Make sure that this examination has 4 numbered pages

University of Regina
Department of Mathematics & Statistics
Final Examination
201610
(April 27, 2016)

Statistics 252
Introduction to Statistical Inference

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.** Points will be deducted for incoherent, incorrect, and/or irrelevant statements. Unless otherwise noted, you must answer all questions in the test booklets provided.*

*You are permitted to have **ONE** 8.5×11 page of handwritten notes (double-sided) for your personal use. No other aids are allowed. The order of the test questions is essentially random; they are not intentionally written easiest-to-hardest.*

*This test has **4** numbered pages with **8** questions totalling **150** points. The number of points per question is indicated. For questions with multiple parts, all parts are equally weighted.*

1. (*30 points*) Suppose that $Y_1, \ldots, Y_n$ is a random sample from a population having common density function

$$f(y|\theta) = \frac{3y^2}{\theta^3} \exp\left\{-\frac{y^3}{\theta^3}\right\}, \quad y > 0,$$

where $\theta > 0$ is a parameter.

(a) Determine the likelihood function $L(\theta)$ for this random sample.

(b) Show that the maximum likelihood estimator of θ is

$$\hat{\theta}_{\text{MLE}} = \left(\frac{1}{n}\sum_{i=1}^{n} Y_i^3\right)^{1/3}.$$

(c) Use the factorization theorem to show that $\hat{\theta}_{\text{MLE}}$ is a sufficient statistic for the estimation of θ.

(d) Find the Fisher information $I(\theta)$ in a single observation from this density. Note that if $Y \sim f(y|\theta)$, then $\mathbb{E}(Y^3) = \theta^3$.

(e) Using the standard normal approximation for the distribution of a maximum likelihood estimator based on the Fisher information, construct an approximate 95% confidence interval for θ. Note that $z_{0.025} = 1.96$.

2. (*12 points*) Suppose that $Y_1, \ldots, Y_n$ is a random sample from a population having common density function

$$f(y|\theta) = \frac{3y^2}{\theta^3} \exp\left\{-\frac{y^3}{\theta^3}\right\}, \quad y > 0,$$

where $\theta > 0$ is a parameter. Consider testing $H_0 : \theta = \theta_0$ against $H_A : \theta \neq \theta_0$. Determine $\Lambda = \Lambda(Y_1, \ldots, Y_n)$, the generalized likelihood ratio for this hypothesis testing problem. Recall from problem **1. (b)** that the maximum likelihood estimator of θ is

$$\hat{\theta}_{\text{MLE}} = \left(\frac{1}{n}\sum_{i=1}^{n} Y_i^3\right)^{1/3}.$$

3. (*12 points*) Let Y be a random variable whose density function is

$$f(y|\theta) = \frac{3y^2}{\theta^3} \exp\left\{-\frac{y^3}{\theta^3}\right\}, \quad y > 0,$$

where $\theta > 0$ is a parameter. Using an appropriate pivotal quantity, verify that

$$\left[\frac{Y}{(-\log\alpha)^{1/3}}, \frac{Y}{(-\log(1-\alpha))^{1/3}}\right]$$

is a confidence interval for θ with coverage probability $1 - 2\alpha$.

4. (*14 points*) Suppose that $Y_1, \ldots, Y_n$ is a random sample from a Uniform$[0, \theta]$ population so that their common density function is

$$f(y|\theta) = \frac{2y}{\theta^2}, \quad 0 \le y \le \theta,$$

where $\theta > 0$ is a parameter. Let $\hat{\theta} = \max\{Y_1, \ldots, Y_n\}$.

(a) Determine the bias of $\hat{\theta}$.

(b) Determine the mean-squared error of $\hat{\theta}$.

5. (*16 points*) Suppose that Y is a random variable whose density function is

$$f(y|\theta) = \frac{1}{2} - \frac{\theta}{2}(y-1), \ \ 0 < y < 2,$$

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

(a) Based on only this single Y, determine the rejection region of the likelihood ratio test of $H_0 : \theta = 0$ vs. $H_A : \theta = 1/2$ at significance level α.

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

6. (*20 points*) Suppose that a random variable X has a Gamma$(a, 1/b)$ distribution with $a > 0$ and $b > 0$ so that the density function of X is

$$f_X(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, \quad x > 0.$$

(a) Determine the density function of the random variable $1/X$. A random variable with this density function is said to have an *inverse Gamma distribution with parameters a and b.*

Suppose that Y is a $\mathcal{N}(0, \theta)$ random variable where $\theta > 0$ an unknown parameter. Note that θ is the variance of this normal distribution so that

$$f(y \,|\, \theta) = \frac{1}{\sqrt{2\pi\theta}} \exp\left\{-\frac{y^2}{2\theta}\right\}, \quad -\infty < y < \infty.$$

(b) If the prior distribution for θ is an inverse Gamma distribution with parameters a and b, determine an exact, closed-form expression for the posterior density $f(\theta \,|\, y)$.

7. (*24 points*) In the context of Bayesian statistics, it is important to decide on a prior density function. In the 1940s, Sir Harold Jeffreys proposed using a prior density that is proportional to the square root of the Fisher information. Suppose that Y is a Poisson(θ) random variables so that its density function is

$$f(y|\theta) = \frac{\theta^y e^{-\theta}}{y!}$$

for $y = 0, 1, 2, \ldots$ where $\theta > 0$ is a parameter.

(a) Find the Fisher information $I(\theta)$ for the Poisson(θ) density. Recall that if $Y \sim$ Poisson(θ), then $\mathbb{E}(Y) = \mathrm{Var}(Y) = \theta$.

(b) Let $J(\theta) = \sqrt{I(\theta)}$ denote the Jeffreys prior. Observe that $J(\theta)$ does *not* define a legitimate density function. Nonetheless, the usual Bayes' rule formula for the posterior density function using the Jeffreys prior, namely

$$f(\theta \,|\, y) = \frac{f(y \,|\, \theta) J(\theta)}{\displaystyle\int_{-\infty}^{\infty} f(y \,|\, \theta) J(\theta)\, \mathrm{d}\theta} \qquad (*)$$

does define a legitimate density function. Compute $f(\theta \,|\, y)$ as defined by (**??**).

(c) Compute $\hat{\theta}_{\mathrm{BAYES}}$, the Bayes estimator of θ. Recall that

$$\hat{\theta}_{\mathrm{BAYES}} = \mathbb{E}[\theta \,|\, y] = \int_{-\infty}^{\infty} \theta f(\theta \,|\, y)\, \mathrm{d}\theta.$$

8. (*22 points*) Decide whether each of the following statements about hypothesis testing is either true (**T**) or false (**F**). Submit this page with your exam booklet. You do *not* need to justify your answers.

T or **F** The significance level of a hypothesis test is equal to the probability of a type II error.

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

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

T or **F** If a test is rejected at significance level α, then the probability that the null hypothesis is true equals α.

T or **F** The probability that the null hypothesis is falsely rejected is equal to the power of the test.

T or **F** A type II error occurs when the test statistic falls in the rejection region of the test.

T or **F** The p-value of a test is the probability that the null hypothesis is correct.

T or **F** If the p-value is 0.03, then the corresponding test will reject at the significance level 0.03.

T or **F** If a test rejects at significance level 0.06, then the p-value is less than or equal to 0.06.

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.

(*The end. Enjoy your summer!*)