Imperial College London

MATH97016 MATH97094

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May-June 2020

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

Computational Stochastic Processes

Date: 15th May 2020

Time: 13.00pm - 15.30pm (BST)

Time Allowed: 2 Hours 30 Minutes

Upload Time Allowed: 30 Minutes

This paper has 5 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 ONE PDF TO THE RELEVANT DROPBOX ON BLACKBOARD INCLUDING A COMPLETED COVERSHEET WITH YOUR CID NUMBER, QUESTION NUMBERS ANSWERED AND PAGE NUMBERS PER QUESTION.

.

1. (a) For $L > 0$, consider the truncated exponential distribution with density

$$f_L(x) = \frac{e^{-x}}{1 - e^{-L}} \, \mathbb{1}_{[0,L]}(x), \quad x \in \mathbb{R}. \tag{1}$$

(i) Using the inverse transform method, derive a scheme for generating samples with distribution f_L from uniformly distributed samples, $\mathcal{U}(0,1)$. (5 marks)

(ii) Let $u_L(x) = \frac{1}{L} \, \mathbb{1}_{[0,L]}(x)$ denote the density of the uniform distribution on $[0, L]$. Find the minimum value of M such that

$$f_L(x) \leq M u_L(x) \qquad \forall x \in \mathbb{R},$$

and describe the rejection sampling algorithm for generating samples from the distribution f_L using the proposal density u_L. Write down without proof the acceptance probability as a function of L and comment briefly. (5 marks)

(b) Suppose that we wish to estimate $I = \mathbb{E}[f(X)]$, where the random variable X has probability density function $p(x)$ on $\mathbb{R}$.

(i) Let $\hat{I}_n$ be the Monte Carlo estimator for I defined by

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^{n} f(X_i),$$

where $X_i \sim p(x)$ are independent, identically distributed random variables. Calculate the bias and the variance of the estimator $\hat{I}_n$ in terms of n and $\sigma_f^2 := \mathrm{Var}[f(X_1)]$.

(2 marks)

(ii) Let $\alpha \in (0,1)$. Denoting by $\Phi(x)$ the cumulative distribution function of the standard normal distribution and by Φ^{-1} its inverse, and assuming that $n \gg 1$, use the central limit theorem to find $b = b(n, \sigma_f^2, \alpha)$ such that

$$\mathbb{P}\left[I \in [\hat{I}_n - b, \hat{I}_n + b]\right] \approx 1 - \alpha,$$

(4 marks)

(iii) We define the estimator

$$\hat{I}_n^{cv} = \frac{1}{n} \sum_{i=1}^{n} \Big(f(X_i) + \beta \, g(X_i) \Big),$$

where g is a function such that $\mathbb{E}[g(X_1)] = 0$ and $\sigma_g^2 := \mathrm{Var}[g(X_1)] > 0$. Derive an expression, as a function of σ_f^2, σ_g^2 and $\mathrm{Cov}[f(X_1), g(X_1)]$, for the value of $\beta \in \mathbb{R}$ for which the variance of $\hat{I}_n^{cv}$ is minimized. (4 marks)

(Total: 20 marks)

2. Let W_t be a standard one-dimensional Brownian motion. We define the Brownian bridge on $[0,1]$ to be the process

$$B_t = W_t - tW_1, \qquad t \in [0,1].$$

(a) Show that $\mathbb{E}[B_t] = 0$ and that $\mathrm{Cov}(B_s, B_t) = \min(s,t) - st$ for any $s, t \in [0,1]$. (4 marks)

(b) Let $\mathcal{K} : L^2(T) \to L^2(T)$, with $T := [0,1]$, be the operator defined by

$$(\mathcal{K}f)(t) = \int_0^1 C(s,t)\, f(s)\, \mathrm{d}s,$$

where $C(s,t) := \mathrm{Cov}(B_s, B_t)$. Show that any eigenpair (φ_n, λ_n) of $\mathcal{K}$, where φ_n is an eigenfunction *normalized in* $L^2(T)$ and λ_n is the associated eigenvalue, satisfies

$$\lambda_n\, \varphi_n''(s) = -\varphi_n(s), \tag{2}$$

with boundary conditions $\varphi_n(0) = \varphi_n(1) = 0$. (4 marks)

(c) Calculate the solutions (φ_n, λ_n) to (2). Arrange the eigenvalues in decreasing order and normalize the eigenfunctions, i.e. ensure that $\int_0^1 |\varphi_n(s)|^2\, \mathrm{d}s = 1$. (4 marks)

(d) Let us define, for $n \in \mathbb{N}$,

$$Z_n = \int_0^1 B_t\, \varphi_n(t)\ \mathrm{d}t.$$

Show that $\mathbb{E}[Z_n] = 0$ and $\mathbb{E}[Z_m\, Z_n] = \lambda_n\, \delta_{mn}$ for all $m, n \in \mathbb{N}$. (4 marks)

(e) Consider the truncated Karhunen–Loève expansion

$$S_N = \sum_{n=1}^{N} Z_n\, \varphi_n.$$

Show that, for $M \geq N$,

$$\int_0^1 \mathbb{E}\left[|S_N(t) - S_M(t)|^2\right]\ \mathrm{d}t = \sum_{n=N+1}^{M} \lambda_n.$$

(4 marks)

(Total: 20 marks)

3. (a) Consider the scalar Itô SDE without drift

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0, \tag{3}$$

where W_t is a standard one-dimensional Brownian motion and $\sigma(x)$ is a smooth function.

