

McGill University

Faculty of Science

MATH 557

MATHEMATICAL STATISTICS II

Final Examination

Date: 14th April 2008

Time: 2pm-5pm

Examiner: Dr. David A. Stephens

Associate Examiner: Dr. Masoud Asgharian

Please write your answers in the answer booklets provided.

This paper contains five questions. Each question carries 20 marks. Credit will be given for all questions attempted. The total mark available is 100 but rescaling of the final mark may occur.

Candidates may take one double-sided sheet of Letter-sized (216×279 mm) or A4-sized (210×297 mm) paper with handwritten notes into the examination room.

Calculators may be used. A formula sheet is provided.

Dictionaries and Translation dictionaries are permitted.

1. (a) Write down the form of the joint probability density function (pdf) or probability mass function (pmf) for a random sample $X_1, \dots, X_n$ from a one-parameter Exponential Family distribution parameterised by natural parameter $\theta \in \Theta \subseteq \mathbb{R}$.

2 MARKS

- (b) Identify a complete minimal sufficient statistic for θ , giving definitions of each of the terms *complete* and *minimal sufficient*.

3 MARKS

- (c) Identify whether or not each of the following are Exponential Family Distributions. For those that are Exponential Family distributions, write the joint pmf/pdf in the form from part (a) for a random sample $X_1, \dots, X_n$, and find a complete minimal sufficient statistic. For those that are not Exponential Family distributions, find a minimal sufficient statistic. Justify your answers in each case.

- (i) $X \sim \text{Uniform}(-\theta, \theta)$ where $\theta > 0$

3 MARKS

- (ii) For $x > \theta$,

$$f_{X|\theta}(x|\theta) = \exp\{-(x - \theta)\}$$

where $\theta > 0$.

3 MARKS

- (iii) For $0 < x < 1$,

$$f_{X|\theta}(x|\theta) = 2 \frac{x + \theta}{1 + 2\theta}$$

where $\theta > 0$.

3 MARKS

- (iv) For $0 < x < 1$,

$$f_{X|\theta}(x|\theta) = \frac{\theta^x \ln \theta}{\theta - 1}$$

where $\theta > 0$.

3 MARKS

- (d) Show that if statistic $T(\underline{X})$ is sufficient for θ , then the Bayesian posterior distribution for θ , $\pi_{\theta|\underline{X}}(\theta|\underline{x})$, depends on $x_1, \dots, x_n$ only through $T(\underline{x})$.

3 MARKS

2. (a) Suppose that $X_1, \dots, X_n$ is a random sample from probability model with pmf/pdf $f_{X|\theta}(x|\theta)$. State a necessary and sufficient condition for an estimator $T = T(\underline{X})$ of univariate parameter θ to attain the Cramér-Rao lower bound on variance.

6 MARKS

- (b) Suppose that $X_1, \dots, X_n$ are a random sample from a Poisson(θ) distribution.

- (i) Is the sample mean $\bar{X}$ the *best unbiased estimator (minimum variance unbiased estimator)* of θ ? Justify your answer.

2 MARKS

- (ii) Let $\phi = \theta(\theta + 2)$. Find the Fisher information for ϕ , $\mathcal{I}(\phi)$.

3 MARKS

- (iv) Show that there does not exist an unbiased estimator of ϕ whose variance attains the Cramér-Rao lower bound for ϕ .

4 MARKS

- (v) Find a random variable T (a function of $\bar{X}$), which is unbiased for ϕ and show that

$$\text{Var}_{f_{T|\theta}}[T|\theta] > \frac{4\theta(\theta + 1)^2}{n}.$$

Hint: Compute the expectation of $\bar{X}^2$.

5 MARKS

3. (a) State and prove the Rao-Blackwell Theorem that relates the concepts of sufficiency, unbiasedness and the variance of estimators.

8 MARKS

- (b) Each of the two independent, identically distributed observations X and Y has an Exponential distribution with unknown parameter θ .

- (i) Show that $S = X + Y$ is a sufficient statistic for θ and state the distribution of S .

