STAT 638 — Exam 1 **Name:** ______________________________
Fall 2014

INSTRUCTIONS FOR THE STUDENT:

1. You have exactly 1 hour to complete the exam.
2. There are 5 pages including this cover sheet, and 14 questions.
3. Each question is worth 7 points, which means that everyone gets a bonus of two points. **Please circle the letter of the correct answer for each question.**
4. Do not discuss or provide any information to anyone concerning any of the questions on this exam or your solutions until I post the solutions.
5. The only materials you may use are a calculator, an 8 1/2 by 11 inch formula sheet of your own making (with writing on one side only), and a copy of the common distributions on pp. 253-258 of Hoff. **Do not use the textbook or class notes.**

I attest that I spent no more than 1 hour to complete the exam. I used only the allowed materials as described above. I did not receive assistance from anyone during the taking of this exam.

Student's Signature ______________________________

INSTRUCTIONS FOR THE PROCTOR:

(1) Record the time at which the student starts the exam: ____________________

(2) Record the time at which the student ends the exam: ____________________

(3) Immediately after the student completes the exam, please scan the exam to a .pdf file and have the student upload it to WebAssign.

(4) Collect all portions of this exam at its conclusion. Do not allow the student to take any portion of the exam.

(5) Please keep these materials until October 15, 2014, at which time you may either dispose of them or return them to the student.

I attest that the student has followed all the INSTRUCTIONS FOR THE STUDENT listed above and that the exam was scanned into a pdf and uploaded to WebAssign in my presence.

Proctor's Signature ______________________________

1. To a frequentist, data sets that might have been observed but were not are

(a) irrelevant.

(b) relevant, but do not affect how he or she does statistical inference.

(c) used to determine HPD regions.

(d) the basis for defining measures of uncertainty.

(e) the basis for Google's innovative analytics methods.

2. The posterior density for θ in a certain binomial experiment is known to have the form

$$\pi(\theta|y) = C\theta^{19}(1-\theta)^{13}I_{(0,1)}(\theta),$$

where C is a constant. Given that $\Gamma(n+1) = n!$ for all integers $n \geq 1$, the value of C

(a) is $(33 \cdot 32 \cdots 20)/13!$.

(b) is $13!/(33 \cdot 32 \cdots 20)$.

(c) is $(31 \cdot 30 \cdots 19)/12!$.

(d) cannot be determined from the information given.

(e) is quite small given the current state of the US economy.

3. Suppose we are to observe a random sample from the density

$$f(y|\theta) = (\theta+1)2^{-(\theta+1)}y^{\theta}I_{(0,2)}(y),$$

where θ is an unknown parameter that can be any positive number. The Jeffrey's noninformative prior for θ is

(a) proportional to $(\theta+1)^{-1}I_{(0,\infty)}(\theta)$ and improper.

(b) proportional to $(\theta+1)^{-1}I_{(0,\infty)}(\theta)$ and proper.

(c) proportional to $\theta^{-1}I_{(0,\infty)}(\theta)$ and improper.

(d) not attainable from the information given.

(e) impossible to determine given the amount of sleep I had last night.

4. The posterior odds ratio is

(a) the same as the Bayes factor.

(b) equivalent to the likelihood ratio when testing two simple hypotheses against each other.

(c) never used in Bayesian hypothesis testing.

(d) the ratio of posterior probabilities of two hypotheses.

(e) derisively referred to as the "posterior odd ratio" by scornful frequentists.

5. Suppose we have observed a random sample $y_1, \ldots, y_n$ from some distribution depending on unknown parameters. The next value we will observe is Y_{n+1}. Consider the density $m(y_{n+1}|y_1, \ldots, y_n)$, which is the conditional density of Y_{n+1} given $y_1, \ldots, y_n$. This density is useful

(a) for determining the mode of the posterior density.

(b) only for model checking.

(c) only for predicting Y_{n+1}.

(d) both for model checking and predicting Y_{n+1}.

(e) for predicting when donkies will fly.