(i) Show that, for $n = 0, 1, \dots, N-1$,

$$\int_{n\Delta t}^{(n+1)\Delta t} \int_{n\Delta t}^{s} \mathrm{d}W_u\,\mathrm{d}W_s = \frac{1}{2}\Big(\Delta W_n^2 - \Delta t\Big),$$

where $\Delta W_n = W_{(n+1)\Delta t} - W_{n\Delta t}$ and Δt is a time step. (5 marks)

(ii) Use Part (i) and Itô's formula to derive Milstein's scheme for (3):

$$X_{n+1} = X_n + \sigma(X_n)\sqrt{\Delta t}\,\xi_n + \frac{1}{2}\,\sigma(X_n)\,\sigma'(X_n)\,\Delta t\Big(\xi_n^2 - 1\Big), \tag{4}$$

where $\xi_n \sim \mathcal{N}(0,1)$ are independent. (7 marks)

(b) Consider the scalar Itô SDE

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \tag{5}$$

where W_t is a standard one-dimensional Brownian motion and $b(x)$ and $\sigma(x)$ are smooth functions. The Euler θ method for (5) with N time steps over the interval $[0,T]$ reads

$$X_{n+1}^N = X_n^N + \Big(\theta\, b(X_{n+1}^N) + (1-\theta)\, b(X_n^N)\Big)\,\Delta t + \sigma(X_n^N)\,\xi_n\sqrt{\Delta t}, \qquad \Delta t = T/N. \tag{6}$$

where $\xi_n \sim \mathcal{N}(0,1)$ are independent, $\theta \in [0,1]$ is a parameter encoding the degree of implicitness of the scheme, and $n = 0, \dots, N-1$.

(i) Show that the θ Euler scheme, when applied to the Itô SDE

$$\mathrm{d}X_t = X_t\,\mathrm{d}t + X_t\,\mathrm{d}W_t, \qquad X_0 = 1, \tag{7}$$

can be written in the form

$$X_{n+1}^N = G(\Delta t, \xi_n, \theta)\,X_n^N.$$

(4 marks)

(ii) Show that, for fixed T,

$$\Big|\mathbb{E}[X_N^N - X_T]\Big| \to 0 \quad \text{as } N \to \infty,$$

where X_N^N denotes the solution of the numerical scheme (6) at the final time. You may use the formula

$$e^x = \lim_{N\to\infty}\left(1 + \frac{x}{N}\right)^N,$$

and the fact that the exact solution to (7) is such that, for any $t \geq 0$, $\mathbb{E}[X_t] = e^t$. (4 marks)

(Total: 20 marks)

4. (a) Consider the Itô SDE

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad X_0 = x_0, \qquad t \in [0,T], \tag{8}$$

where W_t is a standard Brownian motion, x_0 is a deterministic initial condition, $b(x)$ is a smooth bounded function and T is a fixed final time. Let $\Delta t = T/N$ and $t_i = i\,\Delta t$ for $i = 0, \dots, N$, where N is a number of intervals and T is a fixed final time. We define

$$\widehat{\sigma}_N^2 = \frac{1}{T}\sum_{i=0}^{N-1}(X_{t_{i+1}} - X_{t_i})^2$$

and, for $i = 0, \dots, N-1$,

$$I_i = \int_{i\Delta t}^{(i+1)\Delta t} b(X_s)\,\mathrm{d}s, \qquad M_i = \sigma\Delta W_i = \sigma(W_{t_{i+1}} - W_{t_i}).$$

(i) Use the Cauchy-Schwarz inequality to prove that

$$\mathbb{E}(I_i^2) \le C(\Delta t)^2.$$

for a constant C independent of Δt. (3 marks)

(ii) Use Young's inequality with an appropriate choice of ϵ to show that

$$\mathbb{E}(|I_i\,M_i|) \le C(\Delta t)^{3/2},$$

for a constant C independent of Δt. (3 marks)

(iii) Using the fact that you can write $X_{t_{i+1}} = X_{t_i} + I_i + M_i$, prove that

$$|\mathbb{E}(\widehat{\sigma}_N^2) - \sigma^2| \le C(\Delta t + \sqrt{\Delta t}),$$

for a constant C independent of Δt. (4 marks)

(b) Let X_t denote the solution to

$$\mathrm{d}X_t = -\theta\,b(X_t)\,\mathrm{d}t + \mathrm{d}W_t, \qquad X_0 = x_0 := 0, \tag{9}$$

where θ is a parameter and $b(x)$ is a smooth function, and consider the Euler–Maruyama discretization $\hat{X}_n$ of (9) with time step Δt, which is defined by the iteration

$$\hat{X}_{n+1} = \hat{X}_n - \theta\,b(\hat{X}_n)\,\Delta t + \sqrt{\Delta t}\,\xi_n, \qquad n = 0, \dots, N-1, \qquad \hat{X}_0 = x_0. \tag{10}$$

Here $\xi_n \sim \mathcal{N}(0,1)$ are independent.

