Make sure that this examination has 11 numbered pages

University of Regina
Department of Mathematics & Statistics
Final Examination
200710
(April 20, 2007)

Statistics 252-001
Mathematical Statistics

Name: ______________________ **Student Number:** ____________

Instructor: Michael Kozdron **Time:** 3 hours

Read all of the following information before starting the exam.

*You have **3** hours 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.** Several problems require written explanations in context. Only complete solutions written in the context specified by the problem will be awarded full points, and points will be deducted for incoherent, incorrect, and/or irrelevant statements.*

You may use standard notation; however, any new notations or abbreviations that you introduce must be clearly defined.

*Calculators are permitted; however, you must still show all your work. You are also permitted to have **TWO** 8.5×11 pages of handwritten notes (double-sided) for your personal use. Other than these exceptions, no other aids are allowed.*

Note that blank space is not an indication of a question's difficulty. The order of the test questions is essentially random; they are not intentionally written easiest-to-hardest.

*This test has **11** numbered pages with **10** questions totalling **150** points. The number of points per question is indicated.*

DO NOT WRITE BELOW THIS LINE

Problem 1	______	Problem 2	______	Problem 3	______
Problem 4	______	Problem 5	______	Problem 6	______
Problem 7	______	Problem 8	______	Problem 9	______
Problem 10	______				
				TOTAL	______

1. (*24 points*) Suppose that $Y_1, \ldots, Y_n$ are independent and identically distributed Weibull(θ) random variables. That is, each Y_i has density function

$$f_Y(y|\theta) = 2\,\theta^2\, y\, e^{-\theta^2 y^2}, \quad y > 0,$$

where $\theta > 0$ is a parameter.

(*2 pts*) **(a)** Compute the likelihood function $L(\theta)$ for this random sample.

(*6 pts*) **(b)** Show that the maximum likelihood estimator of θ is $\hat{\theta}_{\text{MLE}} = \sqrt{\dfrac{n}{\sum_{i=1}^{n} Y_i^2}}$.

(*4 pts*) **(c)** Show that $\hat{\theta}_{\text{MLE}}$ is a sufficient statistic for the estimation of θ.

(continued)

Recall that $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed with each Y_i having density function

$$f_Y(y|\theta) = 2\,\theta^2\,y\,e^{-\theta^2 y^2}, \quad y > 0,$$

where $\theta > 0$ is a parameter.

(*8 pts*) **(d)** Find the Fisher information $I(\theta)$ in a single observation from this density.

(*4 pts*) **(e)** Using the standard approximation for the distribution of a maximum likelihood estimator based on the Fisher information, construct an approximate 99% confidence interval for θ.

2. (*18 points*) Let Y be a Fisk(θ) random variable so that the density function of Y is

$$f_Y(y|\theta) = \frac{2\,\theta\,y}{(1+\theta y^2)^2}, \quad y > 0,$$

where $\theta > 0$ is a parameter.

(*14 pts*) **(a)** Using an appropriate pivotal quantity, verify that $\left[\dfrac{\alpha}{(2-\alpha)\,Y^2}\,,\ \dfrac{2-\alpha}{\alpha\,Y^2}\right]$ is a confidence interval for θ with coverage probability $1-\alpha$.

(*4 pts*) **(b)** Suppose that a single observation produces $y = 2$. Based on this observation and on the confidence interval you constructed in **(a)**, is there sufficient evidence to reject $H_0 : \theta = 3$ in favour of $H_A : \theta \neq 3$ at the $\alpha = 0.10$ significance level?

3. (*20 points*) Let Y be a random variable whose density function is

$$f_Y(y|\theta) = 1 - \theta^2(y - 1/2), \quad 0 < y < 1,$$

where $-1 < \theta < 1$ is a parameter.

(*8 pts*) **(a)** Based on only this single Y, determine the rejection region of the uniformly most powerful significance level α test of $H_0 : \theta = 0$ vs. $H_A : \theta = 1/2$.

(*8 pts*) **(b)** Determine the power of the test you constructed in **(a)**? (Express your answer in terms of α.)

(*4 pts*) **(c)** Suppose that a single observation produces $y = 0.25$. Based on this observation and on the test you constructed in **(a)**, is there sufficient evidence to reject $H_0 : \theta = 0$ in favour of $H_A : \theta = 1/2$ at the $\alpha = 0.10$ significance level?

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

$$f_Y(y|\theta) = \theta\, y^{\theta-1}, \;\; 0 < y < 1,$$

where $\theta > 0$ is a parameter. Consider testing $H_0 : \theta = 1$ against $H_A : \theta \neq 1$.

(*6 pts*) **(a)** Determine Λ, the generalized likelihood ratio for this hypothesis testing problem. It is known that $\hat{\theta}_{\text{MLE}} = \dfrac{-n}{\sum_{i=1}^{n} \log Y_i}$.

