Name: ______________________ ID: ______________

McGill University

Faculty of Science

Final Examination

MATH 556: Mathematical Statistics I

Examiner: Professor J. Nešlehová

Associate Examiner: Professor D. A. Stephens

Date: Tuesday, December 7, 2010

Time: 9:00 A.M. – 12:00 P.M.

Instructions

- **This is a closed book exam.**
- **Answer all six questions in the examination booklets provided.**
- **Calculators and translation dictionaries are permitted.**
- **A formula sheet is provided.**

Good Luck!

Problem 1

Recall that the $F_{p,q}$ distribution has density

$$f(x|p,q) = \frac{\Gamma\left(\frac{p+q}{2}\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)} \left(\frac{p}{q}\right)^{p/2} x^{p/2-1} \left(1 + \left(\frac{p}{q}\right) x\right)^{-(p+q)/2}, \quad x > 0.$$

(a) Suppose that X has an $F_{p,q}$ distribution. Find the distribution of

$$Y = \frac{pX}{q + pX}$$

and compute its expectation. **(5 marks)**

(b) Suppose that U and V are independent random variables with densities f_U and f_V, respectively. Prove that the density of $W = U/V$ is given by

$$f_W(w) = \int_{-\infty}^{\infty} |v| f_U(wv) f_V(v) \mathrm{d}v.$$

(4 marks)

(c) Prove that if $U \sim \chi^2_p$ and $V \sim \chi^2_q$ are independent,

$$Z = \frac{U/p}{V/q} \sim F_{p,q}.$$

(4 marks)

(d) Suppose that S_1^2 and S_2^2 are the sample variances of the random samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$, respectively. State all the necessary conditions under which the ratio S_1^2/S_2^2 has an $F_{p,q}$ distribution. **(4 marks)**

Problem 2

Let X and Y be independent, $X \sim \text{Gamma}(\alpha, 1)$ and $Y \sim \text{Gamma}(\beta, 1)$ and define

$$T = X + Y, \quad Z = \frac{X}{X+Y}, \quad W = \frac{Y}{X+Y}.$$

(a) Compute the joint distribution of (T, Z). **(5 marks)**

(b) Compute the (marginal) distributions of Z and W. **(4 marks)**

(c) Compute the correlation coefficients $\text{cor}(T, Z)$ and $\text{cor}(Z, W)$. What can you say about the independence of T and Z, and of Z and W, respectively? *Hint: computing the joint distribution of (Z, W) is NOT necessary.* **(4 marks)**

Problem 3

Consider the Pareto family with densities

$$f(x|\alpha) = \alpha \left(\frac{1}{1+x} \right)^{\alpha+1}, \quad x > 0.$$

(a) Show that the family $f(x|\alpha)$ is an exponential family. Determine the natural parametrization and the natural parameter space. **(4 marks)**

(b) Suppose that X has density $f(x|\alpha)$. Compute the mean and the variance of $\log(X+1)$. **(3 marks)**

(c) Explain how a new exponential family $g(x|t)$ can be constructed from some arbitrary density g by exponential tilting. **(3 marks)**

(d) Can a new exponential family be constructed by tilting of the Pareto density with some $\alpha > 0$? If yes, give it, if not, explain why. **(4 marks)**

(e) Give an example of a family of distributions which is not an exponential family. Provide a thorough explanation for your choice. **(4 marks)**

Problem 4

(a) Suppose that each of a random number $N \sim \text{Poisson}(\lambda)$ of independent patients is testing a drug. For each patient, the success of a drug is described by a Bernoulli variable X_i, independent of N. Because the patients are different, we are reluctant to assume that the success probabilities are constant. Instead, we assume that $X_i|P_i$ is Bernoulli(P_i), where $P_i \sim \text{Beta}(\alpha, \beta)$ for some fixed parameters $\alpha > 0$, $\beta > 0$, and $P_1, P_2, \ldots$ are independent.