(i) Write down the probability distribution function $f(x_1, \dots, x_N; \theta)$ of $\hat{X} := \{\hat{X}_n\}_{n=1}^N$, viewed as a random variable with values in $\mathbb{R}^N$. (2 marks)

(ii) Obtain an expression for the maximum likelihood estimator for θ given an observation $\hat{X}$ of the *discrete-time solution* defined by (10). (4 marks)

(iii) In the particular case where $b(x) = 1$, calculate the mean-square error of the maximum likelihood estimator obtained in Part (ii). (4 marks)

(Total: 20 marks)

5. Let $\pi(x)$ be a probability distribution on $\mathbb{Z}$, known up to the normalization constant, such that

$$\pi(x) = \pi(-x) \quad \forall x \in \mathbb{Z}.$$

Let also $q(\cdot \,|\, x)$ be a proposal density such that $q(y \,|\, x) = g(y - x)$, for some function g satisfying

$$\sum_{x \in \mathbb{Z}} g(x) = 1 \qquad \text{and} \qquad g(x) = g(-x) \quad \forall x \in \mathbb{Z}.$$

Given an initial value $X_0 \in \mathbb{Z}$, consider the Markov chain generated by the following algorithm:

- Given the current state X_n, generate Y according to the proposal $q(\cdot \,|\, X_n)$.
- Let $u \sim U(0, 1)$.
- If $u \leq \alpha(X_n, Y)$, then set $X_{n+1} = Y$. Otherwise, set $X_{n+1} = -X_n$.

The acceptance probability $\alpha(x, y)$ is given by

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

For Parts (b) and (d) below, you may find it useful to consider separately the case $y \neq -x$ and the case $y = -x$.

(a) Given $Z \sim g(\cdot)$, how can we generate $Y \sim q(\cdot \,|\, x)$? (2 marks)

(b) For $x, y \in \mathbb{Z}$, write down the transition probability $p(x, y) = \mathbb{P}[X_{n+1} = y \,|\, X_n = x]$ for the Markov chain $\{X_n\}_{n \in \mathbb{N}}$. (7 marks)

(c) Define what it means for a distribution ψ to be a stationary distribution of the Markov chain $\{X_n\}_{n \in \mathbb{N}}$. (Do not use the explicit expression of $p(x, y)$ found in Part (b).) (3 marks)

(d) Show that π is a reversible measure for the Markov chain $\{X_n\}_{n \in \mathbb{N}}$, i.e. show that π satisfies the *detailed balance condition*:

$$\pi(x)\, p(x, y) = \pi(y)\, p(y, x) \qquad \text{for all } x, y \in \mathbb{Z}.$$

(6 marks)

(e) Given that π is a reversible measure for the Markov chain $\{X_n\}_{n \in \mathbb{N}}$, show that π is a stationary distribution of $\{X_n\}_{n \in \mathbb{N}}$. (2 marks)

(Total: 20 marks)

1. (a) (i) The CDF associated with f_L is

$$F_L = \begin{cases} 0 & x < 0, \\ \frac{1-e^{-x}}{1-e^{-L}} & x \in [0, L], \\ 1 & x > L. \end{cases}$$

The generalized inverse of F_L is

$$G_L(u) = -\log\left(1 - (1 - e^{-L})\, u\right), \qquad u \in [0, 1].$$

If $U \sim \mathcal{U}(0, 1)$, then $X = G_L(U) \sim f_L$, as required.

[SIMILAR]

(ii) Since f_L is a decreasing function, the minimum value of M can be obtained from the inequality at $x = 0$, i.e.

$$M = \sup_{x \in [0,L]} \frac{f_L(x)}{u_L(x)} = \frac{f_L(0)}{u_L(0)} = L(1 - e^{-L}).$$

Rejection sampling works as follows:

1. Generate a sample x from density $u_L(\cdot)$ and let $u \sim U(0, 1)$.
2. If $u < \frac{f_L(x)}{M\, u_L(x)}$, then accept the sample $X = x$ and stop.
3. Otherwise, reject the sample and return to step 1.

The acceptance probability is $\frac{1}{M}$, which is small when L is large.