2 MARKS

- (ii) Given a fixed value $x > 0$, define statistic T_x , a function of statistic S , by

$$T_x(S) = \begin{cases} \frac{x}{S} & S > x, \\ 1 & \text{otherwise.} \end{cases}$$

Show, by direct calculation or otherwise, that

$$\text{E}_{f_{T_x|\theta}}[T_x|\theta] \equiv \text{E}_{f_{S|\theta}}[T_x(S)|\theta] = \tau(\theta)$$

where $\tau(\theta) = F_{X|\theta}(x|\theta)$, the cdf of X evaluated at x .

4 MARKS

- (iii) It can be shown that S is a complete statistic. Show that $T_x(S)$ is the unique function of S which is unbiased for $\tau(\theta)$, quoting without proof any relevant results.

3 MARKS

- (iv) Explain the relevance of part (iii) to optimal estimation of $\tau(\theta)$.

3 MARKS

4. (a) State the Neyman-Pearson Lemma.

4 MARKS

- (b) Suppose that the observations $X_1, \dots, X_n$ form a random sample from the distribution with probability density function

$$f_{X|\theta}(x|\theta) = 2\theta x e^{-\theta x^2} \quad x > 0,$$

and zero otherwise, for parameter $\theta > 0$. Suppose that a test of hypotheses

$$H_0 : \theta = \frac{1}{2}$$

$$H_1 : \theta > \frac{1}{2}$$

is required. Find the precise form of the uniformly most powerful (UMP) test at level α .

Hint: Find the distribution of X^2 .

6 MARKS

- (c) For the random sample in (b), find the maximum likelihood estimator $\hat{\theta}_n$, of θ , and a large sample approximation to its distribution.

6 MARKS

Using the large sample approximation, find an approximate $(1 - \alpha)$ confidence interval for θ .

4 MARKS

5. (a) Prove that, under regularity conditions **to be stated**, if the true value of parameter θ in a probability model, $f_{X|\theta}(x|\theta)$, is θ_0 , then

$$\Pr[L(\theta_0|\underline{X}) > L(\theta|\underline{X})|\theta_0] \longrightarrow 1$$

as $n \longrightarrow \infty$, where $L(\theta|\underline{X})$ is the likelihood for θ .

8 MARKS

- (b) Consider the probability model with pdf

$$f_{X|\theta}(x|\theta) = \frac{e^{-(x-\theta)}}{(1 + e^{-(x-\theta)})^2} \quad -\infty < x < \infty$$

for parameter $\theta \in \mathbb{R}$.

- (i) Does an efficient estimator for θ exist for this model? Justify your answer.

3 MARKS

- (ii) Write down a consistent but possibly inefficient estimator for θ .

2 MARKS

- (iii) Explain briefly the idea of *one-step estimation* for producing efficient estimators, illustrating the procedure in the model above.

4 MARKS

- (iv) Show that the Fisher information for θ , $\mathcal{I}(\theta)$, is constant in θ .

3 MARKS

DISCRETE DISTRIBUTIONS							
	RANGE $\mathbb{X}$	PARAMETERS	MASS FUNCTION f_X	CDF F_X	$E_{f_X} [X]$	$\text{Var}_{f_X} [X]$	MGF M_X
$Bernoulli(\theta)$	$\{0, 1\}$	$\theta \in (0, 1)$	$\theta^x(1 - \theta)^{1-x}$		θ	$\theta(1 - \theta)$	$1 - \theta + \theta e^t$
$Binomial(n, \theta)$	$\{0, 1, ..., n\}$	$n \in \mathbb{Z}^+, \theta \in (0, 1)$	$\binom{n}{x} \theta^x (1 - \theta)^{n-x}$		$n\theta$	$n\theta(1 - \theta)$	$(1 - \theta + \theta e^t)^n$
$Poisson(\lambda)$	$\{0, 1, 2, ... \}$	$\lambda \in \mathbb{R}^+$	$\frac{e^{-\lambda} \lambda^x}{x!}$		λ	λ	$\exp \{ \lambda (e^t - 1) \}$
$Geometric(\theta)$	$\{1, 2, ... \}$	$\theta \in (0, 1)$	$(1 - \theta)^{x-1} \theta$	$1 - (1 - \theta)^x$	$\frac{1}{\theta}$	$\frac{(1 - \theta)}{\theta^2}$	$\frac{\theta e^t}{1 - e^t(1 - \theta)}$
$NegBinom(n, \theta)$	$\{n, n + 1, ... \}$	$n \in \mathbb{Z}^+, \theta \in (0, 1)$	$\binom{x-1}{n-1} \theta^n (1 - \theta)^{x-n}$		$\frac{n}{\theta}$	$\frac{n(1 - \theta)}{\theta^2}$	$\left(\frac{\theta e^t}{1 - e^t(1 - \theta)} \right)^n$
or	$\{0, 1, 2, ... \}$	$n \in \mathbb{Z}^+, \theta \in (0, 1)$	$\binom{n+x-1}{x} \theta^n (1 - \theta)^x$		$\frac{n(1 - \theta)}{\theta}$	$\frac{n(1 - \theta)}{\theta^2}$	$\left(\frac{\theta}{1 - e^t(1 - \theta)} \right)^n$

