UNIVERSITY OF CAMBRIDGE

M. PHIL. IN STATISTICAL SCIENCE

Friday 4 June, 2004 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.

Note: *The following properties of the Gamma and Beta distributions may be used without proof:*

If $X \sim \Gamma(a, b)$ *then*

$$f_X(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, \quad x \geqslant 0$$

and $\mathbb{E}(X) = \frac{a}{b}$, *with* $Var(X) = \frac{a}{b^2}$.

If $X \sim \beta(a, b)$ *then*

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad 0 < x < 1.$$

and $\mathbb{E}(X) = \frac{a}{a+b}$, *with* $Var(X) = \frac{ab}{(a+b)^2(a+b+1)}$.

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

1 Time Series

a) Let X be a second-order stationary process. Define its autocorrelations, correlogram, and sample partial autocorrelations and explain how they can be used to diagnose AR and MA processes.

b) Define the spectral density $f_X(\omega)$ of X. Suppose $Y_t = \sum_{s=-\infty}^{+\infty} a_s X_{t-s}$, for a sequence of real numbers $\{a_s\}$ such that $\sum_{s=-\infty}^{+\infty} |a_s| < \infty$. Let $A(z) = \sum_{s=-\infty}^{+\infty} a_s z^s$, $|z| \leq 1$. Show that the process Y is second-order stationary and has the spectral density $f_Y(\omega) = |A(e^{i\omega})|^2 f_X(\omega)$.

c) Suppose the moving average $\frac{1}{6}[-1, 2, 4, 2, -1]$ is applied k times to the white noise series $\{Z_t\}$, where $\mathsf{E}Z_t = 0$, $\mathsf{E}Z_t^2 = \sigma^2$. Find the spectral density of the smoothed series, say $f_k(\omega)$. Show that if $\omega \neq \pi/3$ then $f_k(\omega)/f_k(\pi/3) \to 0$ as $k \to \infty$. Comment on the effect produced by repeated smoothing.

2 Time Series

a) Consider the state space model,

$$X_t = S_t + v_t, \qquad S_t = S_{t-1} + w_t,$$

where X_t and S_t are both scalars, X_t is observed, S_t is unobserved, and $\{v_t\}$, $\{w_t\}$ are independent Gaussian white noise sequences with variances V and W respectively. Show that X_t is an ARMA $(1,1)$ process.

b) Denote $\mathrm{F}_{t-1} = \sigma(X_1, \ldots, X_{t-1})$. Suppose we know that the conditional distribution of S_{t-1} given F_{t-1} is $\mathrm{N}(\widehat{S}_{t-1}, P_{t-1})$, i.e., normal with mean $\widehat{S}_{t-1}$ and variance P_{t-1}. Derive the Kalman filtering equations for $\widehat{S}_t$ and P_t.

c) Show that $P_t \equiv P$ (independently of t) if and only if $P^2 + PW = WV$, and deduce that in this case the Kalman filter for $\widehat{S}_t$ is equivalent to exponential smoothing.

3 Monte Carlo Inference

Describe how, given an infinite series of standard uniform variates $U_1, U_2, \ldots$, you could

(i) sample from a $Bin(n, p)$ distribution;

(ii) sample from an $exp(\lambda)$ distribution via inversion;

(iii) sample from a $\beta(a, b)$ distribution via rejection sampling for $a, b \geqslant 1$;

(iv) sample from a $N(0, \sigma^2)$ via the ratio of uniforms method.

4 Monte Carlo Inference

(i) Explain how the method of importance sampling may be used to estimate $\mu = \mathbb{E}_f(\theta(x))$ from a sample $x_1, \ldots, x_n \sim g(x)$, where $f(x)$ and $g(x)$ are normalised densities with common support and $\theta(x)$ denotes any general scalar function of x.

(ii) How would your description in (i) change if the normalisation constant for f and/or g were unknown?

(iii) Show that the variance of the importance sampling estimator

$$\hat{\mu}_g = \frac{1}{n} \sum_{i=1}^{n} \frac{f(x_i)}{g(x_i)} \theta(x_i)$$

is given by

$$\mathrm{Var}(\hat{\mu}_g) = \frac{1}{n} \int \frac{f^2(x)\theta^2(x)}{g(x)}\, dx - \frac{\mu^2}{n} \ .$$

(iv) Suppose that $f(x) = \dfrac{1}{\pi(1+x^2)}$ and $\mu = \mathbb{P}_f(X \geqslant k)$. What function $\theta(x)$ could be used to estimate μ via importance sampling?

(v) Take $g(x) \propto 1$, for $0 \leqslant x \leqslant k$. Show that

$$\frac{1}{2} - \frac{k}{n} \sum_{i=1}^{n} \frac{1}{\pi(1+x_i^2)}$$

is an importance sampling estimate for μ with finite variance, where $x_1, \ldots, x_n \sim g(x)$.

(vi) Describe the method of antithetic variables to improve Monte Carlo estimation. How could it be used to improve the estimator in (v)?

5 Monte Carlo Inference

Suppose we observe data x_{ijt}, $i = 1, \ldots, I$, $j = 1, 2$ and $t = 1, 2$ which we model by assuming that

$$x_{ijt} \sim Poisson(\lambda_{jt})$$

where

$$\log \lambda_{jt} = \mu + \alpha_j + \beta_t.$$

(i) What does it mean when we say that "we set $\alpha_1 = \beta_1 = 0$ for identifiability"?

Adopting the identifiability constraint in (i) and taking $\theta = \exp \mu$ and priors,

$$\theta \sim \Gamma(a, b), \quad \alpha_2 \sim N(0, \sigma_1^2), \quad \beta_2 \sim N(0, \sigma_2^2)\,,$$

answer each of the following questions.

(ii) What is the posterior conditional distribution for θ?

(iii) How would you update the parameter μ in an MCMC algorithm to sample from the posterior distribution $\pi(\mu, \alpha_2, \beta_2|\mathbf{x})$?

(iv) Why would you use a Metropolis Hastings update for α_2 and β_2? What might be a sensible proposal and why?

(v) Explain how you would introduce a Reversible Jump MCMC step to decide whether or not to include the constant term μ in the model.

(vi) How would you use Reversible Jump MCMC to determine the posterior probability that $\alpha_1 = \alpha_2 = 0$?

6 Monte Carlo Inference

Suppose 120 individuals are each assigned to 4 political parties (L, C, D & O) with probabilities $\left(\frac{1}{4} - \frac{\theta}{4}, \frac{1}{6} - \frac{\theta}{6}, \frac{\theta}{12}, \frac{7}{12} + \frac{\theta}{3},\right)$ respectively.

(i) Show that the MLE for θ is a solution to

$$\theta^2(-4x_1 - 4x_2 - 4x_3 - 4x_4) + \theta(-7x_1 - 7x_2 - 3x_3 + 4x_4) + 7x_3 = 0$$

where (x_1, x_2, x_3, x_4) are the numbers observed in categories (L, C, D & O) respectively.

(ii) Why might it be useful to divide the fourth cell into two with probabilities $\frac{7}{12}$ and $\frac{\theta}{3}$?

(iii) Describe the EM algorithm for maximising a likelihood in the presence of "missing data".

(iv) Given data $(x_1 = 20, x_2 = 10, x_3 = 5, x_4 = 85)$ derive an EM algorithm for estimating θ.