**Stat 532** **Name:** ______________________
**HW 2**
**Due 9/26/16**

---

Please prepare your solutions using LaTeX or another word processing software.

1. (40 points) Consider the binomial sampling distribution $p(x) = \binom{N}{x} p^x (1-p)^{N-x}$. In classical inference, once you observe data ($x$), you might write down the 95% confidence interval for $p$ as

$$\hat{p} \pm 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}, \tag{1}$$

where $\hat{p} = \frac{x}{N}$. This follows from the asymptotic normality assumption of the estimator $\hat{p}$.

Generate a uniform spacing of $p$'s from .01 to .99, with a step size of .01. Under each value of $p$, generate a Binomial random number with N=30. Now construct the confidence interval specified above. Since you know the actual value of $p$, you can confirm if this interval does in facr cover the *true* parameter.

Repeat this 10,000 times for each value of p, and keep track of the frequency of times each of the 10,000 repetitions cover the known parameter. Show a plot of the behavior of the "true" coverage probabilities for the full range of p's (The y-axis will have frequency on it and the x-axis will be over p).

Repeat the exercise with N = {50,100}. Conclude with your thoughts on the experiment. Are you surprised?

2. Consider $X \sim Binomial(N, \theta)$. Consider the uniform prior $p(\theta) = 1$, where $0 \leq \theta \leq 1$. The posterior sampling distribution should be clear. The prior that you have specified places an equal amount of weight on every possible value of p.

However, some people like to work with the log-odds, which we write as $\Lambda = \log\left(\frac{\theta}{1-\theta}\right)$.

   (a) (5 points) Find $p(\Lambda)$. That is, find the pdf for $\Lambda$ (this is just a simple transformation problem).

   (b) (5 points) Now, do the problem the other way around. Consider placing a uniform prior on $\Lambda$. That is let $p(\Lambda) \propto 1$. What is the implied prior distribution on $\theta$?

   (c) (10 points) Reflect on the results in part a and part b.

3. (Hoff Exercise 3.3.) Tumor counts: A cancer laboratory is estimating the rate of tumorigenesis in two strains of mice, A and B. They have tumor count data for 10 mice in strain A and 13 mice in strain B. Type A mice have been well studied, and information from other labs suggests that type A mice have tumor counts that are approximately Poisson-distributed with a mean of 12. Tumor counts for type B mice are unknown, but type B mice are related to type A mice. The observed tumor counts for the two populations are:

$$\begin{aligned} \boldsymbol{y}_A &= (12, 9, 12, 14, 13, 13, 15, 8, 15, 6) \\ \boldsymbol{y}_B &= (11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7) \end{aligned}$$

   (a) (15 points) Find the posterior distributions, means, variances, and 95% credible intervals for $\theta_A$ and $\theta_B$, assuming a Poisson sampling distribution for each group and the following prior distribution:

$$\theta_A \sim gamma(120, 10), \theta_B \sim gamma(12, 1)$$

   (b) (15 points) Compute and plot the posterior expectation for $\theta_B$ under the prior distribution $\theta_B \sim gamma(12 \times n_0, n_0)$ for each value of $n_o \in \{1, 2, ..., 50\}$. Describe what sort of prior beliefs about $\theta_B$ would be necessary in order for the posterior expectation of $\theta_B$ to be close to that of $\theta_A$.

4. (Hoff Exercise 3.12) Jeffreys' prior: Jeffreys suggested a default rule for generating a prior distribution of a parameter $\theta$ in a sampling model $p(y|\theta)$. Jeffreys' prior is given by $p_j(\theta) \propto \sqrt{I(\theta)}$, where $I(\theta) = -E[\delta^2 \log p(Y|\theta)/\delta\theta^2|\theta]$ is the *Fisher information.*

   (a) (10 points) Let $Y \sim binomial(n, \theta)$. Obtain Jeffreys' prior distribution $p_j(\theta)$ for this model.