Stat 5102 Final Exam **May 8, 2001**

Name ______________________ Student ID ______________

Secret Code ________________ (if you want your final grade posted).

The exam is open book, including handouts. It is closed notes. You may use one $8\frac{1}{2} \times 11$ sheet of paper with formulas, etc. 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 numbers with no indication of where they came from.

The points for the questions total to 200. There are 4 pages and 8 problems.

1. [25 pts.] Suppose X_1, X_2, $\ldots$, X_n are i. i. d. $\text{Gam}(\theta, \theta)$ random variables, where $0 < \theta$ is an unknown parameter. That is the X_i are $\text{Gam}(\alpha, \lambda)$ with $\alpha = \lambda = \theta$. Find a method of moments estimator of θ.

2. [25 pts.] The function

$$g(y) = \frac{e^y}{(1 + e^y)^2}, \qquad -\infty < y < \infty$$

is a probability density, which is symmetric about zero and has variance

$$\text{var}(Y) = \frac{\pi^2}{3}.$$

Suppose X_1, X_2, $\ldots$, X_n are i. i. d. random variables with density

$$f(x \mid \mu, \sigma) = \frac{1}{\sigma} g\left(\frac{x - \mu}{\sigma}\right), \qquad -\infty < x < \infty,$$

where g is defined above. What is the asymptotic efficiency of the sample median relative to the sample mean as estimators of μ?

3. [25 pts.] Suppose X_1, X_2, $\ldots$, X_n are i. i. d. random variables having mean

$$E(X_i) = \frac{\theta}{2 + \theta}$$

and variance

$$\operatorname{var}(X_i) = \frac{2\theta}{(2+\theta)^2(3+\theta)}$$

and, as usual, $\overline{X}_n$ denotes the sample mean. A method of moments estimator of θ is

$$\hat{\theta}_n = \frac{2\overline{X}_n}{1 - \overline{X}_n}$$

What is the asymptotic distribution of this estimator? You must give the parameters of the asymptotic distribution as functions of θ for full credit.

4. [25 pts.] Suppose X and Y are independent Poisson random variables with means μ_X and μ_Y, respectively. Suppose we observe $X = 320$ and $Y = 410$. Find an asymptotic (large sample) 95% confidence interval for $\mu_Y - \mu_X$.

5. [25 pts.] Suppose X_1, X_2, ..., X_n are i. i. d. Beta$(\theta, 1)$ random random variables. Find the observed and expected Fisher information for estimating θ.

6. [25 pts.] Suppose X_1, X_2, ..., X_n are i. i. d. type II geometric distribution random variables having density

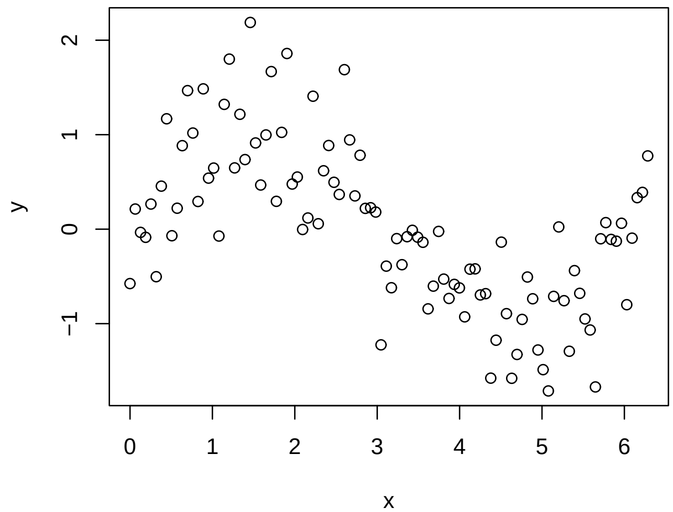
$$f(x) = p(1-p)^x, \qquad x = 0, 1, \ldots$$

(Section B.1.8 of the appendices to the notes) and we want to do a Bayesian analysis with a $\mathcal{U}(0,1)$ prior distribution for p.

   (a) Find the posterior distribution for p, giving it either as a *normalized* probability density function or as a named distribution with specified values of the hyperparameters.

   (b) Find the posterior mean of p.

7. [25 pts.] Several models were fitted to the data for which the scatter plot is shown below.



The output of the R `anova` command was

```
Analysis of Variance Table

Model 1: y ~ x
Model 2: y ~ x + I(x^2)
Model 3: y ~ x + I(x^2) + I(x^3)
Model 4: y ~ x + I(x^2) + I(x^3) + I(x^4)
Model 5: y ~ x + I(x^2) + I(x^3) + I(x^4) + I(x^5)
  Res.Df    RSS Df Sum of Sq        F Pr(>F)
1     98 48.906
2     97 48.851  1     0.054   0.2189 0.6410
3     96 23.442  1    25.410 102.1661 <2e-16 ***
4     95 23.439  1     0.003   0.0109 0.9170
5     94 23.379  1     0.060   0.2428 0.6233
---
Signif. codes:  0 ‘***’  0.001  ‘**’  0.01  ‘*’  0.05  ‘.’  0.1  ‘ ’  1
```

(a) Explain what models are involved in the ANOVA table printout.

(b) Explain why these are nested models.

(c) Explain what conclusions about the fit of these models can be drawn from the printout. Say which models appear to be correct (contain the true regression function) and which appear to be incorrect (do not contain the true regression function).

8. [25 pts.] Show that the type I geometric distribution with density

$$f(y \mid p) = p(1-p)^{y-1}, \qquad x = 1, 2, \ldots.$$

has the form of a GLM model given by equation (12.78) in the notes, which is repeated below

$$f(y \mid \theta, \phi) = \exp\left(\frac{y\theta - b(\theta)}{\phi / w_i} - c(y, \phi)\right)$$

Indicate clearly in your answer

- The canonical parameter θ.
- The dispersion parameter ϕ.
- The weight w.
- The function $b(\theta)$.