Imperial College London

MATH97085 MATH70047

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May-June 2022

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science

**Stochastic Simulation**

Date: 13 May 2022

Time: 09:00 – 11:00 (BST)

Time Allowed: 2:00 hours

Upload Time Allowed: 30 minutes

**This paper has 4 Questions.**

Candidates should start their solutions to each question on a new sheet of paper.

Each sheet of paper should have your CID, Question Number and Page Number on the top.

Only use 1 side of the paper.

Allow margins for marking.

Any required additional material(s) will be provided.

Credit will be given for all questions attempted.

Each question carries equal weight.

**SUBMIT YOUR ANSWERS AS SEPARATE PDFs TO THE RELEVANT DROPBOXES ON BLACKBOARD (ONE FOR EACH QUESTION) WITH COMPLETED COVERSHEETS WITH YOUR CID NUMBER, QUESTION NUMBERS ANSWERED AND PAGE NUMBERS PER QUESTION.**

1. (a) Consider the following congruential generator

$$X_n = (21X_{n-1} + b) \mod 16, \quad U_n = \frac{X_n}{m}.$$

(i) If $b \neq 0$, determine a constraint on $b$ such that this generator has full period.

(3 marks)

(ii) If $b = 5$, comment on the validity, and determine the period for the following choices for $X_0$: $X_0 = 1, X_0 = 7$ and $X_0 = 16$. (5 marks)

(iii) If $b = 0$, what can we say about the maximal period for this congruential generator? Justify your answer with a detailed explanation.

(2 marks)

(b) Let $q \in (0, 1)$. Let $X$ be a discrete random variable with probability mass function

$$f_X(x) = q(1-q)^{x-1}, x = 1, 2, \ldots$$

(i) Show that $F_X(x) = 1 - (1-q)^x$, $x = 1, 2, 3, \ldots$ and determine the generalised inverse of $F_X(x)$. (3 marks)

(ii) Provide the algorithm and the missing code ***** in the R code below to guarantee that the function `MyInverseGen` returns an i.i.d sample $X_1, \ldots, X_n$ of $X$. (2 marks)

```
MyInverseGen <- function(n,q)
{
U <- runif(n)
X <- *****
return(X)
}
```

**Question 1 continues on next page**

(c) Let $\theta \in (0,1)$. Consider the following pdf

$$f_X(x) = Kx^{\theta-1}(1-x)^{\theta-1}\mathbf{1}_{\{x\in(\frac{1}{2},1)\}},$$

where $K$ is an appropriate normalising constant.

(i) Describe in detail a rejection algorithm to sample from $f_X$ using one of the following two distributions to construct your envelope:

$$g_Y(y) \propto (1-y)^{\theta-1}, \qquad \text{or} \qquad h_Z(z) \propto z^{\theta-1}, \qquad \text{for} \quad y,z \in (1/2,1).$$

You can assume that you can sample directly from $g_Y$ and $h_Z$. (3 marks)

(ii) Assume that, for any $n \geq 1$, you can draw a sample of size $n$ from $Y$ and $Z$ using the hypothetical R functions `GenY(n)` and `GenZ(n)`, respectively. Provide the missing code in *** and ***** in the R code below to guarantee that the function `MyRejectionMethod` returns an i.i.d sample $X_1, \ldots, X_n$ of $X$. Define the appropriate value of the input parameter $M$. (2 marks)

```
MyRejectionMethod <- function(n,M,K,theta)
{
 x <- rep(0,n)
 i = 1
 while(i <= n){
    ***
    U <- runif(1)
    if(*****){
      x[i] <- y
      i <- i+1}
  }
return(x)
}
```

(Total: 20 marks)

2. (a) Determine whether a rejection algorithm can be implemented to sample from a Cauchy distribution using a standard normal distribution to construct the corresponding envelope. Justify your answer. Recall that the the probability density function (pdf) of a Cauchy random variable, $X$, is

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R},$$

and the pdf of a standard normal random variable, $Z$, is

$$g_Z(z) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}z^2}, \quad z \in \mathbb{R}$$

(2 marks)

(b) Let $U_1$ and $U_2$ be independent uniform random variates on $[0,1]$. Describe the estimators below and compare their expected values and variances. (4 marks)

(i) $\theta_1 = -\frac{1}{2}[\log(U_1) + \log(U_2)]$.

