UNIVERSITY OF CAMBRIDGE

M. PHIL. IN STATISTICAL SCIENCE

Monday 11 June 2007 9.00 to 12.00

TIME SERIES AND MONTE CARLO INFERENCE

Attempt ***FOUR*** *questions.*

There are ***SIX*** *questions in total.*

The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury Tag
Script paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1 Time Series

Suppose that

$$X_t = \sum_{k=1}^{p} \phi_k X_{t-k} + \epsilon_t + \sum_{k=1}^{q} \theta_k \, \epsilon_{t-k}, \qquad (*)$$

where $\{\epsilon_t\}$ is a white noise process with variance σ^2. Write down conditions under which $(*)$ has a unique stationary causal solution $X_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j}$ which is an invertible ARMA(p,q) process.

Assuming these conditions are satisfied and quoting results from lectures as necessary, show that the c_j's satisfy the recursion

$$c_j = \theta_j + \sum_{k=1}^{p} \phi_k c_{j-k}, \quad j = 0, 1, \ldots$$

where we define $\theta_0 = 1$, $\theta_j = 0$ for $j > q$, and $c_j = 0$ for $j < 0$.

Find a similar recursion for the d_j's in $\epsilon_t = \sum_{j=0}^{\infty} d_j X_{t-j}$.

For the process

$$X_t = \phi X_{t-1} + \epsilon_t + \theta \epsilon_{t-1}, \qquad (**)$$

write down conditions on ϕ and θ under which $(**)$ has a unique stationary causal solution which is an invertible ARMA$(1,1)$ process. Find the c_j's and the d_j's.

2 Time Series

Let $X_t = \epsilon_t + \theta \epsilon_{t-1}$ where $\{\epsilon_t\}$ is a white noise process with variance σ^2 and θ is real. Show that the process $\{X_t\}$ is weakly stationary. Find its autocovariance function and its spectral density function $f(\lambda)$. Show that $f(-\lambda) = f(\lambda)$.

For a bivariate process $Z_t = (X_t, Y_t)^T$ with $\mathbb{E}(X_t) = \mathbb{E}(Y_t) = 0$ for all t, let $\operatorname{cov}(Z_t, Z_{t+h})$ be the matrix

$$\begin{pmatrix} \mathbb{E}(X_t X_{t+h}) & \mathbb{E}(X_t Y_{t+h}) \\ \mathbb{E}(X_{t+h} Y_t) & \mathbb{E}(Y_t Y_{t+h}) \end{pmatrix}.$$

Find $\operatorname{cov}(W_t, W_{t+h})$ where $W_t = (U_t, V_t)^T$, and $\{U_t\}$, $\{V_t\}$ are uncorrelated white noise processes with variances σ_u^2 and σ_v^2 respectively.

Suppose that $Z_t = (X_t, Y_t)^T$ where $X_t = U_t + \theta_{11} U_{t-1} + \theta_{12} V_{t-1}$, $Y_t = V_t + \theta_{21} U_{t-1} + \theta_{22} V_{t-1}$, and the θ_{ij}'s are real constants.

Show that $\mathbb{E} X_t = \mathbb{E} Y_t = 0$ and find $\operatorname{cov}(Z_t, Z_{t+h})$ for all integers h.

Find $f_{XY}(\lambda) = \dfrac{1}{2\pi} \displaystyle\sum_{h=-\infty}^{\infty} \mathbb{E}(X_t Y_{t+h}) e^{-ih\lambda}$ and show that $f_{XY}(-\lambda) = \overline{f_{XY}(\lambda)}$, where $\bar{z}$ denotes the complex conjugate of z.

3 Monte Carlo Inference

Let F be a distribution function. Define the quantile function $F^{-1} : (0,1] \to \mathbb{R}$. Prove that if $U \sim U(0,1]$ then $X = F^{-1}(U)$ has distribution function F.

Describe the rejection algorithm for simulating a random variable with density f. Prove that the output of the algorithm does indeed have density f.

Starting from a sequence (U_n) of independent $U(0,1]$ random variables, in each of the two cases below give an algorithm to simulate from the given density.

i) $f(x) = \frac{1}{\pi x^{1/2}(1-x)^{1/2}}, \quad x \in (0,1)$

ii) $f(x) = \frac{p+\pi(1-p)|x|+px^2}{\pi(1+x^2)^2}, \quad x \in \mathbb{R}$, where $p \in (0,1)$ is known.

