ST430 Spring 2002 **Final Exam** May 09, 2001, 11:20am–1:20pm

There are ten questions, with a total of 110 points (but 100 is the full!). Please start a new page for each question and do **not** use both sides. Lastly, enjoy the summer!

1. [Total 15 points]

   Let $Y_1, \ldots, Y_n$ be a random sample from

$$f_Y(y;\theta) = \begin{cases} \frac{3y^2}{\theta^3} & \text{if } 0 < y < \theta, \\ 0 & \text{otherwise.} \end{cases}$$

   (a) [5] Show that $\hat{\theta} = \frac{4}{3}\bar{Y}$ is an unbiased estimator of θ, where $\bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$.

   (b) [5] First find $E(Y_i)$ and $E(Y_i^2)$. Then find $\mathrm{Var}(\hat{\theta})$.

   (c) [5] Find the Cramer–Rao lower bound for $f_Y(y;\theta)$. Compare your answer with $\mathrm{Var}(\hat{\theta})$. Any comments?

2. [Total 8 points]

   Suppose that one observation from the exponential pdf, $f_Y(y) = \lambda e^{-\lambda y}$, $y > 0$, is to be used to test $H_0 : \lambda = 1$ versus $H_1 : \lambda < 1$. The decision rule calls for the null hypothesis to be rejected if $y \geq \ln 10$. Find β as a function of λ.

3. [Total 8 points]

   Assume that $X_1, \ldots, X_m$ are iid $N(\mu, \sigma^2)$ and $Y_1, \ldots, Y_n$ are iid $N(0, \sigma^2)$. Further assume that $X_1, \ldots, X_m$ are independent of $Y_1, \ldots, Y_n$. What is the distribution of

$$S = \frac{1}{\sigma^2}\left\{\sum_{i=1}^n Y_i^2 + \sum_{i=1}^m (X_i - \bar{X})^2\right\},$$

   where $\bar{X} = \frac{1}{m}\sum_{i=1}^m X_i$?

4. [Total 12 points]

   Suppose a random sample $Y_1, \ldots, Y_n$ is to be drawn from $N(\mu, 7^2)$. We wish to test $H_0 : \mu = 50$ vs $H_1 : \mu > 50$ at the $\alpha = 0.05$ level of significance and want $1 - \beta = 0.60$ when $\mu = 52$. What is the smallest sample size (i.e., the smallest n) that will achieve that objective? Hint: first set up two simultaneous equations, one from H_0 while the other from H_1.

5. [Total 10 points]

   The variability in milk production for a 305–day lactation period was observed for a random sample of 15 Holstein cows. Use the milk–yield data below to estimate σ^2, the population variance of milk yields. Attach a 95% confidence interval to your estimate. You may assume that the milk yields follow a normal distribution.

Milk Production Data (in 1,000 pounds)

12.928	13.812	11.036	12.120	14.358
9.248	14.972	8.998	9.980	14.044
10.620	11.990	14.788	14.744	14.786

—— please turn over ——

6. [Total 10 points]

A national survey was conducted to obtain information on the alcohol consumption patterns of U.S. adults by marital status. A random sample of 1772 residents, 18 years old and over, yielded the data displayed below. Conduct a suitable statistical test for testing whether marital status and alcohol consumption are associated or not. Use $\alpha = 0.05$.

marital	drinks per month		
status	Abstain	1 to 60	over 60
single	67	213	74
married	411	633	129
widowed	85	51	7
divorced	27	60	15

7. [Total 10 points]

Is it possible for two random variables X and Y to have the following properties: $E(X) = 3$, $E(Y) = 2$, $E(X^2) = 10$, $E(Y^2) = 29$, and $E(XY) = 0$? Hint: find $\rho(X, Y)$.

8. [Total 12 points]

   (a) [6] Let X and Y be two independent random variables. Show that $\text{Cov}(X, XY) = E(Y)\text{Var}(X)$.

   (b) [6] Let X, Y and Z be three random variables. Show that $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z)$ if Y and Z are independent.

9. [Total 10 points]

Suppose that two quantities Y and x are related by $Y = \beta_0 + \beta_1 x$. A total of 25 pairs of (x_i, y_i) were measured, giving $\sum x_i = 66.075$, $\sum y_i = 50.12$, $\sum x_i^2 = 174.673$, $\sum y_i^2 = 100.499$ and $\sum x_i y_i = 132.491$. Estimate β_0, β_1 and $\text{var}(\hat{\beta}_0)$. Note: when estimating $\text{var}(\hat{\beta}_0)$, use $\frac{n}{n-2}\hat{\sigma}^2$ to estimate σ^2, where

$$\hat{\sigma}^2 = \frac{1}{n}\sum(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = \frac{1}{n}(\sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i).$$

10. [Total 15 points]

Suppose we would like to summarize the data points $(x_1, y_1), \ldots, (x_n, y_n)$ with the equation $y = \beta_1 x$; i.e., a straight line passing through the origin.

   (a) [5] Show that the least–squares estimate for β_1 is $\hat{\beta}_1 = \sum x_i y_i / \sum x_i^2$.

   (b) [5] Show that $\hat{\beta}_1$ is unbiased for β_1.

   (c) [5] Find $\text{Var}(\hat{\beta}_1)$.

—— End of Exam ——