(ii) $\theta_2 = -\frac{1}{2}[\log(U_1) + \log(1-U_1)]$.

(c) Let $X$ be a random variable with pdf $f_X(\cdot)$. Suppose that the cumulative distribution $F_X(\cdot)$ and the quantile function $F_X^{-1}(\cdot)$ are known. Let $Y$ be the random variable $X$ truncated to the interval $I = (a,b)$ with $a < b$ and such that its pdf, denoted by $g_Y(\cdot)$, is proportional to $f_X(\cdot)\mathbf{1}_{(a,b)}(\cdot)$.

(i) Determine the normalising constant for $g_Y(\cdot)$. (1 mark)

(ii) Determine the cdf, $G_Y(\cdot)$, and the quantile function, $G_Y^{-1}(\cdot)$, of $Y$ in terms of $a, b, F_X(\cdot)$ and $F_X^{-1}(\cdot)$. (3 marks)

(iii) Using your calculations above, provide an algorithm to sample from $Y$. (2 marks)

(d) Consider the following integral

$$\theta := \int_0^\infty \int_{-\infty}^x \exp\{-(x + y^2/2)\} dy\, dx$$

Giving a detailed explanation, justify whether a Monte Carlo estimator can be used to approximate the integral above. If so, provide an algorithm to estimate $\theta$. (5 marks)

(e) Let $Z$ be an exponential random variable with parameter 1 and let $g : z \mapsto (z+2)^3$, $z \geq 0$. To estimate $\mathbf{E}[g(Z)]$, you are given the following estimators

$$\theta_1 = g\left(\frac{X+Y}{2}\right), \quad \text{and} \quad \theta_2 = \frac{g(X)+g(Y)}{2},$$

where $X$ and $Y$ are both exponential, antithetic variates with parameter 1. Based on their mean, explain which estimator you prefer. (3 marks)

(Total: 20 marks)

3. (a) Let $g_X(\cdot)$ be the pdf of $X$, a real-valued random variable and let $D$ be a bounded subset in $\mathbb{R}$. Giving all the details (including the necessary conditions for the method to work), provide an importance sampling estimator for the integral

(3 marks)

$$A = \int_D g_X(x)dx.$$

(b) Consider a random variable $X$ with pdf $f_X(x) \propto (1 + 2x^2)^{-3/4}$, $x \in \mathbb{R}^+$. Giving a detailed explanation, determine whether the method of ratio of uniforms can be used to sample from $X$ and, if so, outline the algorithm. (5 marks)

(c) Suppose that $Z = (X, Y)^T \sim MVN(\mu, \Sigma)$, where $\mu = (0, 0)^T$ and

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \text{with} \quad |\rho| < 1.$$

Recall that the corresponding pdf $f_Z(\cdot)$ is given by

$$f_Z(z) = \frac{1}{(2\pi)|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(z-\mu)^T\Sigma^{-1}(z-\mu)\right], \quad z \in \mathbb{R}^2,$$

where $|\Sigma|$ denotes the determinant of $\Sigma$.

(i) Describe, in detail, a Gibbs sampling method to obtain a sample $Z_i = (X_i, Y_i) \sim Z,\ i = 1, \ldots, n$. Write down the explicit form of the $n$th iteration update. You may assume that you can sample directly from one-dimensional distributions. (5 marks)

(ii) Giving a detailed explanation, use your algorithm above to provide an estimator for $\mathbb{P}[a \leq ||Z|| < b]$, for any $b > a > 0$. (4 marks)

(d) Consider the Metropolis-Hastings algorithm associated with a target pdf $f_X(\cdot)$ and a proposal distribution $q(y|x)$. Now consider a rejection algorithm to sample from $f_X(\cdot)$ using $g_Y(\cdot) := q(\cdot|x)$ to construct the corresponding envelope with $M = \sup_x f_X(x)/g_Y(x)$. Prove that, given a current state $x$, the Metropolis-Hastings algorithm accepts the proposed value $y$ more often than the rejection algorithm.