6. Some people have criticized Bayesian statistics as being too subjective. This criticism can be countered by saying

(a) that subjectivity is ok in certain circumstances.

(b) that one may use a noninformative prior, which is not subjective.

(c) both (a) and (b).

(d) that priors are inherently *objective*.

(e) yo, frequentists, get over it!

7. The random variables X and Y have the following joint density:

$$f(x, y) = \phi(x)\frac{1}{2}\phi\left(\frac{y}{2}\right) \quad \text{for all } x \text{ and } y,$$

where ϕ denotes the standard normal density function. These two random variables are

(a) identically distributed.

(b) exchangeable but not independent.

(c) independent but not exchangeable.

(d) both independent and exchangeable.

(e) BFF.

8. In a binomial experiment that uses a beta(a, b) prior for the unknown success proportion, a nice interpretation of a and b is that

(a) a is like the number of failures in a prior study with $a + b$ trials.

(b) a is like the number of successes in a prior study with $a + b$ trials.

(c) a/b is a prior estimate of the success proportion.

(d) $a - b$ is a prior estimate of the difference between the success and failure proportions.

(e) they have been to every Aggie home game since 1978.

9. An investigator uses a normal distribution as her prior for an unknown parameter θ. She observes a single set of data and finds that the *posterior* distribution of θ is also a normal distribution. In this case

(a) it is definitely true that the normal distribution is a conjugate family for the investigator's likelihood.

(b) it is definitely *not* true that the normal distribution is a conjugate family for the investigator's likelihood.

(c) it might be true that the normal distribution is a conjugate family for the investigator's likelihood.

(d) her prior is noninformative.

(e) the investigator will undoubtedly receive the Nobel prize.

10. A nice property of the posterior distribution is that it

(a) will always coincide strongly with the investigator's prior opinions.

(b) will never be strongly affected by the prior distribution.

(c) depends on the data only through sufficient statistics.

(d) satisfies both (b) and (c).

(e) makes for great dinner conversation when rump roast is on the menu.

11. Given θ, the observations Y_1 and Y_2 are independent and identically distributed Poisson(θ) random variables. A gamma$(2, 1)$ prior is to be used for θ. Suppose that Y_1 and Y_2 are observed to be 0 and 5, respectively. An expression for the posterior probability that θ is closer to 5 than to 0

(a) is $(3^7/\Gamma(7)) \int_{2.5}^{\infty} \theta^6 e^{-3\theta}\, d\theta$.

(b) is $(3^7/\Gamma(7)) \int_{0}^{2.5} \theta^6 e^{-3\theta}\, d\theta$.

(c) is $\int_{2.5}^{\infty} \theta^6 e^{-3\theta}\, d\theta$.

(d) cannot be obtained from the information given.

(e) cannot be obtained from the information given unless your name is Jim Berger.

12. Most Bayesians agree that an improper prior distribution is ok to use only if

(a) it is a Jeffreys prior.

(b) the corresponding posterior is proper.

(c) the corresponding posterior is noninformative.

(d) it is uniform over the entire parameter space.

(e) those in the room are not overly sensitive.

13. Suppose that Y_f and Y are independent given θ. To generate data from the posterior predictive distribution of Y_f given $Y = y$

(a) one must have an explicit expression for $m(y_f|y)$.

(b) it is necessary to know how to generate values from the posterior and values from the distribution of Y_f given θ.

(c) it suffices to know how to generate values from the posterior and values from the distribution of Y_f given θ.

(d) one needs to know how to draw pairs (y, θ) from the joint distribution of Y and θ.

(e) one should draw numbers randomly from a hat.

14. Ten independent and identically distributed observations are obtained from a gamma$(1, 1/\theta)$ density. An inverse-gamma$(1, 1)$ prior is used for θ. The ten observations turned out to have sample mean 4.7. A reasonable Bayesian point estimate for θ

(a) is 4.8.

(b) is 4.

(c) is either (a) or (b).

(d) cannot be determined from the information given.

(e) cannot be computed unless one has access to powerful parallel computing.