(*6 pts*) **(b)** Suppose that a random sample of size $n = 8$ is conducted and the observed data yield $\sum_{i=1}^{8} \log y_i = -4.0$. Perform the generalized likelihood ratio test for this hypothesis testing problem at the approximate significance level $\alpha = 0.10$.

5. (*8 points*) In order for a Saskatchewan farmer to sell his wheat, he must sell it to the Canadian Wheat Board. A farmer does this by loading his truck and driving it to the nearest grain terminal. He must then wait at the grain terminal until his truck can be unloaded. Suppose that Farmer Mike wants to sell his grain at the Prairie Inland Grain Terminal (PIGT) in the town of Balgonie. It is known that the time taken to unload a truck at the PIGT is well-modelled by a Normal distribution with mean 4 and variance 2; that is, $\mathcal{N}(4, 2)$. It is also known that the times to unload successive trucks are independent. Suppose that Farmer Mike arrives one morning at the PIGT. If he has the third truck in line, what is the probability that Farmer Mike will need to wait more than 13 hours for his truck to be unloaded?

Hint: Let the waiting times of the three trucks be Y_1, Y_2, Y_3, respectively, and then compute $P(Y_1 + Y_2 + Y_3 > 13)$.

6. (*8 points*) The opening price per share Y_1 and Y_2 of two similar stocks are independent random variables, each with density function

$$f_Y(y) = \frac{1}{2} \exp\left\{-\frac{1}{2}(y-4)\right\}, \quad y \geq 4.$$

On a given morning, an investor is going to buy shares of whichever stock is less expensive. Determine the density function of $\min\{Y_1, Y_2\}$ which represents the price per share that the investor will pay.

7. (*10 points*) Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent, identically distributed Uniform(α, β) random variables where *both* α and β are unknown. That is, each Y_i has common density function

$$f_Y(y) = \frac{1}{\beta - \alpha}, \quad \alpha \le y \le \beta,$$

where $-\infty < \alpha < \beta < \infty$ are parameters. Determine the method of moments estimators $\hat{\alpha}_{\text{MOM}}$ and $\hat{\beta}_{\text{MOM}}$.

8. (*24 points*) Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a population with common probability mass function

$$f_Y(y|\theta) := P(Y = y) = \theta^y (1-\theta)^{1-y}, \quad y = 0 \text{ or } y = 1,$$

where $0 \leq \theta \leq 1/2$ is a parameter.

(*4 pts*) **(a)** Find $\hat{\theta}_{\text{MOM}}$, the method of moments estimator of θ.

(*3 pts*) **(b)** Show that $\hat{\theta}_{\text{MOM}}$ is an unbiased estimator of θ.

(*3 pts*) **(c)** Compute the variance of $\hat{\theta}_{\text{MOM}}$.

(*6 pts*) **(d)** Find $\hat{\theta}_{\text{MLE}}$, the maximum likelihood estimator of θ. *Hint*: Maximize $\ell(\theta)$, the log-likelihood function.

(continued)

(*4 pts*) **(e)** It turns out that the same formula for the Fisher information in a single observation can be used in this case. Prove that

$$I(\theta) = \frac{1}{\theta(1-\theta)}.$$

(*4 pts*) **(f)** Prove that $\hat{\theta}_{\text{MOM}}$ is the minimum variance unbiased estimator (MVUE) of θ.

9. (*14 points*) In the context of Stat 252, provide *brief* answers to each of the following questions.

(*7 pts*) **(a)** A researcher who tried to learn statistics without taking a formal course does a hypothesis test and gets a p-value of 0.024. He says, "There is a 97.6% chance that the alternative hypothesis is false, so the null hypothesis is true." What, if anything, is wrong with his statement?

(*7 pts*) **(b)** You perform a hypothesis test using a sample size of four units, and you do not reject the null hypothesis. Your research colleague says this statistical test provides conclusive evidence against the hypothesis. Do you agree or disagree with his conclusion?

10. (*12 points*) Decide whether each of the following statements is either true (**T**) or false (**F**). Clearly circle your choice. You do *not* need to justify your answers.

T or **F** The significance level of a hypothesis test is determined by the null distribution of the test statistic.

T or **F** The statement *"the 95% confidence interval for the population mean is* $[350, 400]$*"* is equivalent to the statement *"there is a 95% probability that the population mean is between 350 and 400."*

T or **F** To reduce the width of a confidence interval by a factor of two (i.e., in half), you have to quadruple the sample size.

T or **F** A p-value of 0.08 is more evidence against the null hypothesis than a p-value of 0.04.

T or **F** If two independent studies are done on the same population with the purpose of testing the same hypotheses, the study with the larger sample size is more likely to have a smaller p-value than the study with the smaller sample size.

T or **F** The statement *"the p-value is* 0.003*"* is equivalent to the statement *"there is a* 0.3% *probability that the null hypothesis is true."*

(*The end. Enjoy your summer!*)