(3 marks)

(Total: 20 marks)

4. (a) (i) In terms of the required input information, give a detailed explanation of an advantage of the slice sampling algorithm over the Gibbs sampling method and the Metropolis-Hasting algorithm.

(3 marks)

(ii) Explain the mathematical principle behind the slice sampling method for univariate distributions. (2 marks)

(iii) Explain how a Markov chain can be constructed to sample from the bivariate uniform distribution involved in the slice sampling method to sample from a target density $f_X(x)$.

(3 marks)

(iv) In terms of implementation, explain the main drawback of the 2D slice sampling method.

(2 marks)

(b) Consider the following unnormalised density

$$f_X(x) = \exp\left(-\frac{1}{2}x^2\right) \mathbf{1}_{(\alpha,\infty)}(x),$$

with $\alpha > 0$.

(i) Giving a detailed explanation, provide an algorithm to obtain a sample of length $n$ from the distribution above using slice sampling.

(5 marks)

(ii) Let $\beta > \alpha$. Consider the integral

$$\theta := \int_{\alpha}^{\beta} \exp(-y^2/2)dy.$$

Giving a detailed explanation, justify whether a Monte Carlo estimator can be used to approximate the integral above. If so, provide the corresponding algorithm to estimate $\theta$ using the slice sampling method outlined above. (5 marks)

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2022

This paper is also taken for the relevant examination for the Associateship.

# MATH60047/MATH97085

# Stochastic Simulation (Solutions)

| Setter's signature | Checker's signature | Editor's signature |
|---|---|---|
| . . . . . . . . . . . . . . . . | . . . . . . . . . . . . . . . . | . . . . . . . . . . . . . . . . |

1. (a) (i) For full period, we have $21 = 1 \bmod 2$ (2 is the only prime factor of 16) and $21 = 1 \bmod 4$ (needed as 4 divides 16), so just need $b$ relatively prime to $m$ (i.e. gcd($b$,16) = 1), as $m$ is a power of 2 we just need $b$ to be odd. 

sim. seen ⇓

3, A

(ii) If b=5, then this is odd, so the generator has full period for all valid $X_0$, so full period of $X_0 = 1$ and 7, but $X_0 = 16$ is not valid (for $X_n \in \{0, 1, 2 \ldots, 15\}$).

5, A

(iii) If the shift parameter equals zero, then a congruential generator cannot reach full period as, in such a case, the generator is a multiplicative congruential generator

1, A

whose maximum possible period is given by $m-1$ (called the *maximal period*).

1, A

(b) (i) Note that the cumulative distribution function (cdf) is given as

meth seen ⇓

$$F_X(x) = \mathbb{P}[X \leq x] = \sum_{m=1}^{x} q(1-q)^{m-1} = 1 - (1-q)^x, \qquad x = 1, 2, \ldots.$$

We know that, by definition, the generalised inverse $F_X^- : [0,1] \mapsto \mathbb{R}$ of the cdf $F_X(\cdot)$ is

1, A

1, B

$$F^-(y) = \min\{x : F_X(x) \geq y\}.$$

Then, it follows that

1, B

$$x = F^-(y) = \left\lfloor \frac{\ln(1-y)}{\ln(1-q)} \right\rfloor + 1.$$

(ii) Using the calculations above and the inversion method for discrete random variables, it follows that the algorithm to generate from $X$ is

1. Generate an independent sample $U_i \sim U(0,1)$, $i = 1, \ldots n$.
2. Return $X_i = \left\lfloor \frac{\ln(1-U_i)}{\ln(1-q)} \right\rfloor + 1$, $i = 1, \ldots, n$.

