Statistics 252 Midterm #2 – Winter 2005

This exam has 5 problems and is worth 50 points. **Instructor:** Michael Kozdron

You have 50 minutes to complete this exam. Please read all instructions carefully, and check your answers. Show all work neatly and in order, and clearly indicate your final answers. Answers must be justified whenever possible in order to earn full credit. ***Unless otherwise specified, no credit will be given for unsupported answers, even if your final answer is correct.*** *Points will be deducted for incoherent, incorrect, and/or irrelevant statements. Calculators are permitted, as well as an* $8\frac{1}{2} \times 11$ *double-sided page of handwritten notes.*

1. (*12 points*) Suppose that $Y_1, \ldots, Y_n$ are independent and identically distributed $\mathcal{N}(\mu, \sigma^2)$ random variables. It is known that $\sigma^2 = 9$, but μ is unknown. You want to test $H_0 : \mu = 0$ against $H_A : \mu > 0$ by rejecting H_0 when $\overline{Y} > c$.

(a) Find c so that this test has significance level $\alpha = 0.05$. (Of course, c will depend on the sample size n.)

(b) Using the test determined in **(a)**, find the power of the test when $\mu = 1$ and $n = 36$.

(c) Again, using the test determined in **(a)**, show that the power when $\mu = 1$ will increase as the sample size increases. Regardless of whether or not you can show this result, does it make sense intuitively? Comment **very** briefly.

2. (*8 points*) John carries out hypothesis tests using 0.01 as the significance level, while George uses 0.05 as his significance level. Ringo carries out an experiment to compare two hypotheses, and computes the p-value. But all Ringo tells John and George is that the p-value is smaller than 0.03. Can John make his decision to accept or reject the null? If so, what is the decision? What about George? Justify your answers **very** briefly.

3. (*6 points*) In the context of Stat 252, discuss clearly and briefly what is meant by the *duality* of hypothesis tests and confidence intervals.

4. (*12 points*) Suppose that $Y_1, \ldots, Y_n$ are independent and identically distributed with density function

$$f(y|\theta) = \theta^2 \, y \, \exp(-\theta y)$$

where $y > 0$ and $\theta > 0$. As usual, let $\overline{Y}$ denote the sample mean given by

$$\overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i.$$

It is known that the maximum likelihood estimator of θ is $\hat{\theta}_{\text{MLE}} = 2\,/\,\overline{Y}$.

(a) Consider testing $H_0 : \theta = 1$ against $H_A : \theta \neq 1$. Verify that the generalized likelihood ratio for this hypothesis testing problem is

$$\Lambda = \left(\frac{\overline{Y}}{2}\right)^{2n} \exp(2n - n\overline{Y}).$$

(b) Suppose that a random sample of size $n = 5$ is conducted and the observed data yield $\overline{Y} = 1.0$. Perform the generalized likelihood ratio test at the approximate significance level $\alpha = 0.05$.

5. (*12 points*) The goal of this problem is to study a modified linear regression model. Suppose that X and Y are random variables with $\mathbb{E}(X) = \mu_x$, $\text{Var}(X) = \sigma_x^2$, and $\mathbb{E}(Y) = \mu_y$, $\text{Var}(Y) = \sigma_y^2$. Suppose further that $\text{Cov}(X, Y) = \sigma_{xy}$.

Suppose that the random variable $\hat{Y}$ is used to predict Y where $\hat{Y} = \beta_0 + \beta_1 X$, and the parameters β_0 and β_1 are chosen to minimize $\mathbb{E}[(Y - \hat{Y})^2]$.

Show that the minimizing values of β_0 and β_1 are

$$\beta_0 = \mu_y - \beta_1 \mu_x \quad \text{and} \quad \beta_1 = \frac{\sigma_{xy}}{\sigma_x^2}.$$

You do *not* need to check the second derivative test.

Hint: Use the facts that

- $\mathbb{E}[(Y - \hat{Y})^2] = [\mathbb{E}(Y) - \mathbb{E}(\hat{Y})]^2 + \text{Var}(Y - \hat{Y})$;
- $\text{Var}(Y - \hat{Y}) = \text{Var}(Y) + \text{Var}(\hat{Y}) - 2\,\text{Cov}(Y, \hat{Y})$.