UNIVERSITY OF OSLO

Faculty of Mathematics and Natural Sciences

Examination in	STK4050 — Statistical simulations and computation.
Day of examination:	Friday, December 4, 2009.
Examination hours:	14.30 – 17.30.
This problem set consists of 3 pages.	
Appendices:	None.
Permitted aids:	Approved calculator.

Please make sure that your copy of the problem set is complete before you attempt to answer anything.

Problem 1

Consider the standard *logistic distribution* which has the density

$$g(x) = \frac{e^{-x}}{[1+e^{-x}]^2}$$

and cumulative distribution function

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

(a) Explain how a variable can be generated from the logistic distribution using the inversion method.

Assume now we want to simulate from the standard normal distribution with density

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-0.5x^2}$$

(b) Define $h(x) = \log f(x) - \log g(x)$. Show that $h(x)$ has a maximum point at $x = 0$.

(c) Explain how to use the accept-reject method for simulation of the standard normal distribution using the logistic distribution as a proposal distribution.

What will be the acceptance rate of this method?

(Continued on page 2.)

Problem 2

Consider the bi-variate density

$$f(x_1, x_2) \propto \exp\left[-0.5x_1(1 + x_2^2)\right] \quad 0 < x_1, -\infty < x_2 < \infty$$

(a) Explain the main idea about Markov chain Monte Carlo algorithms and discuss their strengths and weaknesses.

(b) Propose a random walk Metropolis-Hastings algorithm for simulation from $f(x_1, x_2)$. Specify your choice of proposal distribution and write a pseudo-code for the simulations.

(c) Now construct a Gibbs sampler for simulation from $f(x_1, x_2)$.

Derive the necessary conditional distributions and write a pseudo-code for the simulations. (You can assume that routines are available for the conditional distributions involved).

(d) Discuss advantages and disadvantages of the Gibbs sampler compared to the Metropolis-Hastings algorithm.

Problem 3

Consider a dynamic state space model

$$\begin{aligned} x_t &\sim p(x_t|x_{t-1}) && \text{latent process} \\ y_t &\sim p(y_t|x_t) && \text{Observations} \end{aligned}$$

Define $\boldsymbol{y}_{1:t} = (y_1, ..., y_t)$. Our interest will be in sequential simulation from $p(x_t|\boldsymbol{y}_{1:t})$ as t increases.

(a) Explain why it can be easier to consider simulation from $p(\boldsymbol{x}_{1:t}|\boldsymbol{y}_{1:t})$.

Consider a sequential importance sampling method where $\boldsymbol{x}_{1:t}$ is generated through a sequential proposal distribution

$$q_t(\boldsymbol{x}_{1:t}|\boldsymbol{y}_{1:t}) = q_1(x_1|y_1)\prod_{s=2}^{t} q_s(x_s|x_{s-1}, y_s)$$

(Note that we allow the proposal distribution to depend on the observations.) The corresponding importance weights become

$$w_t(\boldsymbol{x}_{1:t}) = \frac{p(\boldsymbol{x}_{1:t}|\boldsymbol{y}_{1:t})}{q_t(\boldsymbol{x}_{1:t}|\boldsymbol{y}_{1:t})}$$

(b) Consider first $t = 1$. Assume $q_1(x_1|y_1) = p(x_1|y_1)$ (i.e. the conditional distribution for x_1 given y_1 based on the assumed model) and derive the importance weights $w_1(x_1)$ in this case. What is the advantage of this choice of proposal distribution compared to other choices?

(Continued on page 3.)

Assume now $q_t(x_t|x_{t-1}, y_t) = p(x_t|x_{t-1}, y_t)$ (the conditional distribution for x_t given both x_{t-1} and y_t based on the assumed model) for all t. One can show that for this specific choice of proposal distribution,

$$w_t(\boldsymbol{x}_{1:t}) = w_{t-1}(\boldsymbol{x}_{1:t-1})p(y_t|x_{t-1}). \qquad (*)$$

An alternative case is to use $q_t(x_t|x_{t-1}, y_t) = p(x_t|x_{t-1})$ which result in weights updated by

$$w_t(\boldsymbol{x}_{1:t}) = w_{t-1}(\boldsymbol{x}_{1:t-1})p(y_t|x_t). \qquad (**)$$

One can further show that the weights based on (*) will always have lower variance than the weights based on (**). None of these results you need to show.

(c) Given the better theoretical properties of the weights based on (*), discuss possible practical complications with this approach compared to the other one.

END