Stat 532 **Name:** _____
In class midterm
10/18/2017

---

1. Suppose you are interested in determining if MSU students hike more than an average of 10 days a year. To answer this question, you have data on twenty-one students that are randomly selected and polled on their hike behavior (number of days spent hiking).

   (a) (4 points) How would you answer this question using classical statistics techniques?

   (b) (4 points) How would you answer this question from a Bayesian viewpoint?

   (c) (4 points) Describe the inferences that could be made in each case and highlight how they differ.

Short Answer Questions:

For the short answer questions please try to keep your answers to a maximum of 4-5 sentences.

1. (4 points) Given a data set, describe how you would verify that your prior and sampling model are reasonable.

2. (4 points) Describe the steps in conducting a Bayesian data analysis.

3. (4 points) What is the purpose of MCMC and how are MCMC samples used to make inferences?

Computational and Mathematical Questions:

1. (5 points) Assume you have data that can be modeled with a normal distribution. State appropriate priors such that you can implement a Gibbs sampler and then sketch out the algorithm. You do not need to derive the form of the full conditional distributions.

2. (5 points) Suppose you have a kernel of a normal distribution with

$$p(\tilde{\theta}|\tilde{y}, X) \propto \exp\left[-\frac{1}{2}\left(\tilde{y} - X\tilde{\theta}\right)^T \Sigma^{-1}\left(\tilde{y} - X\tilde{\theta})\right)\right],$$

where X is a matrix. What are the mean and variance of this distribution for θ? Note this is a regression setting with a uniform prior on θ.

3. (5 points) Suppose $y_1, \ldots, y_n \sim Bernoulli(\theta)$. You use a $Uniform(0, 1)$ prior for θ, what is the distribution for $p(\theta|y_1, \ldots, y_n)$? (Include the parameters in this distribution).

4. (5 points) Define a Highest Posterior Density region (HPD) and then, using the image and table below that shows the probability for each area, select the 90% HPD region in this setting. Note this should consist of a set of letters (e.g. {A, B, H}).

Area	Probability
A	15
B	18
C	10
D	7
E	6
F	4
G	18
H	22