Candidates are admitted to the examination room ten minutes before the start of the examination. On admission to the examination room, you are permitted to acquaint yourself with the instructions below and to read the question paper.

Do not 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.

April 2015 **MTMD03/ 2014/15/ 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.

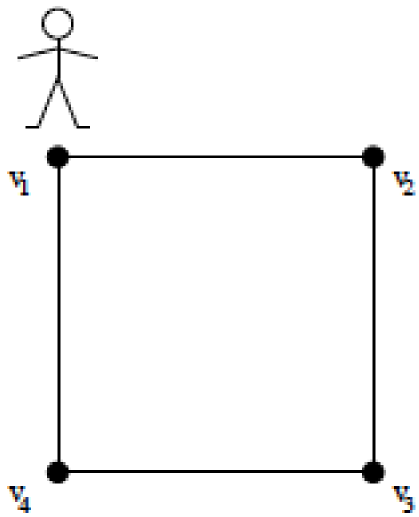
1. a) Show how to derive samples from the density *f(x)* defined on $[0, \infty)$

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

using the transformation method, also called probability integral transform method, starting from samples from the uniform distribution U[0,1]. (Hint: use the variable $y = e^{-x}$ to help with the integration.)

[10 marks]

A random walker in a very small town with 4 squares uses a fair coin to decide where to go next.



We can model this system via a Markov chain with the following transition matrix:

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

in which P_{11} denotes the probability that the walker remain where he/she *is* P_{12} is the probability that the walker starts from position 1 and moves to position 2, etc.

b) Show that P is a stochastic matrix.

[5 marks]

c) Assume that the probability density of the position of the walker at time zero is

$$p^{(0)} = (1, 0, 0, 0)^T$$

Show that

$$\Pr(X^{(1)} = 2 \mid X^{(0)} = 1) = \frac{1}{2}$$

in which $X^{(1)} = 2$ means that the walker is in position 2.
and that

$$\Pr(X^{(1)} = 3 \mid X^{(0)} = 1) = 0$$

[5 marks]

d) Calculate P^2. How do you interpret P^2 ?

[5 marks]

e) Prove by induction that if $p^{(0)}$ *is given as in c) that*

$$p^{(n)} = (0, \tfrac{1}{2}, 0, \tfrac{1}{2})^T$$

for $n=1,3,5,...$

$$p^{(n)} = (\tfrac{1}{2}, 0, \tfrac{1}{2}, 0)^T$$

for n=2,4,6,…

.

[10 marks]

f) What is the invariant distribution?

[5 marks]

g) Suppose we modify the random walk such that the random walker tosses two coins. The first coin is to decide whether to stay or to go. The second coin is used to decide on a clockwise or counterclockwise move.
Write down the transition matrix for this case.

[10 marks]

2. a) Explain why Gibbs sampling can be seen as a special kind of Metropolis-Hastings sampling.

[10 marks]

Assume we want to sample from the following density:

$$p(x) = xe^{-x} \quad \text{for} \quad x \geq 0$$

using Metropolis-Hastings with a Gaussian proposal density.

b) Show that the proposal density can be written as

$$q(x^* \mid x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x^* - x)^2}{2} \right)$$

in which x is the previous sample and x^* is the proposed sample.

[10 marks]

c) We want to investigate what the influence is of a proposal density that has different support from the target density. So, in this case the target is defined on the non-negative domain, while the proposal is defined on the whole real axis.

Show that straight-forward implementation of the above, i.e. ignoring that *x* is nonnegative in the target density, leads to an acceptance rate given by:

$$\alpha = \min\left\{ 1, \frac{p(x^*)}{p(x)} \right\}$$

[10 marks]