1, A

Hence, *** should be replaced by

```
X <- floor(log(1-U)/log(1-q)) + 1
```

1, A

(c) (i) First, we will verify that the conditions of the rejection method hold. $\boxed{\text{meth seen} \Downarrow}$

- Note that, as required, both $g_Y$ and $h_Z$ have the same support than $f_X$.
- Second, we need to verify that there exists a finite $M > 0$ such that for all $x \in [1/2, 1)$, $f(x) \leq M g_Y(x)$. $\boxed{1, \text{B}}$
  As usual, we can take $M = \sup_x \frac{f_X(x)}{g_Y(x)}$. Indeed, $M_1 = \sup_{x \in (1/2,1)} K\, x^{\theta-1}$ and $M_2 = \sup_{x \in (1/2,1)} K(\theta)(1-x)^{\theta-1}$. Note that $M_1 = K\, \left(\frac{1}{2}\right)^{\theta-1}$ is finite, whereas $M_2 \to \infty$ as $x \to 1$, therefore we should choose the distribution $g_Y(\cdot)$ to construct our rejection algorithm. $\boxed{1, \text{B}}$

Therefore, the rejection algorithm to sample from $f_X$ is as follows:

1. Set $M_1 = K\, \left(\frac{1}{2}\right)^{\theta-1}$
2. Generate a random variate $Y = y \sim g$.
3. Generate a uniform random variate $U = u \sim (0, 1)$.
4. If $u \leq \frac{K\, y^{\theta-1}}{M_1}$, then set $X = y$. Otherwise, GOTO 2. $\boxed{1, \text{B}}$

**Remark:** The candidate may also provide a rejection algorithm using the unnormalised densities.

(ii) Considering the observations above, the missing code in *** is

`y <= GenY(1)` $\boxed{1, \text{A}}$

whereas the missing code in ***** is

`U<=K*(y**(theta-1))/M` where $K$ denotes the normalising constant for $f_X$, and $M$ corresponds to the constant $M_1$ defined above. $\boxed{1, \text{A}}$

2. (a) It is not possible to use the normal distribution to construct an envelope for a Cauchy distribution, since doing so would require to find a finite constant $M$ such that

meth seen ⇓

$$\frac{\sqrt{2\pi}}{\pi(1+x^2)} e^{x^2/2} \leq M.$$

1, A

However, since the exponential function grows faster than a polynomial function as $x \to \infty$, there is no least upper bound $M$ for any $x$.

1, A

(b) (i) $\theta_1$ is a crude Monte Carlo estimator (sample size $n = 2$) of $\mathbb{E}[X]$ where $X \sim Exp(1)$. It is the average of two independent exponential random variates with parameter 1.

meth seen ⇓

1, C

**Remark:** Similarly, the candidate may also identify $\theta_1$ and $\theta_2$ as estimators of $\mathbb{E}[-\log(U)]$, $U \sim (0, 1)$.

(ii) Since $-\log(\cdot)$ is a monotonic function, $\theta_2$ is an average of two estimators (of size $n = 1$) given by antithetic random variates.

1, C

Hence, as seen in lectures, both estimators have the same mean,

1, C

but $var(\theta_2) < var(\theta_1)$.

1, C

**Remark:** Similarly, the candidate may also identify $\theta_1$ and $\theta_2$ as estimators of $\mathbb{E}[-\log(U)]$, $U \sim (0, 1)$.

meth seen ⇓

(c) (i) By definition of pdf and cdf, the normalising constant $K$ is given by

$$K = \int_{-\infty}^{\infty} g_Y(y)dy = \int_a^b f_X(x)dx = F_X(b) - F_X(a).$$

1, A

(ii) By definition,

$$\begin{aligned} G_Y(y) &= \int_a^y g_Y(z)dz \\ &= \frac{F_X(y) - F_X(a)}{F_X(b) - F_X(a)}, \qquad y \in (a, b). \end{aligned}$$

1, B

**Remark:** Give the full mark only if the candidate states the range of $y$.

For the quantile function $Q_Y$ (which in this case coincides with the inverse of $G_Y(\cdot)$),

1, B

we use that $F_X^{-1}(\cdot)$ is known to obtain

$$Q_Y(u) = G_Y^{-1}(u) = F_X^{-1}\left(u[F_X(b) - F_X(a)] + F_X(a)\right), \quad 0 < u < 1.$$

