McGill University

Faculty of Science

MATH 556

MATHEMATICAL STATISTICS I

Final Examination

Date: 4th December 2008 Time: 9am-12pm

Examiner: Prof. D. A. Stephens Associate Examiner: Prof. R. J. Steele

This paper contains six questions.

Credit will be given for all questions attempted.

Calculators may not be used. A Formula Sheet is provided.

1. (a) Suppose that X_1 and X_2 are independent random variables with $Gamma\,(\alpha_1, 1)$ and $Gamma\,(\alpha_2, 1)$ distributions respectively, for parameters $\alpha_1, \alpha_2 > 0$. Let

$$Y_1 = \frac{1}{X_1} \qquad\qquad Y_2 = \frac{1}{X_2}$$

(i) Find the pdf of Y_1.

4 MARKS

(ii) Find the expectation and variance of Y_2. State explicitly conditions for the expectation and variance to exist.

6 MARKS

The distribution of Y_1 and Y_2 is termed the *inverse-Gamma* distribution.

(b) Find the joint pdf of random variables

$$Z_1 = X_1 + X_2 \qquad\qquad Z_2 = Y_1 + Y_2$$

and show that marginally Z_1 has a Gamma distribution.

5 MARKS

Marginally, does Z_2 have an inverse-Gamma distribution ? Justify your answer.

5 MARKS

2. (a) For a scalar random variable X, the *cumulant generating function* (cgf), K_X, is defined in terms of the moment generating function, M_X, by

$$K_X(t) = \log M_X(t) \qquad t \in (-h, h), \text{ some } h > 0$$

Find expressions for the expectation and variance of X in terms of K_X.

6 MARKS

(b) For random variable X, consider the one parameter Exponential Family distribution,

$$f_X(x|\eta) = h(x)c(\eta)\exp\{\eta\, x\}$$

Find the form of $K_X(t)$, and an expression for $\mathbb{E}_{f_X}[X]$ in terms of one or more of the functions and parameters that appear in the pdf.

8 MARKS

(c) Show how the cumulant generating function plays a role in the *exponential tilting construction* of the Exponential Family of distributions. Illustrate your description with specific reference to the $Normal(\theta, 1)$ distribution.

6 MARKS

3. (a) Consider scalar random variable X.

(i) Define the *characteristic function* of X, $C_X(t)$.

2 MARKS

(ii) Show that $|C_X(t)| \leq 1$ for all $t \in \mathbb{R}$.

2 MARKS

(iii) Describe how to diagnose whether a characteristic function corresponds to a discrete or a continuous distribution.

4 MARKS

(iv) State the *inversion formula* for a characteristic function known to belong to a **discrete** distribution.

2 MARKS

(b) Find $C_X(t)$ if X is a continuous random variable with

$$f_X(x) = \exp\{-x - e^{-x}\} \qquad x \in \mathbb{R}.$$

5 MARKS

(c) Find $f_X(x)$ if $C_X(t)$ is given by

$$C_X(t) = 1 - |t| \qquad -1 < t < 1$$

and zero otherwise.

5 MARKS

4. (a) State and prove Jensen's Inequality in the univariate case. Define the Kullback-Leibler divergence between two pdfs f_1 and f_2 each with support $\mathbb{R}$, $\mathbb{K}(f_1, f_2)$, and prove that

$$\mathbb{K}(f_1, f_2) \geq 0.$$

10 MARKS

(b) Suppose that $X \sim Normal(0, \sigma^2)$. Find a function of σ, $l(\sigma)$, such that

$$P_X[\,-2 \leq X \leq 2\,] \geq \, l(\sigma)$$

which does **not** involve the standard normal cdf. Justify your answer.

4 MARKS

(c) If $X \sim Poisson(\mu)$, show that

$$P_X[\,X \geq \, 2\mu\,] \leq e^{\mu(e-3)}$$

Justify your answer.

6 MARKS

5. (a) A *finite mixture model* is a specific form of hierarchical model whose density, f_Y, takes the form

$$f_Y(y|\underset{\sim}{\pi}, \underset{\sim}{\theta}, L) = \sum_{l=1}^{L} f_l(y|\theta_l)\pi_l.$$

Explain the components of this specification, that is, the functions $f_1, \ldots, f_L$, the parameters $\theta_1, \ldots, \theta_L$, and the parameters $\pi_1, \ldots, \pi_L$.

4 MARKS

Find the form of the moment generating function, M_Y, corresponding to f_Y.

4 MARKS

(b) Suppose that $X_1, \ldots, X_n$ is a random sample from a $Normal(\mu, 1)$ distribution. Find the distribution of the statistic

$$V = \sum_{i=1}^{n} (X_i - \overline{X})^2$$

You may quote without proof results concerning the sampling distribution of statistics derived from a normal random sample.

8 MARKS

(c) Suppose that $X_1, \ldots, X_n$ is a random sample from an $Exponential(\lambda)$ distribution. Find the distribution of the statistic

$$Y_1 = \min\{X_1, \ldots, X_n\}$$

You may quote without proof results concerning the sampling distribution of order statistics derived from a random sample.

4 MARKS

6. (a) Suppose $\{X_n\}$ are an independent sequence of random variables with cdf

$$F_X(x) = \frac{1}{1 + e^{-x}} \qquad x \in \mathbb{R}$$

Let $Y_n = \max\{X_1, \ldots, X_n\}$. Show that for large n and $y > 0$,

$$P[Y_n > y] \simeq 1 - \exp\{-ne^{-y}\}.$$

6 MARKS

(b) Suppose $\{X_n\}$ are an independent sequence of $Poisson(\lambda)$ random variables. Let

$$M_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Show that $M_n \xrightarrow{p} \lambda$ as $n \longrightarrow \infty$. If random variable T_n is defined by $T_n = e^{-M_n}$, find an approximation to the distribution of T_n for large n.

8 MARKS

(c) Suppose that $X \sim Gamma(\alpha, 1)$. By considering the variable

$$Z_\alpha = \frac{X - \alpha}{\sqrt{\alpha}}$$

or otherwise, construct an approximation to the distribution of Z_α for large α.

6 MARKS

DISCRETE DISTRIBUTIONS

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

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

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

CONTINUOUS DISTRIBUTIONS

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