THE UNIVERSITY OF NEW SOUTH WALES

SCHOOL OF MATHEMATICS AND STATISTICS

Final Exam Term 1 2023

# MATH5905
# Statistical Inference

(1) TIME ALLOWED – 2 HOURS. READING TIME-10 MINUTES

(2) TOTAL NUMBER OF QUESTIONS TO BE ANSWERED – 5

(3) ANSWER ALL QUESTIONS

(4) THE QUESTIONS ARE **NOT** OF EQUAL VALUE

(5) TOTAL NUMBER OF MARKS – 60

(6) USE SEPARATE BOOKLET FOR EACH QUESTION

(7) THIS PAPER MAY BE RETAINED BY THE CANDIDATE

(8) CANDIDATES MAY USE ANY CALCULATOR, RULER, ANY PRINTED NOTES OR TEXTBOOKS

All answers must be written in ink. Except where they are expressly required pencils may only be used for drawing, sketching or graphical work.

1. **[13 marks]** Consider a decision problem with parameter space $\Theta = \{\theta_1, \theta_2\}$ and a set of non randomized decisions $D = \{d_i, 1 \le i \le 6\}$ with risk points $\{R(\theta_1, d_i), R(\theta_2, d_i)\}$ as follows:

| | $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| x | $R(\theta_1, d_i)$ | 0 | 1 | 2 | 4 | 2.5 | 4 |
| y | $R(\theta_2, d_i)$ | 6 | 5 | 3 | 1 | 5 | 4 |

a) **[2 marks]** Find the minimax rule(s) amongst the **nonrandomized** rules $D = \{d_1, d_2, d_3, d_4, d_5, d_6\}$;

b) **[2 marks]** Draw the risk plot of the set of randomized decision rules $\mathcal{D}$ generated by the set of rules in $D$.

c) **[2 marks]** Obtain the minimax rule(s) in the set of **randomized** rules $\mathcal{D}$ generated by the set of rules in $D$. State the minimax risk of the rule(s).

d) **[3 marks]** Find the Bayes rule and the Bayes risk for the prior $(\frac{1}{3}, \frac{2}{3})$ on $(\theta_1, \theta_2)$.

e) **[2 marks]** Express the randomized decision rule with risk point (2,5) using the given non-randomized decision rules.

f) **[2 marks]** Calculate all priors for which $d_1$ is a Bayes rule.

9

**Start a new page clearly marked Question 2**

2. **[12 marks]** Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be independent and identically distributed random variables from a population with a density

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

where $\theta > 0$ is an unknown parameter.

a) **[2 marks]** Show that $T = X_{(1)}$ is a minimal sufficient statistic for $\theta$.

b) **[2 marks]** Show that the density of $T = X_{(1)}$ is

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

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

c) **[2 marks]** Find the maximum likelihood estimator of $\theta$ and provide justification.

d) **[2 marks]** Show that $T = X_{(1)}$ is complete for $\theta$.

e) **[2 marks]** Find the UMVUE of $\theta$.

**Hint:** Calculate $\mathbb{E}(X_{(1)})$ first and start from there.

f) **[2 marks]** Given that the family $\{L(\mathbf{X},\theta\}, \theta > 0$ has a monotone likelihood ratio in $T = X_{(1)}$, find the uniformly most powerful $\alpha$-size test $\varphi^*$ of $H_0 : \theta \leq 1$ versus $H_1 : \theta > 1$.

**Start a new page clearly marked Question 3**

3. **[12 marks]**

Consider $n$ independent identically distribution (i.i.d.) observations from Gamma$(\alpha, \beta)$ distribution.

a) **[2 marks]** Determine the cumulant generating function for a single observation from a Gamma$(\alpha, \beta)$.