1, B

**Remark:** Give the full mark only if the candidate states the values for $u$.

(iii) Since we obtained the inverse function $G_Y^{-1}$, we can use the inverse method as follows:

1. Generate a random variate $U = u \sim U(0,1)$.
2. Return $Y = F_X^{-1}\left(u[F_X(b) - F_X(a)] + F_X(a)\right)$

2, B

(d) To be able to approximate the integral above, we need to guarantee that $\theta$ is finite, but this is true because

$$\int_0^\infty \int_{-\infty}^\infty \exp\{-x - y^2/2\} dy\, dx = \sqrt{2\pi}.$$

Moreover, we can rewrite $\theta$ as an expectation of independent random variables. Indeed, since 1, B

$$\begin{aligned}\theta &= \int_0^\infty \int_{-\infty}^\infty \exp\{-x - y^2/2\} \mathbf{1}_{\{y<x\}} dy\, dx \\ &= \sqrt{2\pi}\, \mathbf{E}\left[\mathbf{1}_{\{Y<X\}},\right]\end{aligned}$$

where $X$ and $Y$ are independent random variables with $X \sim Exp(1)$ and $Y \sim N(0,1)$. 1, B

1, B

**Remark:** Give full mark only if the candidate states clearly that $X$ and $Y$ are independent.

Therefore, a Monte Carlo estimator $\hat{\theta}$ can be obtained as follows:

1. Generate $n$ i.i.d. random variables $X_i \sim \exp(1)$ (e.g. using inversion method).
2. Generate $n$ i.i.d. random variables $Y_i \sim N(0,1)$ (e.g. using Box-Muller method).
3. Return
$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i<Y_i\}}.$$

2, B

**Remark:** Give full mark only if the candidate has stated clearly the assumptions of independence among all variables involved.

unseen ⇓

(e) Observe that $\mathbf{E}[\theta_2] = \mathbf{E}[g(X)]$ since $X$ and $Y$ have the same distribution. 1, C

Since $g$ is not linear, it follows that $\mathbf{E}[\theta_1] \neq \mathbf{E}[\theta_2]$ 1, C

and, thus, both estimators cannot be compared. 1, C

3. (a) An importance sampling (IS) method for the integral above assumes that

meth seen ⇓

* we are unable to sample directly from $g_X$,
* we can sample from an auxiliary density $h_Y$ such that $g_X(x) > 0$ implies $h_Y(x) > 0$, and
* we can pointwise evaluate both $g_X(\cdot)$ and $h_Y(\cdot)$.

2, A

**Remark:** Give only 2 marks if all conditions above are stated. The IS estimator takes the form:

$$\tilde{A} = \frac{1}{n}\sum_{i=1}^{n} \frac{g_X(Y_i)\mathbf{1}_{\{Y_i \in D\}}}{h_Y(Y_i)}, \qquad Y_1, \ldots, Y_n \text{ i.i.d sample from } h_Y.$$

1, A

(b) Define $h : x \mapsto (1 + 2x^2)^{-3/4}$, $x \in \mathbb{R}^+$. For the implementation of the ratio of uniforms method, we need to enclose the area

sim. seen ⇓

1, A

$$C_h = \{(u, v) : 0 \le u \le \sqrt{h\left(\frac{u}{v}\right)}\}$$

by a rectangle $B = [0, \sup_x \sqrt{h(x)}] \times [\inf_{x \le 0} x\sqrt{h(x)}, \sup_{x \ge 0} x\sqrt{h(x)}]$,

2, A

which (as seen in lectures) can be obtained whenever $h(x)$ and $x^2 h(x)$ are bounded in the domain of $x$ .

1, C

However, in this case, $x^2 h(x)$ is unbounded and so the method is not appropriate to sample from $h$ .