[SIMILAR]

(b) (i) By linearity of the expectation, the bias is zero:

$$\mathbb{E}[\hat{I}_n] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[f(X_i)] = \frac{1}{n} \sum_{i=1}^{n} I = I.$$

Since the summands are independent and identically distributed, the variance is

$$\begin{aligned} \mathrm{Var}[\hat{I}_n] &= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}[f(X_i), f(X_j)], \\ &= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{ij}\, \mathrm{Var}[f(X_i)] = \frac{1}{n} \mathrm{Var}[f(X_1)] = \frac{\sigma_f^2}{n}. \end{aligned}$$

[SEEN]

(ii) By the central limit theorem,

$$\frac{\hat{I}_n - I}{\sigma_f / \sqrt{n}} \to \mathcal{N}(0, 1)$$

in distribution as $n \to \infty$. Therefore, for n sufficiently large, the left-hand side is approximately distributed according to $\mathcal{N}(0, 1)$, so

$$\mathbb{P}\left[\Phi^{-1}(\alpha/2) \leq \frac{\hat{I}_n - I}{\sigma_f / \sqrt{n}} \leq \Phi^{-1}(1 - \alpha/2)\right] \approx 1 - \alpha/2 - \alpha/2 = 1 - \alpha.$$

Since $\Phi^{-1}(x) = -\Phi^{-1}(1-x)$ for any $x \in (0,1)$, this implies

$$\mathbb{P}\left[\frac{|\hat{I}_n - I|}{\sigma_f/\sqrt{n}} \leq \Phi^{-1}(1-\alpha/2)\right] \approx 1-\alpha.$$

To conclude, note that

$$\frac{|\hat{I}_n - I|}{\sigma_f/\sqrt{n}} \leq \Phi^{-1}(1-\alpha/2) \Leftrightarrow I \in [\hat{I}_n - b, \hat{I}_n + b],$$

for $b = \frac{\sigma_f}{\sqrt{n}}\,\Phi^{-1}(1-\alpha/2)$.
[SEEN]

(iii) Employing the same reasoning as above,

$$\operatorname{Var}[\hat{I}_n^{cv}] = \frac{1}{n}\Big(\sigma_f^2 + \beta^2\,\sigma_g^2 + 2\,\beta\,\operatorname{Cov}[f(X_1), g(X_1)]\Big).$$

This is a convex quadratic polynomial in β, so the variance is minimized for

$$\beta = -\frac{\operatorname{Cov}[f(X_1), g(X_1)]}{\sigma_g^2}.$$

[SEEN]

2. (a) Since Brownian increments have mean zero and $W_0 = 0$ by definition,

$$\mathbb{E}[B_t] = \mathbb{E}[W_t - tW_1] = \mathbb{E}[W_t - W_0] - t\,\mathbb{E}[W_1 - W_0] = 0.$$

For the covariance, we have

$$\mathrm{Cov}(B_s, B_t) = \mathbb{E}[B_s\, B_t] = \mathbb{E}[W_s\, W_t - t\, W_s\, W_1 - s\, W_t\, W_1 + s\, t\, W_1\, W_1]. \tag{1}$$

For any $0 \leq a \leq b$,

$$\mathbb{E}[W_a\, W_b] = \mathbb{E}[W_a\, (W_b - W_a) + W_a^2] = 0 + \mathbb{E}[W_a^2] = a$$

because Brownian increments are independent. Therefore for any $a, b \geq 0$

$$\mathbb{E}[W_a\, W_b] = \mathbb{E}[W_{\min(a,b)}\, W_{\max(a,b)}] = \min(a, b),$$

Going back to (1)

$$\mathrm{Cov}(B_s, B_t) = \min(s, t) - t\, s - s\, t + s\, t = \min(s, t) - s\, t.$$

[SIMILAR]

(b) If (φ_n, λ_n) is an eigenpair of $\mathcal{K}$, then

$$(\mathcal{K}\varphi_n) = \lambda_n\, \varphi_n.$$

Employing the definition of $\mathcal{K}$,

$$\int_0^1 \Big(\min(s, t) - st\Big)\, \varphi_n(s)\, \mathrm{d}s = \lambda_n\, \varphi_n(t),$$

so

$$\int_0^t \Big(s - st\Big)\, \varphi_n(s)\, \mathrm{d}s + \int_t^1 \Big(t - st\Big)\, \varphi_n(s)\, \mathrm{d}s = \lambda_n\, \varphi_n(t).$$

Evaluating both sides at $t = 0$ and $t = 1$, we deduce that φ_n must satisfy $\varphi_n(0) = \varphi_n(1) = 0$. Differentiating both sides, we obtain

$$-\int_0^t s\, \varphi_n(s)\, \mathrm{d}s + \int_t^1 (1 - s)\, \varphi_n(s)\, \mathrm{d}s = \lambda_n\, \varphi_n'(t).$$

Differentiating again, we obtain

$$-t\, \varphi_n(t) - (1 - t)\, \varphi_n(t) = \lambda_n\, \varphi_n''(t),$$

which after simplification is the required equation.

[SIMILAR]

(c) Either $\lambda_n = 0$ or $\lambda_n < 0$ or $\lambda_n \geq 0$.

* If $\lambda_n = 0$, then $\varphi_n = 0$ by the differential equation. There is therefore no associated eigenfunction.

