priors on September 9 to discuss how one might go about specifying a weakly informative prior for that situation.

4. (6 pts) Suppose you are making inference in a Beta-Binomial setting and want to specify a $Beta(a, b)$ prior distribution for π that reflects the following prior knowledge: (1) the median of the distribution is near 0.3 and (2) the probability that π is greater than 0.75 is about 10%. Computationally and/or graphically find values for a and b giving a prior that is consistent with your prior knowledge. Include the values for a and b, a plot of the prior distribution, and the R-code you used to find a and b.

5. The Poisson distribution is often useful for modeling counts of events occurring over time or space. We are usually interested in making inference about the mean of the Poisson distribution, λ.

   (a) (3 pts) Write the Poisson probability mass function in a form that makes it obvious the conjugate prior density for λ is a $Gamma(\alpha, \beta)$ distribution. Point out the connection between properties of exponential family distributions and natural conjugate prior distributions.

   (b) (2 pts) Starting with a Poisson likelihood and a $Gamma(\alpha, \beta)$ prior, derive the posterior distribution for λ.

   (c) (3 pts) Briefly explain how you can think of the parameters of the $Gamma(\alpha, \beta)$ prior in terms of the number of observations added by the prior?

   (d) (4 pts) Based on previous research and your knowledge, suppose you have strong reason to believe that λ is probably between 15 and 25, though values between 10 and 35 are certainly plausible, and there is even a small chance is could be as small as 1 or as big as 50. Specify a reasonable conjugate prior distribution and justify your choice (there is no right answer here, but clearly justify your choice and provide a plot).

   (e) Simulate data by first drawing **a single** random value for λ from your prior distribution and then drawing 20 independent observations from a $Poisson(\lambda)$ distribution. (Make sure you save your data or use set.seed() so that you can repeat the procedure for the same data).

   (f) (2 pts) Provide a plot of your "observed data"' with a vertical line showing the value used for λ.

   (g) (6 pts) Using your prior distribution in (d), obtain the posterior distribution analytically (show work). Plot the likelihood, posterior, and prior together and briefly comment on the plot. Include your plot in the body of the HW and your code in the Appendix.

   (h) (2 pts) Provide a 90% posterior interval for λ. (Show the one line of R-code used)

   (i) (2 pts) What is the probability that λ is between 10 and 20 given the observed data and prior knowledge? (Show R code used)

   (j) (2 pts) What is the probability that λ is less than 5 given the observed data and prior knowledge? (Show R code used)

   (k) (4 pts) Let's show how we can propogate uncertainty from one distribution into another through simulation. Compare and briefly discuss plots that summarize the following two distributions –

      i. Draw 1000 observations from the $Poisson(\lambda)$ distribution you used in part (e) - using your single value of λ.

ii. Draw 1000 realizations of λ from your prior, then for each value of λ draw a value from the Poisson distribution with that mean.

6. For the following suppose we observe 8 successes out of 30 independent trials and we are trying to make inference about the probability of a success.

   (a) (4 pts) Consider the logit transformation for the probability of a success ($g(\theta) = logit(\theta) = \eta$). Find the Jeffreys prior for η up to a proportionality constant. Use R to approximate the normalizing constant and plot the prior distribution.

   (b) (3 pts) Write an R function to calculate the likelihood function for η. Plot the normalized likelihood function on the same plot as the prior distribution from (a).

   (c) (4 pts) Write down the posterior distribution for η up to a proportionality constant. Use R to approximate the normalizing constant using a grid approximation and plot the normalized posterior distribution on the same plot as the likelihood function and the prior distribution. Are you surprised by the results?

   (d) (2 pts) Now, obtain 10000 independent draws from the posterior distribution of π using the Jeffreys prior for π and results we have obtained previously for this conjugate prior. Transform the draws back to the logit scale and plot as a histogram (play around with number of bins - don't just settle for the default). Plot the posterior distribution density you obtained in (c) over the histogram. Briefly comment on the results

   (e) (6 pts) Repeat parts (a)-(d) using a uniform prior on η and the Uniform(0,1) prior for π. Compare results and discuss.

   (f) (3 pts) Logit transform draws from the posterior distribution obtained by using the $Beta(0,0)$ prior and compare to the posterior for η obtained from the uniform prior on η. Compare results and discuss.

   (g) (3 pts) Construct a 92% confidence interval for the probability of a success using non-Bayesian methods. Compare this to the 92% posterior interval obtained using appropriate quantiles from the posterior distributions under the priors the $Beta(0,0)$, $Beta(0.5,0.5)$, and $Beta(1,1)$. Comment/discuss what you find.

7. (3 pts) Draw a diagram (e.g. labeled picture) describing how numerical integration works.

8. Suppose you decide to use the $Beta(0,0)$ prior on the probability of success informed by a study design that justifies use of the binomial distribution with $m = 40$.

   (a) (2 pts) If you were to observe 10 successes, is the resulting posterior distribution proper? Show work.

(b) (2 pts) If you were to observe 0 successes or 40 successes is the resulting posterior distribution proper? Show work

(c) (4 pts) For 0 successes, show how you could use computation to help check if the posterior distribution is proper. Provide nicely organized code and plots. A list of steps to annotate your code might also be helpful.

9. EXTRA CREDIT: Look into one of the software packages meant to help researchers elicit reasonable prior distributions. The options are BEEP, SHELF, or ROBEO. Run through an example and briefly show results and discuss what you learned from it. (Do not spend a lot of time on this unless you really want to)