1, C

(c) (i) Let $Z = (X, Y)^T$.

meth seen ⇓

First, compute the conditional density of $X$ given $Y = y$:

2, D

$$\begin{aligned} f_{X|Y=y}(x|y) &= \frac{f_{X,Y}(x, y)}{f_Y(y)} \\ &= \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left\{-\frac{(x-\rho y)^2}{2(1-\rho^2)}\right\}. \end{aligned}$$

Hence, $X|Y = y \sim N(\rho y, 1-\rho^2)$ and, similarly, $Y|X = x \sim N(\rho x, 1-\rho^2)$.

1, D

The algorithm to obtain a sample $Z_1, \ldots, Z_n$, is given as follows:

1. Initialise $X_0$ and set $m = 1$.
2. Given $Z_{m-1} := (X_{m-1}, Y_{m-1})$ $(= (x_{m-1}, y_{m-1}) =: z_{m-1})$, for $m \ge 1$, sample a new value from the full conditionals
   $X_m \sim N(\rho y_{m-1}, 1-\rho^2)$. 1, D
   $Y_m \sim N(\rho x_m, 1-\rho^2)$. 1, D
3. Replace $m$ by $m + 1$ and GOTO step 2 if $m \le n$. Otherwise, return $Z_1, \ldots, Z_n$.

**Remark:** Give 2 marks only if the candidate did not identify the conditionals as normal distributions.

(ii) Since $\boxed{1, \mathrm{D}}$

$$\theta := \mathbb{P}[a \leq ||Z|| < b] = \mathbb{E}\left[\mathbf{1}_{||Z|| \in [a,b)}\right],$$

and the sample generated by the estimator above is not independent, $\boxed{1, \mathrm{D}}$
a Monte Carlo estimator can be obtained as follows:

1. Sample from $Z$ using the algorithm above allowing a burn-in period $B$.
2. Discard the sample during the burn-in period.
3. Use the rest of the sample to compute the

$$\tilde{\theta} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{||Z_{i+B}|| \in [a,b)\}}.$$

$\boxed{2, \mathrm{D}}$

(d) Compare the acceptance probabilities for both methods: $\boxed{\text{meth seen} \Downarrow}$

∗ For $(x, y)$, the acceptance probability of the proposed value $y$, given the current state $x$, for the MH algorithm is

$$\alpha(x, y) = \min\left\{\frac{f_X(y) g_Y(x)}{f_X(x) g_Y(y)}, 1\right\}.$$

$\boxed{1, \mathrm{D}}$

∗ The corresponding acceptance probability for the rejection algorithm is

$$\gamma(x, y) = \frac{f_X(y)/g_Y(y)}{\sup f_X(z)/g_Y(z)}$$

$\boxed{1, \mathrm{D}}$

Since $\alpha(x, y) \geq \gamma(x, y)$ for all $(x, y)$, it follows that MH accepts more often than the rejection algorithm. $\boxed{1, \mathrm{D}}$

4. (a) (i) Slice sampling only need the target distribution, whereas M-H and Gibbs sampling require to derive (or choose) additional distributions to implement either method. unseen ⇓ 1, M

Indeed, M-H algorithm requires a proposal transition distribution to update the states of the Markov chain constructed to sample from the target distribution. Slice sampling only need the target distribution, whereas M-H and Gibbs sampling require to derive (or choose) additional distributions to implement either method. 1, M

Gibbs sampling requires the computation of full conditionals of the given joint distribution. Slice sampling only need the target distribution, whereas M-H and Gibbs sampling require to derive (or choose) additional distributions to implement either method. 1, M

(ii) The slice sampling method relies on the principle that one can sample from a target density $f_X(\cdot)$ by sampling uniformly from the subgraph of the target density function, i.e. $U_f := \{(x, y) : x \in \text{ supp } f, 0 < y < f_X(x)\}$, and then look only at the horizontal coordinates of the sample points. Slice sampling only need the target distribution, whereas M-H and Gibbs sampling require to derive (or choose) additional distributions to implement either method. 2, M

