On admission to the examination room, you should acquaint yourself with the instructions below. You <u>must</u> listen carefully to all instructions given by the invigilators. You may read the question paper, but must <u>not</u> write anything until the invigilator informs you that you may start the examination.

You will be given five minutes at the end of the examination to complete the front of any answer books used.

DO NOT REMOVE THIS QUESTION PAPER FROM THE EXAM ROOM

April 2016 **MTMD03 2015/16 A001**

Answer Book
Any bilingual English language dictionary permitted
Any non-programmable calculator is permitted

UNIVERSITY OF READING

Monte-Carlo Methods and Particle Filters (MTMD03)

Two hours

Answer **ANY TWO** questions

The marks for the individual components of each question are given in [ ] brackets. The total mark for the paper is 100.

1. a) Show how to derive samples from the density *f(x)* defined on $[0,\infty)$

$$f(x)=\frac{1}{(1+x)^2}$$

using the transformation method, also called probability integral transform method, starting from samples from the uniform distribution U(0,1]. (Hint: use y=1+x to calculate the integral.)

[10 marks]

b) A very simple model of weather prediction, taking into account uncertainties in the system, is the Markov chain with the following transition matrix from one day to the next:

$$P=\begin{pmatrix} 0.8 & p \\ q & 0.4 \end{pmatrix}$$

in which P_{11} denotes the probability that a sunny day is followed by another sunny day, P_{22} denotes the probability that a rainy day is followed by a rainy day, $P_{12}=p$ is the probability that a sunny day is followed by a rainy day etc.

Find the values for *p* and *q* using the fact that *P* is a stochastic matrix.

[5 marks]

c) Assume it is sunny on day 0. What is the probability it is sunny at day 2?

[5 marks]

d) Show that the Markov chain is ergodic.

[10 marks]

e) Find the invariant distribution *p(x)*. What does this mean in terms of sunny and rainy days?

[10 marks]

f) Calculate $(P^T)^2$, with *P* defined by

$$P = \begin{pmatrix} 0.8 & 0.2 \\ q & a \end{pmatrix}$$

What should the value of *a* be to ensure that every n-day forecast with n>1 is the same as the one-day forecast?

[10 marks]

2. a) Explain the Metropolis-Hastings algorithm.

[10 marks]

b) Explain why it is important to use a good proposal density in Metropolis-Hastings. What is the ideal choice for the proposal density and why is that choice not practical?

[10 marks]

c) In Adaptive Metropolis-Hastings the proposal density is chosen adaptively using K past samples as follows:

$$q(z \mid x^{n-1}) = N(x^{n-1}, Q),$$

with covariance

$$Q = \frac{1}{K-1} \sum_{k=0}^{K} \left(x^k - \bar{x}\right)\left(x^k - \bar{x}\right)^T$$

and

$$\bar{x} = \frac{1}{K} \sum_{k=0}^{K} x^k.$$

Do you think this will be effective? Why?

[10 marks]

Would you include the burn-in samples? Why?

[10 marks]

d) Assume that the posterior pdf is shaped as in the figure below. Explain why Adaptive Metropolis-Hastings will not be efficient. Which MCMC method would you use and why?

[10 marks]

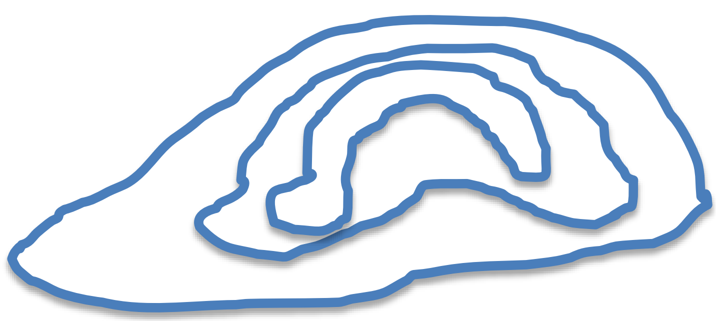

Figure 1 Contour plot of the posterior pdf.

3. Suppose we want to calculate the integral

$$I = \int_{-10}^{10} f(x)dx$$

in which $f(x) = e^{|x|}$

a) Sketch f(x) on the interval [-10,10]

[2 marks]

b) Show how this integral can be evaluated by drawing samples from the uniform density U[-10,10]. Will this be an efficient method to calculate the integral? Explain your answer.

[8 marks]

c) Argue that using the proposal density

$$q(x) = \frac{1}{\pi}\frac{1}{1+x^2}$$

to generate the samples will be even less efficient. Which proposal density would you use?

[10 marks]

d) Explain why standard Particle Filters are not efficient when the likelihood is much more peaked than the prior, and explain how a proposal density can be used to make Particle Filters more efficient.

[10 marks]

e) Give one argument why the number of independent observations is an important factor in the potential degeneracy of particle filters. Assume that the observation errors are Gaussian distributed.

[10 marks]

f) Assume that the number of independent observations is large, and we use a 4DVar on each particle sequentially in the particle filter. Write down the costfunction that should be minimized for each particle, and define all symbols used.

[10 marks]

[End of Question paper]