STAT6106 Assignment 2

Due on Oct. 18th at 7:30pm

(due at this time so that you can get the solution before the midterm)

three problems from the textbook, starting from page 227:

3.2, 3.3, 3.9

four other problems here:

1. Suppose your prior distribution for the proportion of Californians who support the death penalty has mean 0.6 and standard deviation 0.3. Please find a beta distribution to denote this prior knowledge. Then a random sample of 1000 Californians is taken, and the data shows that 65% support the death penalty. Please now derive the posterior. Plot the prior and posterior on the same plot.

2. (Predictive distributions) There are two coins C1 and C2, Pr(heads|C1)=0.6, Pr(heads|C2)=0.4. Now choose one coin randomly and spin it repeatedly. After you observe that the first two spins are tails, if you continue spin the same coin, what is the expectation of the number of additional spins until a head shows up?

3. A random sample of n students is drawn from a large population. Their average weight is 150 pounds. Assume that the weights in the population are normally distributed with mean θ and standard deviation 20. Suppose your prior for θ is normal with mean 180 and standard deviation 40.

3.1 give your posterior distribution for θ (as a function of n)

3.2 If n=10, give 95% posterior interval for θ

3.3 Using computer to plot the prior, the likelihood (as a function of θ) and posterior in the same plot. Observe the relationship of the 3 curves.

4 . There is a loaded die. The probabilities to get the number {1,2,3,4,5,6} are equal to {$\boldsymbol{p_1}$, $\boldsymbol{p_2}$, $\boldsymbol{p_3}$, $\boldsymbol{p_4}$, $\boldsymbol{p_5}$, $\boldsymbol{p_6}$}, respectively. Obviously, $\sum_{i=1}^{6} p_i = 1$. We want to infer these 6 unknown parameters from observed data in a Bayesian way. We independently tossed this loaded die for 100 times, the number of times we got the number {1,2,3,4,5,6} are {10,10,10,20,10,40}, respectively.

(a). Write down the likelihood function.

(b). What type of distribution is the conjugate prior distribution for the unknown parameters here? Give the distribution name and its density function.

(c). We actually do not have any prior information about $\boldsymbol{p}$. How to choose a specific distribution as your prior from the conjugate distribution family in **(b)**?

(d). Derive the posterior distribution using the observed data.

(e). Provide a posterior point estimate and an uncertainty estimate for each of these unknown parameters.

5. An experiment was performed on the effects of magnetic fields on the flow of calcium out of chicken brains. Measurements on an unexposed group of 32 chickens had a sample mean of 1.013 and sample standard deviation of 0.24. Measurements on exposed group of 36 chickens had sample mean of 1.173 and sample sd of 0.20.

(a) Assuming that the measurements in the control (unexposed) group were from a normal distribution with mean μ_C and standard deviation σ_C, what is the posterior distribution of μ_C? Similarly, what is the posterior distribution of the treatment group μ_t? Assume the prior $p(\mu_c,\mu_t, \sigma^2_c, \sigma^2_t)$ is proportional to $1/(\sigma^2_c\sigma^2_t)$

(b) What's the posterior distribution of μ_t-μ_c? You can obtain independent posterior samples of μ_t and μ_c then plot the histogram of the difference. Obtain an approximate 95% posterior interval for μ_t-μ_c (Hint: use the quantile() function in R)