(1) Determine the distribution of Y. **(5 marks)**

(2) Compute the mean and variance of the unconditional distribution of the total number of successes, $Y = \sum_{i=1}^{N} X_i$. **(4 marks)**

(b) Let X and Y be two random variables with finite variances.

(1) Show that X and $Y - \text{E}(Y|X)$ are uncorrelated. **(4 marks)**

(2) Show that $\text{var}(Y - \text{E}(Y|X)) = \text{E}(\text{var}(Y|X))$. **(4 marks)**

Problem 5

Recall without proof the MGF and the mean and variance of the NegBinomial(r, p) distribution with parameters $p \in (0, 1)$ and $r \in \mathbb{N}$, as given on the formula sheet.

(a) Let $X_1, \dots, X_n$ be a random sample from the NegBinomial(r, p) distribution. Determine the distribution of $\overline{X}_n = (1/n)(X_1 + \cdots + X_n)$. **(4 marks)**

(b) Prove Jensen's inequality, that is, for any random variable Y with finite expectation and a convex function g such that $\mathrm{E}|g(X)| < \infty$, $\mathrm{E}g(X) \geq g(\mathrm{E}X)$. You can use, without proof, the fact that the one-sided derivatives of any convex function exist everywhere. **(5 marks)**

(c) In the context of part (a), consider the statistic

$$T_n = \frac{r}{\overline{X}_n + r}$$

Show that $\mathrm{E}(T_n) \geq p$, and that, at the same time, $T_n \to p$ in probability as $n \to \infty$.

(4 marks)

(d) Show how the distribution function of an NegBinomial(r, p) random variable can be approximated by the distribution function of the standard normal distribution for r large.

(5 marks)

Problem 6

Consider an i.i.d. sequence $X_1, X_2, \dots$ from the distribution function

$$F(x|\alpha) = 1 - (1 - x)^\alpha, \quad x \in (0, 1),$$

where $\alpha > 0$ is a parameter.

(a) Show that the distribution function of

$$M_n = \max(X_1, \dots, X_n)$$

is given by $(1 - (1 - x)^\alpha)^n$ if $x \in (0, 1)$. **(3 marks)**

(b) Prove that if a sequence $\{Y_n\}$ of arbitrary random variables satisfies $Y_n \rightsquigarrow a$ as $n \to \infty$ where $a \in \mathbb{R}$ is a constant, then $Y_n \to a$ in probability as $n \to \infty$. **(5 marks)**

(c) Prove that $M_n \to 1$ in probability as $n \to \infty$. **(4 marks)**

(d) Does $n^{1/\alpha}(M_n - 1)$ converge in distribution as $n \to \infty$? If yes, determine the limiting distribution, if not, explain why. **(5 marks)**

DISCRETE DISTRIBUTIONS

	RANGE $\mathbb{X}$	PARAMETERS	MASS FUNCTION f_X	CDF F_X	$\mathrm{E}_{f_X}[X]$	$\mathrm{Var}_{f_X}[X]$	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	$\exp\{\lambda(e^t-1)\}$
$Geometric(\theta)$	$\{1,2,...\}$	$\theta \in (0,1)$	$(1-\theta)^{x-1}\theta$	$1-(1-\theta)^x$	$\dfrac{1}{\theta}$	$\dfrac{(1-\theta)}{\theta^2}$	$\dfrac{\theta e^t}{1-e^t(1-\theta)}$
$NegBinomial(r,p)$	$\{0,1,2,...\}$	$r \in \mathbb{Z}^+, p \in (0,1)$	$\binom{r+x-1}{x}p^r(1-p)^x$		$\dfrac{r(1-p)}{p}$	$\dfrac{r(1-p)}{p^2}$	$\left(\dfrac{p}{1-e^t(1-p)}\right)^r$

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

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

CONTINUOUS DISTRIBUTIONS

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