(iii) A Markov chain that converges to the uniform distribution of the region $U_f := \{(x, y) : 0 < y < f_X(x)\}$ 1, M

can be constructed as follows: starting from a point $(x, y)$ in $U_f$, one can

a) sample alternately from the conditional distribution for $Y|X = x$, and then 1, M

b) sampling the new $x$ from the conditional distribution for $X|Y = y$. 1, M

**Remark:** Alternatively, the candidate may provide the explicit conditional distributions.

(iv) The sampling of the uniform distribution over the slice $S_y := \{x : f_X(x) > y\}$ may be difficult to implement 1, M

when the exact form of $S_y$ is unknown. 1, M

(b) (i) Let $U_f := \left\{(x, y) : 0 < y < \exp\left(-\frac{1}{2}x^2\right) \mathbf{1}_{(\alpha,\infty)}(x)\right\}$.

Given (x,y), Step 1 of the slice sampler requires to sample $y'$ from a uniform distribution $U(0, f(x))$, and Step 2 requires to sample a new $x'$ from a uniform distribution over the region $S = \{x : y' < f_X(x)\} = (\alpha, \sqrt{-2\ln y'})$. $\boxed{2, \mathrm{M}}$

1. Given an initial value $(x_0, y_0) \in U_f$ with $x_0 > \alpha$, set $i = 1$. $\boxed{1, \mathrm{M}}$
2. Sample $Y = y_i | X = x_{i-1}$ from $U(0, f_X(x_{i-1}))$ $\boxed{1, \mathrm{M}}$
3. Sample $X = x_i$ from $U(\alpha, \sqrt{-2\log y_i})$
4. If $i = n$, return $x_1, \ldots, x_n$. Otherwise, set $n = i + 1$ and GOTO 2. $\boxed{1, \mathrm{M}}$

(ii) To be able to approximate the integral above, we need to guarantee that $\theta$ is finite, but this is true because

$$\theta \leq \int_{-\infty}^{\infty} \exp(-y^2/2)dy = \sqrt{2\pi}.$$

$\boxed{1, \mathrm{M}}$

Moreover, we can rewrite $\theta$ as

$$\theta = \mathbf{E}\left[\mathbf{1}_{\{X<\beta\}},\right]$$

where $X$ is a r.v. with the unnormalised density $f_X(x)$ given in (b). $\boxed{1, \mathrm{M}}$

Therefore, a Monte Carlo estimator $\hat{\theta}$ can be obtained as follows:

1. Sample from $X$ using the slice sampling algorithm in (i) allowing a burn-in period $B$. $\boxed{1, \mathrm{M}}$
2. Discard the sample during the burn-in period.
3. Use the rest of the sample to compute and return

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_{i+B}<\beta\}}.$$

$\boxed{2, \mathrm{M}}$

**Review of mark distribution:**

Total A marks: 24 of 24 marks

Total B marks: 15 of 15 marks

Total C marks: 9 of 9 marks

Total D marks: 12 of 12 marks

Total marks: 80 of 60 marks

Total Mastery marks: 20 of 20 marks

**If your module is taught across multiple year levels, you might have received this form for each level of the module. You are only required to fill this out once for each question.**

**Please record below, some brief but non-trivial comments for students about how well (or otherwise) the questions were answered. For example, you may wish to comment on common errors and misconceptions, or areas where students have done well. These comments should note any errors in and corrections to the paper. These comments will be made available to students via the MathsCentral Blackboard site and should not contain any information which identifies individual candidates. Any comments which should be kept confidential should be included as confidential comments for the Exam Board and Externals. If you would like to add formulas, please include a separate pdf file with your email.**

| ExamModuleCode | QuestionNumber | Comments for Students |
|---|---|---|
| Stochastic Simulation_MATH60047 MATH97085 MATH70047 | 1 | |
| Stochastic Simulation_MATH60047 MATH97085 MATH70047 | 2 | |
| Stochastic Simulation_MATH60047 MATH97085 MATH70047 | 3 | This question was well done, but more attention could have been paid to careful description of the rationale and conditions required by the methods in (a) and (b). |
| Stochastic Simulation_MATH60047 MATH97085 MATH70047 | 4 | |
| Stochastic Simulation_MATH60047 MATH97085 MATH70047 | 5 | |