STAC58S: 2016 Final Exam

Name:

Student Number:

Time: 3 hours
Instructions: Write your answers on the exam paper. You do not have to fully evaluate numerical results for questions that ask for these as full marks will be given when it is clear that you have the correct expression.
Aids: The exam is open book. Students may use any notes, books and calculators in writing this exam. You are free to refer to any results derived in class or on the assignments.

1. A researcher is interested in grade point averages (GPA) for students enrolled in a Specialist program in the Department of Computer and Mathematical Sciences and in particular, assessing whether or not there is a relationship between GPA and whether or not a student works part-time.

(a) (5 marks) For a statistical study of this problem identify the population, the response measurement, the predictor measurement and the relevant conditional distributions.

(b) (5 marks) Is it possible to definitively answer the question as to whether or not there is a relationship between the variables? If so describe how you would do this.

(c) (5 marks) Explain why we might use continuous probability distributions when modeling the distribution of GPA.

2. Suppose that a statistical model for a response $x \in \{1, 2, 3\}$ is given by the following table.

x	$\theta = a$	$\theta = b$	$\theta = c$
1	1/2	1/3	1/5
2	1/4	1/6	2/5
3	1/4	1/2	2/5

(a) (5 marks) What is the parameter space?

(b) (5 marks) If the value $x = 2$ is observed, what is the MLE of θ and the 0.9-likelihood region?

(c) (5 marks) Now suppose interest is in making inference about

$$\psi = \Psi(\theta) = \begin{cases} 0 & \theta \in \{a, b\}, \\ 1 & \theta \in \{c\}. \end{cases}$$

Determine the profile likelihoods when $x = 1, 2$ and 3 as well as the profile likelihood MLE of ψ.

(d) (5 marks) For the situation described in (c), determine the sampling distribution of the profile likelihood MLE when $\theta = a, b$ and c.

3. (a) (5 marks) Suppose that $x = (x_1, \ldots, x_n)$ is a sample from f_θ where $\theta \in \Theta$. Prove that the order statistic $(x_{(1)}, \ldots, x_{(n)})$ is a sufficient statistic for this model.

(b) (10 marks) Suppose that $x = (x_1, \ldots, x_n)$ is a sample from a Geometric(θ) distribution where $\theta \in (0, 1]$ is unknown. Determine a minimal sufficient statistic for this model and justify your answer. Note that for $x_i \in \mathbb{N}_0$, then $f_\theta(x_i) = \theta(1-\theta)^{x_i}$.

4. Suppose that $x = (x_1, \ldots, x_n)$ is a sample from a $N(\mu, \sigma^2)$ distribution where $\mu \in R^1, \sigma^2 > 0$ are unknown.

(a) (5 marks) Let $\mu + \sigma z_p$ denote the p-th quantile of the $N(\mu, \sigma^2)$ distribution. What does this mean?

(b) (10 marks) Determine a UMVU estimate of $\mu + \sigma z_p$ and justify your answer.

(c) (5 marks) What is $\frac{1}{n}\sum_{i=1}^{n} I_{(-\infty,c]}(x_i)$ an unbiased estimator of? Indicate (no need to do the computation) how you would obtain the UMVU estimator.

(d) (5 marks) Suppose an investigator wants to assess the hypothesis $H_0 : \mu = 0$ and uses the statistic $T(\bar{x}, s^2) = \sqrt{n}\bar{x}/s$ for this purpose via a p-value where $s = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$. Explain how you would do this. What is the distribution of the p-value when H_0 is true?

5. Suppose that a statistical model is given by the two distributions in the following table.

	$x = 1$	$x = 2$	$x = 3$	$x = 4$
$f_a(x)$	$1/3$	$1/6$	$1/12$	$5/12$
$f_b(x)$	$1/2$	$1/4$	$1/6$	$1/12$

(a) (5 marks) Determine the UMP size 0.10 test for testing $H_0 : \theta = a$ versus $H_a : \theta = b$. What is your conclusion when you observe $x = 3$?

(b) (5 marks) What is the power of the test determined in (a)?

6. Suppose that the statistical model is given by

	$f_\theta(1)$	$f_\theta(2)$	$f_\theta(3)$	$f_\theta(4)$
$\theta = a$	1/3	1/6	1/3	1/6
$\theta = b$	1/2	1/4	1/8	1/8

and that the priors is given by $\pi(a) = 1/3, \pi(b) = 2/3$ and we observe the sample $(x_1, x_2, x_3) = (1, 1, 3)$.

(a) (5 marks) Determine the posterior of θ.

(c) (5 marks) Determine the MAP estimate of θ.

(d) (5 marks) Determine the relative belief estimate of θ.