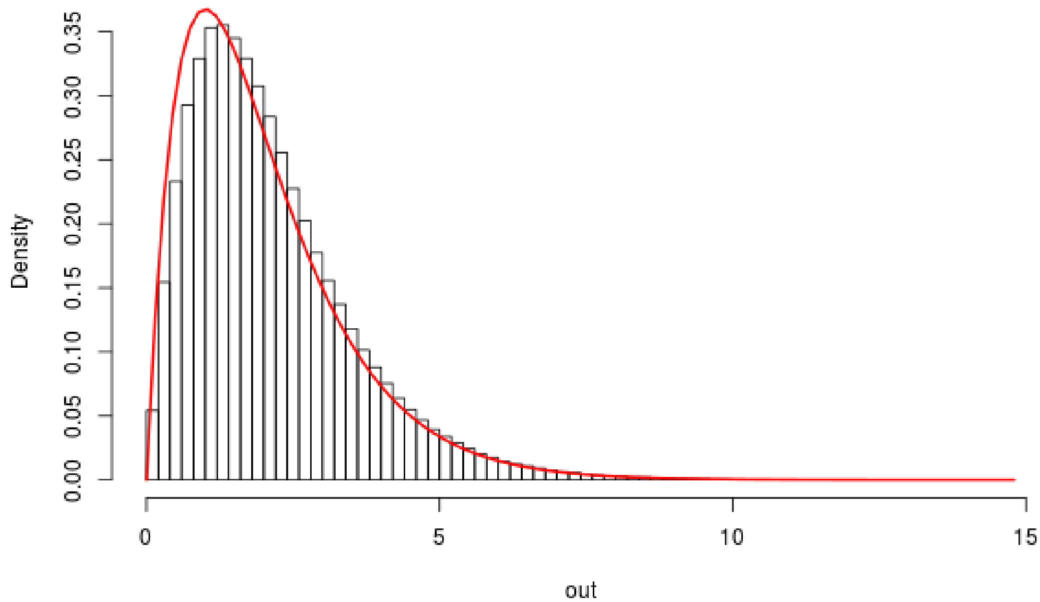
d) The straight-forward implementation used above is 'lucky' in the sense that all negative samples from the proposal lead to a negative acceptance rate that will be neglected anyway.

Explain that eliminating negative samples is a form of sampling too. How is that sampling method called?

[10 marks]

Suppose now that we want to strengthen the robustness of our method and ignore samples that are negative as soon as they appear, so we don't have to calculate their acceptance rate.

Doing this will lead to the following results:



In which the red line denotes the true density and the histogram denotes our estimate. Clearly something went wrong.

e) (i) Show that we have actually sampled from:

$$q(x^* \mid x) = \frac{1}{\Phi(x)} \exp\left(-\frac{(x^* - x)^2}{2} \right) \text{ for } x^* \geq 0$$

in which $\Phi(x)$ is the normalization constant, and give an expression for this normalization constant.

[5 marks]

(ii) How should we modify the acceptance ratio to generate samples from the correct density?

[5 marks]

3. a) Explain how the basic particle filter without resampling works, and why it is degenerate for almost all systems.

[5 marks]

b) Explain what resampling is, why it is useful, and give one example of a resampling scheme in the form of an algorithm.

[10 marks]

c) When the state dimension and the number of independent observations are high resampling will not be enough to avoid degeneracy. This is especially true in systems where the number of model time steps between observations is large. We will investigate the so-called Guided Importance Sampling Particle filter. In this scheme resampling is performed every time step, using the future observations. In this way the particles are guided towards these observations. The scheme works as follows. Consider a model state at time n and observations at time $n+m$. Let us assume, however, that the scheme pretends that the observations are made at every time step between the actual observations.

(i) Write down an expression for the weight a particle would obtain at time n with this assumption, assuming uncorrelated Gaussian observation errors.

[5 marks]

(ii) How will resampling help the particles to get closer to the observations at time $n+m.$?

[5 marks]

d) When calculating these weights the error covariance of the observations is typically multiplied by a factor larger than one. Why do you think that is?

[5 marks]

e) The weighting and resampling we just performed is artificial, and can be seen as a proposal density step. That will introduce a proposal density weight on each particle.

What is that proposal density weight?

[10 marks]

f) (i) Write down an expression for the final weights at the actual observation times.

[5 marks]

(ii) Why will this scheme be less degenerate than the standard particle filter with resampling?

[5 marks]

[End of Question paper]