**Hint:** The Table of Common Distributions at the end of the exam paper may be helpful.

b) **[2 marks]** Compute the first and second derivative of the cumulant generating function.

c) **[4 marks]** Show that the first order saddlepoint approximation for the density of $\bar{X}$ is

$$\hat{f}(\bar{x}) \approx \sqrt{\frac{n\alpha}{2\pi}}(n\alpha\beta)^{-n\alpha}e^{n\alpha}(n\bar{x})^{n\alpha-1}e^{-n\frac{\bar{x}}{\beta}}n.$$

Justify ALL your steps!

**Hint:** The first order saddlepoint approximation is

$$\hat{f}(\bar{x}) \approx \sqrt{\frac{n}{2\pi K''_X(\hat{t})}}e^{\{nK_X(\hat{t})-n\hat{t}\bar{x}\}}.$$

d) **[2 marks]** Hence, determine the first order saddlepoint approximation for the density of the sum of $n$ iid observations $Y = \sum_{i=1}^{n} X_i$ from this distribution.

**Hint:** Consider using the density transformation formula

$$f_Y(y) = f_X(x(y))\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|.$$

e) **[2 marks]** What is the approximate distribution of the saddlepoint approximation in d)? Keep in mind the *Stirling approximation* of the Gamma function:

$$\sqrt{\frac{2\pi}{n\alpha}}(n\alpha)^{n\alpha}e^{-n\alpha} \approx \Gamma(n\alpha).$$

**Start a new page clearly marked Question 4**

4. **[9 marks]** Suppose $X_{(1)} < X_{(2)} < X_{(3)}$ are the order statistics based on a random sample of size 3 from the standard exponential density $f(x) = e^{-x}, x > 0$.

a) **[4 marks]** Show that $E(X_{(2)}) = 5/6$.

b) **[5 marks]** Show that for the density of the mid-range $B = \frac{1}{2}(X_{(1)} + X_{(3)})$ it holds

$$f_B(b) = 12e^{-2b}(1 - e^{-b})^2, b > 0.$$

**Hint:** The following formulae may be helpful:

$$f_{X_{(r)}}(x) = \frac{n!}{(r-1)!(n-r)!}F(x)^{r-1}[1 - F(x)]^{n-r}f(x)$$

$$f_{X_{(i)},X_{(j)}}(u, v) =$$

$$\frac{n!}{(i-1)!(j-1-i)!(n-j)!}f(u)f(v)[F(u)]^{i-1}[F(v) - F(u)]^{j-1-i}[1 - F(v)]^{n-j}$$

for $-\infty < u < v < \infty$.

6

2x2x1

**Start a new page clearly marked Question 5**

5. **[14 marks]**

Let $X_1, X_2, \ldots, X_n$ be a positive exponentially distributed random sample with the density function

$$f(x, \theta) = \frac{1}{\theta} e^{-x/\theta}, \qquad x > 0$$

where $\theta > 0$ is an unknown parameter.

a) **[2 marks]** Argue that the statistic $T = \sum_{i=1}^{n} X_i$ is a complete and sufficient statistic for $\theta$.

b) **[2 marks]** Determine the UMVUE of $\theta$.

c) **[2 marks]** Determine the MLE of $\theta$.

d) **[1 mark]** The above density can also be parameterized in the form

$$f(x, \eta) = \eta e^{-\eta x}, \qquad x > 0$$

where $\eta = \frac{1}{\theta} > 0$ is an unknown parameter. Determine the MLE of $\eta$.

e) **[2 marks]** Show that the UMVUE of $\eta$ is

$$\hat{\eta}_{\text{umvue}} = \frac{n-1}{\sum_{i=1}^{n} X_i}.$$

**Hint:** Consider $1/\bar{X}$ in your search of an unbiased estimator of $\eta$. Calculate $\mathbb{E}(1/\bar{X})$ and start from there. Note also that

$$T = \sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \theta).$$

f) **[2 marks]** Determine the MLE of $h(\theta) = P(X_1 \geq 1)$.

g) **[3 marks]** Determine the asymptotic distribution of $h(\theta) = P(X_1 \geq 1)$.