* If $\lambda_n < 0$, then the general solution to the differential equation is

$$\varphi_n(t) = A_n \sinh(t/\sqrt{-\lambda_n}) + B_n \cosh(t/\sqrt{-\lambda_n}).$$

The boundary condition $\varphi_n(0) = 0$ implies that $B_n = 0$, and the boundary condition $\varphi_n(1) = 0$ implies that λ_n satisfies

$$A_n \sinh(1/\sqrt{-\lambda_n}) = 0,$$

which implies that $A_n = 0$ and therefore $\varphi_n = 0$: there is no eigenfunction with negative associated eigenvalue.

* If $\lambda_n > 0$, then the general solution to the differential equation is

$$\varphi_n(t) = A_n \sin(t/\sqrt{\lambda_n}) + B_n \cos(t/\sqrt{\lambda_n}).$$

The boundary condition $\varphi_n(0) = 0$ implies that $B_n = 0$, and the boundary condition $\varphi_n(1) = 0$ implies

$$\lambda_n = \frac{1}{|k\pi|^2},$$

for some $k \in \mathbb{Z}$. Arranging the eigenvalues in decreasing order and normalizing the eigenfunctions, we deduce

$$\lambda_n = \frac{1}{|n\pi|^2}, \qquad \varphi_n = \pm\sqrt{2}\,\sin(n\,\pi\,t), \qquad n = 1, \dots$$

[UNSEEN]

(d) Using Fubini's theorem,

$$\mathbb{E}[Z_n] = \mathbb{E}\left[\int_0^1 B_t\,\varphi_n(t)\,\mathrm{d}t\right] = \int_0^1 \mathbb{E}[B_t]\,\varphi_n(t)\,\mathrm{d}t = 0, \qquad \forall n \in \mathbb{N}.$$

Similarly, for all $n, m \in \mathbb{N}$ it holds that

$$\begin{aligned}
\mathbb{E}[Z_n\,Z_m] &= \mathbb{E}\left[\int_0^1 B_s\,\varphi_n(s)\,\mathrm{d}s \int_0^1 B_t\,\varphi_m(t)\,\mathrm{d}t\right] = \mathbb{E}\left[\int_0^1\!\!\int_0^1 B_s\,B_t\,\varphi_n(s)\,\varphi_m(t)\,\mathrm{d}s\,\mathrm{d}t\right] \\
&= \int_0^1\!\!\int_0^1 \mathbb{E}[B_s\,B_t]\,\varphi_n(s)\,\varphi_m(t)\,\mathrm{d}s\,\mathrm{d}t = \int_0^1\!\!\int_0^1 C(s,t)\,\varphi_n(s)\,\varphi_m(t)\,\mathrm{d}s\,\mathrm{d}t \\
&= \int_0^1\!\!\int_0^1 C(s,t)\,\varphi_n(s)\,\mathrm{d}s\,\varphi_m(t)\,\mathrm{d}t = \int_0^1 \lambda_n\,\varphi_n(t)\,\varphi_m(t)\,\mathrm{d}t = \lambda_n\,\delta_{nm}.
\end{aligned}$$

[SEEN]

(e) By substituting the expansion,

$$\begin{aligned}
\int_0^1 \mathbb{E}\left[|S_N(t) - S_M(t)|^2\right] \mathrm{d}t &= \int_0^1 \mathbb{E}\left[\left|\sum_{i=N+1}^{M} Z_i\,\varphi_i(t)\right|^2\right] \mathrm{d}t \\
&= \int_0^1 \mathbb{E}\left[\sum_{i=N+1}^{M}\sum_{j=N+1}^{M} Z_i\,Z_j\,\varphi_i(t)\,\varphi_j(t)\right] \mathrm{d}t \\
&= \sum_{i=N+1}^{M}\sum_{j=N+1}^{M} \mathbb{E}[Z_i\,Z_j] \int_0^1 \varphi_i(t)\,\varphi_j(t)\,\mathrm{d}t \\
&= \sum_{i=N+1}^{M}\sum_{j=N+1}^{M} \lambda_i\,\delta_{ij}\,\delta_{ij} \\
&= \sum_{i=N+1}^{M} \lambda_i.
\end{aligned}$$

[UNSEEN]

3. (a) Note first that

$$\int_{n\Delta t}^{(n+1)\Delta t}\int_{n\Delta t}^{s} \mathrm{d}W_u \,\mathrm{d}W_s = \int_{n\Delta t}^{(n+1)\Delta t} (W_s - W_{n\Delta t})\,\mathrm{d}W_s$$
$$= \int_{n\Delta t}^{(n+1)\Delta t} W_s \,\mathrm{d}W_s - W_{n\Delta t}(W_{(n+1)\Delta t} - W_{n\Delta t}).$$

For the first term on the right-hand side, apply Itô's formula:

$$\mathrm{d}W_s^2 = \mathrm{d}s + 2W_s\,\mathrm{d}W_s,$$

which yields

$$\int_{n\Delta t}^{(n+1)\Delta t} W_s\,\mathrm{d}W_s = \frac{1}{2}(W_{(n+1)\Delta t}^2 - W_{n\Delta t}^2 - \Delta t).$$

Gathering the terms,

$$\int_{n\Delta t}^{(n+1)\Delta t}\int_{n\Delta t}^{s} \mathrm{d}W_u \,\mathrm{d}W_s = \frac{1}{2}\left(W_{(n+1)\Delta t}^2 - W_{n\Delta t}^2 - \Delta t\right) - W_{n\Delta t}W_{(n+1)\Delta t} + W_{n\Delta t}^2$$
$$= \frac{1}{2}\left(\Delta W_n^2 - \Delta t\right),$$

as required.

