請勿攜去

Not to be taken away

香 港 中 文 大 學

The Chinese University of Hong Kong

二〇一六至一七年度下學期科目考試

Course Examination 2nd Term, 2016-17

科目編號及名稱 Course Code & Title : STAT3006 Statistical Computing

時間 Time allowed : 3 小時 hours 0 分鐘 minutes

學號 Student I.D. No. : ________ 座號 Seat No. : ________

Q1 (10%): (a) What is the purpose of statistics? (b) What are the two schools of inference methods for statistics? (c) How are the statistical computing methods covered in the course related to the two schools of inference methods?

Q2 (10%): Please write out a functional iteration algorithm to calculate the square root of a positive number a: $\sqrt{a}$.

Q3 (20%): Suppose we have independent and identically distributed (I.I.D.) samples, $X_1, X_2, \ldots, X_n$, that are from the *Cauchy* distribution with the density function

$$f(x|\theta) = \frac{1}{\pi}\frac{1}{1+(x-\theta)^2}.$$

Please derive a Newton's method to search the maximum likelihood estimator (MLE) of θ.

Q4 (20%): We collected n count data samples $X_1, X_2, \ldots, X_n$. Each sample is known to be from either $Pois(\lambda_1)$—a Poisson distribution with mean λ_1 or $Pois(\lambda_2)$—a Poisson distribution with mean λ_2. We associate each sample X_i with a group indicator C_i. If $C_i = 1$, $X_i \sim Pois(\lambda_1)$; if $C_i = 0$, $X_i \sim Pois(\lambda_2)$. The proportion of samples who follow $Pois(\lambda_1)$ is $P(C_i = 1) = p$, and accordingly $P(C_i = 0) = 1 - p$. However, we only observe X_i for sample i without knowing the true group indicator C_i. Based on the information, please write out the log complete-data likelihood function $l(p, \lambda_1, \lambda_2 | X_1, \ldots, X_n, C_1, \ldots, C_n)$ and derive an EM algorithm to estimate the unknown parameters $(p, \lambda_1, \lambda_2)$.

Q5 (10%): There are two known probability density functions (p.d.f.s) $f(x)$ and $g(x)$. Suppose that $f(x) \le 3 \cdot g(x)$ and sampling from $g(x)$ is much easier than directly sampling from $f(x)$, how can we use samples from $g(x)$ to obtain samples from $f(x)$? Please write out detailed steps. If $f(x)$ is a t distribution with degrees of freedom two and $g(x)$ is a normal distribution, does your proposed method work? Please explain why or why not.

Q6 (10%): I.I.D. samples $X_1, X_2, \ldots, X_n$ follow a normal distribution with an unknown mean μ and an unknown variance σ^2. We want to infer (μ, σ^2) in the Bayesian framework. Therefore, we have to assign a prior distribution to both μ and σ^2. The prior distribution is selected as the normal-inverse-chi-square distribution with parameters (m, k, r, s), written as $Normal\chi^{-2}(m, k, r, s)$. A random vector (μ, σ^2) is said to follow $Normal\chi^{-2}(m, k, r, s)$, if its joint p.d.f. $\pi(\mu, \sigma^2) \propto (\sigma^2)^{-\frac{r+3}{2}} exp\left(-\frac{k(\mu-m)^2+rs}{2\sigma^2}\right)$. What is the posterior distribution of (μ, σ^2)?

Q7 (20%): There are six random variables $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$ whose joint p.d.f. is:

$$f(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6) \propto e^{-\frac{3}{2}\theta_1^2 - \theta_1\theta_2 - \frac{3}{2}\theta_2^2} \cdot \frac{\theta_4^{\theta_3+1}}{\theta_3!} e^{-2\theta_4} \cdot n\theta_5 \theta_6^{\theta_5+1}(1-\theta_6)^{n+1-\theta_5}.$$

Based on the joint p.d.f., please (a) derive the full conditional functions $\{f_1, f_2, \ldots, f_6\}$ (b) write out an algorithm to draw samples from this joint p.d.f. (c) Is it possible to parallelize your proposed algorithm? If yes, please explain how; if not, please explain why.
(Note: the p.d.f. of $\Gamma(\alpha, \beta)$ is $\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$; the p.d.f. of $Pois(\lambda)$ is $e^{-\lambda}\frac{\lambda^x}{x!}$; the p.d.f of $Beta(a, b)$ is $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$.)

— END —