The University of New South Wales
Department of Statistics
Session 2, 2017
MATH5905 - Statistical Inference
Assignment 2

The assignment must be submitted by 5PM on Tuesday, 10th October 2017 at the latest (i.e., at the beginning of the lecture in week 11). Please, declare on the first page that the assignment is your own work, except where acknowledged. State also that you have read and understood the University Rules in respect to Student Academic Misconduct.

1. Let $\mathbf{X} = X_1, X_2, \dots, X_n$ be i.i.d. random variables from a population with a density

$$f(x;\theta) = \begin{cases} \frac{2\theta^2}{x^3}, \text{ if } x \geq \theta \\ 0 \text{ else} \end{cases}$$

where $\theta > 0$ is an unknown parameter

a) Show that $T = X_{(1)}$ is a minimal sufficient statistic for θ.

b) Show that the density of $T = X_{(1)}$ is

$$f_{X_{(1)}}(x;\theta) = \begin{cases} \frac{2n\theta^{2n}}{x^{2n+1}}, \text{ if } x \geq \theta \\ 0 \text{ else} \end{cases}$$

Hint: You may use $P(X_{(1)} \leq x) = 1 - P(X_1 \geq x, X_2 \geq x, \dots, X_n \geq x)$.

c) What is the Maximum Likelihood Estimator of θ? Is it unbiased?

d) Show that $T = X_{(1)}$ is complete. Hence find the UMVUE of θ.

e) Show that the family $\{L(\mathbf{X}, \theta\}, \theta > 0$ has a monotone likelihood ratio in $T = X_{(1)}$.

f) Find the uniformly most powerful α-size test φ^* of $H_0 : \theta \leq 1$ versus $H_1 : \theta > 1$.

g) Find the power function of φ^* and sketch a graph as precisely as possible.

2) Let $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2}, \dots, \mathbf{X_n})$ be i.i.d. random variables, each with a density

$$f(x,\theta) = \begin{cases} \frac{1}{\sqrt{2\pi}x\theta} e^{\{-\frac{1}{2}[\frac{\ln(x)}{\theta}]^2\}}, x > 0 \\ 0 \text{ elsewhere} \end{cases}$$

where $\theta > 0$ is a parameter. (This is called the log-normal density.)

a) Prove that the family $L(\mathbf{X}, \theta)$ has a monotone likelihood ratio in $T = \sum_{i=1}^{n} (\ln X_i)^2$.

b) Argue that there is a uniformly most powerful (UMP) $\alpha-$size test of the hypothesis $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ and exhibit its structure.

c) Using the density transformation formula (or otherwise) show that

$$Y_i = \ln X_i$$

has a $N(0, \theta^2)$ distribution.

Note: *Density transformation formula:* For $Y = W(X)$:

$$f_Y(y) = f_X(W^{-1}(y))|\frac{d(W^{-1}(y)}{dy}| = f_X(x)|\frac{dx}{dy}|.$$

d)Using c) (or otherwise), find the threshold constant in the test and hence determine completely the uniformly most powerful $\alpha-$ size test φ^* of

$$H_0 : \theta \leq \theta_0 \text{ versus } H_1 : \theta > \theta_0.$$

Calculate the power function $E_\theta \varphi^*$ and sketch a graph as precisely as possible.

3. (Use `mathStatica` **only** to solve parts a) and b) of this problem.) During the lab in week 8, capabilities of `mathStatica` to deal with distributions of order statistics have been demonstrated. Even more demonstrations can be found in Section 9.4 of the online `MathStatica` textbook (accessible within MATHEMATICA). Examining the files `5905demo_2017.nb` and `introgra phshort.nb`) in the computing subfolder on moodle and possibly examining the material of Section 9.4, solve the following problems and attach the `mathStatica` printout to your assignment:

a) Let $X_{(1)}, X_{(2)}, X_{(3)}$ denote the order statistics of a random sample of size $n = 3$ from $X \sim N(0, 1)$.

i) Obtain the densities of each of these three order statistics. (Note: $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$)

ii) Plot on a *single* diagram the densities of each order statistic (use the interval (-3.5,3.5)).

iii) Determine $E[X_{(r)}]$ for $r = 1, 2, 3$. In particular, you will check in this way the answer to tutorial question 5b) from the Part four exercise sheet.

b) Repeat the steps i), ii), iii) in a) for the case of order statistic of size $n = 3$ from the standard Laplace distribution with density

$$f(x) = \frac{1}{2} \exp(-|x|), x \in (-\infty, \infty).$$

To get a nicer expression about the density of the order statistic, you might need to use the option `FullSimplify`. When plotting on a single diagram, use the interval (-5,5).

4. (For this problem, you are asked to present your analytic derivations. However, if you present a `mathStatica` solution instead, you will get the marks allocated.)

Suppose $X_{(1)} < X_{(2)} < X_{(3)} < X_{(4)}$ are the order statistics based on a random sample of size 4 from the standard exponential density $f(x) = e^{-x}, x > 0$.

i) Find $E(X_{(3)})$.

ii) Find the density of the midrange $B = \frac{1}{2}(X_{(1)} + X_{(4)})$. Using this result (or otherwise), find $P(B > 1)$.