[SEEN]

(ii) Let $\mathcal{L}f(x) = \frac{1}{2}\,\sigma(x)^2\, f''(x)$, so that Itô's formula is

$$\mathrm{d}f(X_t) = \mathcal{L}f(X_t)\,\mathrm{d}t + f'(X_t)\,\sigma(X_t)\,\mathrm{d}W_t.$$

We write

$$X_{(n+1)\Delta t} = X_{n\Delta t} + \int_{n\Delta t}^{(n+1)\Delta t} \sigma(X_s)\,\mathrm{d}W_s. \tag{2}$$

Applying Itô's formula to the diffusion coefficient, we obtain

$$\sigma(X_s) = \sigma(X_{n\Delta t}) + \int_{n\Delta t}^{s} \mathcal{L}\sigma(X_u)\,\mathrm{d}u + \int_{n\Delta t}^{s} \sigma'(X_u)\,\sigma(X_u)\,\mathrm{d}W_u.$$

Substituting these formulas into (2) and writing $X_k = X_{k\Delta t}$, we get

$$X_{n+1} = X_n + \sigma(X_n)\Delta W_n$$
$$+ \int_{n\Delta t}^{(n+1)\Delta t}\int_{n\Delta t}^{s} \mathcal{L}\sigma(X_u)\,\mathrm{d}u\,\mathrm{d}W_s + \int_{n\Delta t}^{(n+1)\Delta t}\int_{n\Delta t}^{s} \sigma'\sigma(X_u)\,\mathrm{d}W_u\,\mathrm{d}W_s$$

We now discard terms of order higher than $O(\Delta t)$, i.e. we discard the first integral which is of order $O(\Delta t^{3/2})$. So we are left with

$$X_{n+1} \approx X_n + \sigma(X_n)\,\Delta W_n + (\sigma'\sigma)(X_n)\int_{n\Delta t}^{(n+1)\Delta t}\int_{n\Delta t}^{s} \mathrm{d}W_u\,\mathrm{d}W_s$$
$$= X_n + \sigma(X_n)\Delta W_n + \frac{1}{2}(\sigma'\sigma)(X_n)(\Delta W_n^2 - \Delta t)$$
$$= X_n + \sigma(X_n)\Delta W_n + \frac{1}{2}\Delta t(\sigma'\sigma)(X_n)(\xi_n^2 - 1),$$

where we used part (ii) and defined $\xi_n \sim \mathcal{N}(0,1)$.

[SIMILAR]

(b) (i) This follows from the definition of the scheme:

$$\begin{aligned} X_{n+1}^N &= X_n^N + \left(\theta\, X_{n+1}^N + (1-\theta)\, X_n^N\right) \Delta t + X_n^N\, \xi_n \sqrt{\Delta t} \\ \Leftrightarrow (1-\theta\Delta t)\, X_{n+1}^N &= (1 + (1-\theta)\,\Delta t + \xi_n \sqrt{\Delta t})\, X_n^N \\ \Leftrightarrow X_{n+1}^N &= \frac{1 + (1-\theta)\,\Delta t + \xi_n \sqrt{\Delta t}}{1-\theta\Delta t}\, X_n^N, \end{aligned}$$

which is of the required form with

$$G(\Delta t, \xi_n, \theta) = \frac{1 + (1-\theta)\,\Delta t + \xi_n \sqrt{\Delta t}}{1-\theta\Delta t}.$$

[SIMILAR]

(ii) From Part (i) and since $x_0 = 1$,

$$\begin{aligned} \left|\mathbb{E}[X_N^N - X_T]\right| &= \left|\mathbb{E}\left[\prod_{i=0}^{N-1} G(\Delta t, \xi_n, \theta) - X_T\right]\right| \\ &= \left|\prod_{i=0}^{N-1} \mathbb{E}[G(\Delta t, \xi_n, \theta)] - e^T\right| = \left|\left(\frac{1 + (1-\theta)\,\Delta t}{1-\theta\,\Delta t}\right)^N - e^T\right|. \end{aligned}$$

Since $\Delta t = T/N$,

$$\left|\mathbb{E}[X_N^N - X_T]\right| = \left|\frac{\left(1 + \frac{(1-\theta)\,T}{N}\right)^N}{\left(1 - \frac{\theta\,T}{N}\right)^N} - e^T\right|$$

By the given characterization of the exponential, as $N \to \infty$ the numerator tends to $e^{(1-\theta)T}$ and the denominator tends to $e^{-\theta T}$, so

$$\lim_{N\to\infty} \left|\mathbb{E}[X_N^N - X_T]\right| = \lim_{N\to\infty} \left|\frac{\left(1 + \frac{(1-\theta)\,T}{N}\right)^N}{\left(1 - \frac{\theta\,T}{N}\right)^N} - e^T\right| = \left|\frac{e^{(1-\theta)T}}{e^{-\theta T}} - e^T\right| = 0.$$

