Stat 5102 Final Exam **December 19, 2016**

Name ______________________ Student ID ______________

The exam is closed book and closed notes. You may use three $8\frac{1}{2} \times 11$ sheets of paper with formulas, etc. You may also use the handouts on "brand name distributions" and Greek letters. You may use a calculator.

Put all of your work on this test form (use the back if necessary). Show your work or give an explanation of your answer. No credit for answers with no indication of where they came from. Leave no undone integrals, expectations, or derivatives in your answers, but other than that requirement there is no unique "correct" simplification. Any correct answer gets full credit, except as explicitly stated in questions.

Abbreviations used: asymptotic relative efficiency (ARE), generalized linear model or models (GLM), independent and identically distributed (IID), maximum likelihood estimate (MLE), probability density function (PDF).

The points for the questions total to 200. There are 10 pages and 8 problems.

Questions begin on the following page.

1. [25 pts.] Suppose $X_1, \ldots, X_n$ are IID from the distribution having PDF

$$f_\sigma(x) = \sqrt{\frac{\sigma}{2\pi}} \cdot x^{-3/2} e^{-\sigma/(2x)}, \qquad 0 < x < \infty,$$

where σ is an unknown parameter satisfying $0 < \sigma < \infty$.

(a) Find the Jeffreys prior for this distribution.

(b) Say whether it is proper or improper.

2. [25 pts.] The following Rweb output fits three GLM and does tests of model comparison between them

```
Rweb:>  out1 <- glm(y ~ x1 + x2, family = binomial)
Rweb:>  out2 <- glm(y ~ poly(x1, x2, degree = 2),
+      family = binomial)
Rweb:>  out3 <- glm(y ~ poly(x1, x2, degree = 3),
+      family = binomial)
Rweb:>  anova(out1, out2, out3, test = "Chisq")
Analysis of Deviance Table

Model 1: y ~ x1 + x2
Model 2: y ~ poly(x1, x2, degree = 2)
Model 3: y ~ poly(x1, x2, degree = 3)
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1        97     106.97
2        94     105.94  3   1.0270   0.7947
3        90     104.59  4   1.3482   0.8531
```

(a) How many hypothesis tests does this code do?

(b) For each of the tests done, say what is the value of the test statistic, what is the distribution of the test statistic under the null hypothesis (which is used to calculate the P-value), and what is the P-value.

(c) For each of the tests done, say what is the null hypothesis and what is the alternative hypothesis.

(d) For each of the tests done, explain why the degrees of freedom of the reference distribution of the test statistic (named in your answer to part (b)) is what it is.

(e) If one has to choose among the models, which does one choose on grounds of simplicity and statistical significance? Explain.

3. [25 pts.] Suppose $X_1, \ldots, X_n$ are IID from the $\mathcal{N}(\mu, \sigma^2)$ where both μ and σ^2 are considered unknown parameters. Show that there is a two-dimensional sufficient statistic, and identify this statistic.

4. [25 pts.] Suppose $X_1, \ldots, X_n$ are IID with PDF

$$f_{\mu,\sigma}(x) = \frac{405}{2\sigma}\left(6 + \frac{(x-\mu)^2}{\sigma^2}\right)^{-7/2}, \qquad -\infty < x < +\infty$$

which has mean μ and variance $3\sigma^2/2$ (you do not have to show this).

(a) Show that this distribution is symmetric and find the center of symmetry.

(b) Find the asymptotic distribution of the sample mean of an IID sample from this distribution.

(c) Find the asymptotic distribution of the sample median of an IID sample from this distribution.

(d) Find the ARE of the sample mean and sample median of an IID sample from this distribution, both considered as estimators of the center of symmetry (if you do this right, the ARE will not depend on unknown parameters). Also state which estimator is better.

5. [25 pts.] Suppose X is $\text{Bin}(n, p)$. We have only the one observation X. Suppose we are interested in the parameter

$$\theta = \log(p) - \log(1 - p).$$

Find an asymptotic 95% confidence interval for θ. Hint: the 0.975 quantile of the standard normal distribution is 1.9600.

6. [25 pts.] Suppose $X_1, \ldots, X_n$ are IID from the distribution having PDF

$$f_\theta(x) = \frac{1+\theta x^2}{1+\frac{\theta}{3}}, \qquad 0 < x < 1,$$

where the parameter θ satisfies $0 < \theta < \infty$. Find a method of moments estimator of θ.

7. [25 pts.] Suppose $X_1, \ldots, X_n$ are IID Poi(μ) and $Y_1, \ldots, Y_n$ are IID Poi(2μ) and all of the X's are also independent of all of the Y's. Here μ is an unknown parameter satisfying $0 < \mu < \infty$.

(a) Find the MLE of μ.

(b) Show that your solution to part (a) is the unique global maximizer of the likelihood if it is. If you cannot show that it is the unique global maximizer, show that it is a local maximizer.

(c) Calculate expected Fisher information for sample size n.

(d) Give the asymptotic distribution of the MLE.

8. [25 pts.] Suppose X_1, ..., X_n are IID from the distribution with PDF

$$f(x \mid \theta) = \theta^2 x^{-3} e^{-\theta/x}, \qquad 0 < x < \infty,$$

where θ is an unknown parameter satisfying $0 < \theta < \infty$. And suppose we use the $\text{Gam}(\alpha, \lambda)$ prior distribution for θ. Find the posterior distribution for θ. If it is a brand name distribution, then you may give just the name and the values of its hyperparameters. Otherwise give its PDF.