For **CONTINUOUS** distributions (see over), define the **GAMMA FUNCTION**

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

and the **LOCATION/SCALE** transformation $Y = \mu + \sigma X$ gives

$$f_Y(y) = f_X\left(\frac{y - \mu}{\sigma}\right) \frac{1}{\sigma} \qquad F_Y(y) = F_X\left(\frac{y - \mu}{\sigma}\right) \qquad M_Y(t) = e^{t\mu} M_X(\sigma t) \qquad E_{f_Y} [Y] = \mu + \sigma E_{f_X} [X] \qquad \text{Var}_{f_Y} [Y] = \sigma^2 \text{Var}_{f_X} [X]$$

CONTINUOUS DISTRIBUTIONS

	PARAMS.	PDF	CDF	$E_{f_X} [X]$	$Var_{f_X} [X]$	MGF
$\mathbb{X}$						
$Uniform(\alpha, \beta)$ (standard model $\alpha = 0, \beta = 1$)	$\alpha < \beta \in \mathbb{R}$	$\frac{1}{\beta - \alpha}$	$\frac{x - \alpha}{\beta - \alpha}$	$\frac{(\alpha + \beta)}{2}$	$\frac{(\beta - \alpha)^2}{12}$	$M_X = \frac{e^{\beta t} - e^{\alpha t}}{t(\beta - \alpha)}$
$Exponential(\lambda)$ (standard model $\lambda = 1$)	$\lambda \in \mathbb{R}^+$	$\lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\left(\frac{\lambda}{\lambda - t}\right)$
$Gamma(\alpha, \beta)$ (standard model $\beta = 1$)	$\alpha, \beta \in \mathbb{R}^+$	$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$		$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\left(\frac{\beta}{\beta - t}\right)^\alpha$
$Weibull(\alpha, \beta)$ (standard model $\beta = 1$)	$\alpha, \beta \in \mathbb{R}^+$	$\alpha \beta x^{\alpha-1} e^{-\beta x^\alpha}$	$1 - e^{-\beta x^\alpha}$	$\frac{\Gamma(1 + 1/\alpha)}{\beta^{1/\alpha}}$	$\frac{\Gamma(1 + 2/\alpha) - \Gamma(1 + 1/\alpha)^2}{\beta^{2/\alpha}}$	
$Normal(\mu, \sigma^2)$ (standard model $\mu = 0, \sigma = 1$)	$\mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\}$		μ	σ^2	$e^{\{\mu t + \sigma^2 t^2 / 2\}}$
$Student(\nu)$	$\nu \in \mathbb{R}^+$	$\frac{(\pi\nu)^{-\frac{1}{2}} \Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \left\{1 + \frac{x^2}{\nu}\right\}^{(\nu+1)/2}}$		0 (if $\nu > 1$)	$\frac{\nu}{\nu - 2}$ (if $\nu > 2$)	
$Pareto(\theta, \alpha)$	$\theta, \alpha \in \mathbb{R}^+$	$\frac{\alpha \theta^\alpha}{(\theta + x)^{\alpha+1}}$	$1 - \left(\frac{\theta}{\theta + x}\right)^\alpha$	$\frac{\theta}{\alpha - 1}$ (if $\alpha > 1$)	$\frac{\alpha \theta^2}{(\alpha - 1)(\alpha - 2)}$ (if $\alpha > 2$)	
$Beta(\alpha, \beta)$	$\alpha, \beta \in \mathbb{R}^+$	$\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1 - x)^{\beta-1}$		$\frac{\alpha}{\alpha + \beta}$	$\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$	