[UNSEEN]

4. (a) (i) Using Cauchy-Schwarz inequality, we have

$$|I_i|^2 \leq \int_{i\Delta t}^{(i+1)\Delta t} 1 \,\mathrm{d}s \int_{i\Delta t}^{(i+1)\Delta t} b(X_s)^2 \,\mathrm{d}s = \Delta t \int_{i\Delta t}^{(i+1)\Delta t} b(X_s)^2 \,\mathrm{d}s,$$

Since b is bounded, we deduce $|I_i|^2 \leq C\Delta t^2$, and therefore also $\mathbb{E}(I_i^2) \leq C\Delta t^2$.

[SIMILAR to question in 2017 exam]

(ii) Using Young's inequality with $\epsilon = \sqrt{\Delta t}$ we obtain

$$2|I_i M_i| \leq \frac{I_i^2}{\sqrt{\Delta t}} + \sqrt{\Delta t}\, M_i^2,$$

and therefore

$$\mathbb{E}[|I_i\, M_i|] \leq \frac{\mathbb{E}(I_i^2)}{2\sqrt{\Delta t}} + \frac{\sqrt{\Delta t}}{2}\mathbb{E}[M_i^2].$$

Using (i) and the fact that $M_i = \sigma \Delta W_i \Rightarrow \mathbb{E}(M_i^2) = \sigma^2 \Delta t$, we obtain the bound $\mathbb{E}(|I_i\, M_i|) \leq C\Delta t^{3/2}$ as required.

[SIMILAR to question in 2017 exam]

(iii) We write $X_{i+1} = X_i + I_i + M_i$, so $(X_{i+1} - X_i)^2 = I_i^2 + M_i^2 + 2I_i\, M_i$. Therefore,

$$\hat{\sigma}_N^2 = \frac{1}{N\Delta t}\left(\sum_{i=0}^{N-1} I_i^2 + 2\sum_{i=0}^{N-1} I_i\, M_i + \sum_{i=0}^{N-1} \sigma^2 \Delta W_i^2\right),$$

so

$$\mathbb{E}(\hat{\sigma}_N^2) = \sum_{i=0}^{N-1} \frac{\mathbb{E}[I_i^2]}{N\Delta t} + 2\sum_{i=0}^{N-1} \frac{\mathbb{E}[I_i\, M_i]}{N\Delta t} + \frac{\sigma^2}{N\Delta t}\underbrace{\sum_{i=0}^{N-1} \mathbb{E}[\Delta W_i^2]}_{=N\Delta t}.$$

We deduce

$$\mathbb{E}(\hat{\sigma}_N^2) - \sigma^2 = \sum_{i=0}^{N-1} \frac{\mathbb{E}(I_i^2)}{N\Delta t} + 2\sum_{i=0}^{N=1} \frac{\mathbb{E}(I_i\, M_i)}{N\Delta t}.$$

Taking absolute values and using parts (i) and (ii), we obtain

$$|\mathbb{E}(\hat{\sigma}_N^2) - \sigma^2| \leq \sum_{i=0}^{N-1} \frac{C\Delta t^2}{N\Delta t} + 2\sum_{i=0}^{N=1} \frac{C\Delta t^{3/2}}{N\Delta t} = C(\Delta t + \sqrt{\Delta t}),$$

as required.

[SIMILAR to question in 2017 exam]

(b) (i) The PDF is given by

$$f(x_1 \ldots x_n; \theta) = \left|\frac{1}{2\pi\Delta t}\right|^{N/2} \exp\left(-\frac{1}{2\,\Delta t}\sum_{k=0}^{N-1} |x_{k+1} - x_k + \theta\, b(x_k)\,\Delta t|^2\right).$$

[SIMILAR]

(ii) The logarithm of the PDF evaluated at $\hat{X}$ is given by

$$-\frac{N}{2}\ln(2\,\pi\,\Delta t) - \frac{1}{2\,\Delta t}\sum_{k=0}^{N-1} \left|\hat{X}_{k+1} - \hat{X}_k + \theta\, b(\hat{X}_k)\,\Delta t\right|^2.$$

Equaling the derivative with respect to θ to 0, we obtain

$$\sum_{k=0}^{N-1} \left(\hat{X}_{k+1} - \hat{X}_k + \hat{\theta}_{MLE}\, b(\hat{X}_k)\, \Delta t\right)\, b(\hat{X}_k) = 0,$$

which gives

$$\hat{\theta}_{MLE} = -\frac{\sum_{k=0}^{N-1} b(\hat{X}_k)\, \left(\hat{X}_{k+1} - \hat{X}_k\right)}{\sum_{k=0}^{N-1} |b(\hat{X}_k)|^2\, \Delta t}.$$

[SIMILAR]