4 Monte Carlo Inference

Let $X = (X_1, \ldots, X_n)$ be a random vector having independent components each with distribution function F. Suppose that $\mathbb{E}_F(X_1^4) < \infty$ and that we are interested in estimating $\theta = \mathbb{E}_F(X_1^2)$. Define an estimator $\hat{\theta}^n = \hat{\theta}^n(X)$ by $\hat{\theta}^n = n^{-1}\sum_{i=1}^n X_i^2$. What is meant by the jackknife estimator $\hat{v}_{JACK}$ of $v = \mathrm{Var}_F \hat{\theta}^n(X)$? Show that $\hat{v}_{JACK}$ is unbiased.

Now let $x = (x_1, \ldots, x_n)$ be a realisation of X. Define the nonparametric bootstrap estimator $\hat{v}_{BOOT}$ of v. Describe a Monte Carlo algorithm for approximating $\hat{v}_{BOOT}$.

What is meant by a bootstrap percentile confidence interval for θ? Give a Monte Carlo procedure for approximating this interval. Describe one alternative, analytic confidence interval for θ.

5 Monte Carlo Inference

(a) Briefly, and qualitatively, compare and contrast the Metropolis–Hastings (MH) algorithm and the Gibbs sampler. You should describe, in particular, when the methods are equivalent, when one might be preferred over the other, and outline the difference between the original Metropolis algorithm and the MH algorithm.

(b) Let $x_1, \ldots, x_N$ be independent Poisson distributed observations for which there is a suspicion of a *change–point* along the observation of the process for some random $m = 1, \ldots, N$. That is, given m we have that $x_i|\lambda \sim Pois(\lambda)$, $i = 1, \ldots, m$ and $x_i|\phi \sim Pois(\phi)$, $i = m+1, \ldots, N$. Our prior assumptions are that λ, ϕ and m are independent with $\lambda \sim \Gamma(\alpha, \beta)$, $\phi \sim \Gamma(\gamma, \delta)$, and that m has a discrete uniform distribution over $\{1, \ldots, N\}$. Treat α, β, γ, and δ as known constants. Recall that the Poisson distribution has probability mass function

$$Pois(x|\lambda) = \tfrac{e^{-\lambda}\lambda^x}{x!}, \qquad x \in \{0, 1, 2, \ldots\},\ \lambda > 0$$

and the gamma distribution has probability density function

$$\Gamma(x|\alpha, \beta) = \tfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \qquad x > 0,\ \alpha > 0,\ \beta > 0.$$

(i) Give the joint posterior distribution of the parameter vector $\theta \equiv (\lambda, \phi, m)$ for data $\mathbf{x} = (x_1, \ldots, x_N)$ up to a constant of proportionality.

(ii) Devise a Markov chain Monte Carlo (MCMC) scheme to sample from the posterior distribution derived in part (i).

(iii) What is the meaning of $\mathbb{P}(m = N|\mathbf{x})$? How would you estimate this quantity using the sample obtained under the scheme you devised in part (ii)?

6 Monte Carlo Inference

(a) Briefly explain the similarities and differences between the data augmentation method and the EM (Expectation–Maximisation) algorithm.

(b) Suppose we observe data $x_1, \ldots, x_N$ that are believed to come from a population of k distinct clusters. Moreover, $x_1, \ldots, x_N$ are assumed to be independent and identically distributed with a likelihood function that is a mixture of k normals with common variance:

$$L(x_i; \alpha, \mu, \sigma^2) = \sum_{j=1}^{k} \alpha_j f(x_i; \mu_j, \sigma^2), \qquad 0 \leqslant \alpha_j \leqslant \sum_{j=1}^{k} \alpha_j = 1,$$

where

$$f(x; \mu, \sigma^2) = \tfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\tfrac{(x-\mu)^2}{2\sigma^2}\right\}.$$

(i) For fixed $k \geqslant 1$, derive an EM algorithm for estimating the parameters in the mixture model

$$\theta_k \equiv (\alpha_1, \ldots, \alpha_k, \mu_1, \ldots, \mu_k, \sigma^2).$$

(ii) Now suppose the number of clusters, k, is unknown. How would you estimate k?

END OF PAPER