(iii) In the case $b(x) = 1$,

$$\hat{X}_n = -\theta\, n\, \Delta t + \sqrt{\Delta t} \sum_{i=0}^{n-1} \xi_i, \qquad n = 0, \dots N,$$

because $\hat{X}_0 = 0$, so

$$\hat{\theta}_{MLE} = -\frac{\hat{X}_N}{N\, \Delta t} = \frac{\theta\, N\, \Delta t - \sqrt{\Delta t}\, \sum_{i=0}^{N-1} \xi_i}{N\, \Delta t} = \theta - \frac{\sum_{i=0}^{N-1} \xi_i}{N\, \sqrt{\Delta t}}.$$

Therefore, $\mathbb{E}[\hat{\theta}_{MLE}] = \theta$ and $\mathrm{Var}[\hat{\theta}_{MLE}] = \frac{1}{N \Delta t}$, so

$$\mathrm{MSE}[\hat{\theta}_{MLE}] = \frac{1}{N \Delta t}.$$

[UNSEEN]

5. (a) Simply set $Y = x + Z$, as in the Random Walk Metropolis Hastings algorithm.
[SEEN]

(b) Consider the probability $p(x, y) = \mathbb{P}[X_{n+1} = y \,|\, X_n = x]$. If $x \neq -y$, then

$$p(x, y) = \alpha(x, y)\, q(y\,|x).$$

If $x = -y$ then it is also possible that $X_{n+1} = y$ is obtained by rejecting another proposal, i.e.

$$p(x, y) = \alpha(x, -x)\, q(-x \,|\, x) + \sum_{z \in \mathbb{Z}} (1 - \alpha(x, z))\, q(z \,|\, x).$$

We can express both cases in the following compact form

$$p(x, y) = \alpha(x, y)\, q(y \,|\, x) + \delta_{-x}(y) \sum_{z \in \mathbb{Z}} (1 - \alpha(x, z))\, q(z \,|\, x).$$

[SIMILAR]

(c) π is a stationary distribution of $\{X_n\}_{n \in \mathbb{N}}$ if

$$\sum_{x \in \mathbb{Z}} \pi(x)\, p(x, y) = \pi(y), \quad \forall y \in \mathbb{Z}.$$

[SEEN]

(d) If $y \neq -x$, then $p(x, y) = q(y|x)\, \alpha(x, y)$ and $p(y, x) = q(x|y)\, \alpha(y, x)$. Since $q(y|x) = g(y - x) = q(x|y)$ because $g(\cdot)$ is even,

$$\begin{aligned}
\pi(x)\, p(x, y) &= \pi(x)\, q(y|x)\, \alpha(x, y) \\
&= \pi(x)\, q(x|y)\, \min\left(1, \frac{\pi(y)}{\pi(x)}\right) = q(x|y)\, \min\left(\pi(x), \pi(y)\right) \\
&= \pi(y)\, q(x|y)\, \min\left(\frac{\pi(x)}{\pi(y)}, 1\right) = \pi(y)\, q(x|y)\, \alpha(y, x) = \pi(y)\, p(y, x).
\end{aligned}$$

If $y = -x$, we need to show

$$\pi(x)\, p(x, -x) = \pi(-x)\, p(-x, x)$$

Since $\pi(x) = \pi(-x)$, it is sufficient to show $p(x, -x) = p(-x, x)$. Noticing that $\alpha(x, -x) = \alpha(-x, x) = 1$, we have

$$\begin{aligned}
p(x, -x) &= g(2x) \, + \sum_{z \in \mathbb{Z}} g(z - x)\, (1 - \alpha(x, z)) \\
p(-x, x) &= g(2x) \, + \sum_{z \in \mathbb{Z}} g(z + x)\, (1 - \alpha(-x, z)).
\end{aligned}$$

Therefore

$$p(x, -x) - p(-x, x) = \sum_{z \in \mathbb{Z}} g(z - x)\, (1 - \alpha(x, z)) - \sum_{z \in \mathbb{Z}} g(z + x)\, (1 - \alpha(-x, z)).$$

Using the substitution $z \to -z$ in the second sum and the fact that $\alpha(-x, -z) = \alpha(x, z)$, we obtain the result.
[UNSEEN]

(e) We write

$$\sum_{x\in\mathbb{Z}} \pi(x)\, p(x,y) = \sum_{x\in\mathbb{Z}} \pi(y)\, p(y,x) = \pi(y) \sum_{x\in\mathbb{Z}} p(y,x) = \pi(y).$$

[SEEN]

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 sperate pdf file with your email.

ExamModuleCode	QuestionNu	Comments for Students	
MATH97016 MATH97094	1	I am very pleased with the performance of the students for this question. In a)ii), many students did not notice that the acceptance probability was a strictly decreasing function of L, and so they reached the wrong conclusion. Also, some students struggled with the theory question b)iii), although the required proof was seen in class.	
MATH97016 MATH97094	2	This question was well answered overall. Although most students derived the correct differential equation in b), few gave a derivation of the associated boundary conditions. I was pleased that many students managed to do e), which was a bit more challenging than the other subquestions.	
MATH97016 MATH97094	3	I am very pleased with the performance of the students for this question. Few students managed to rigorously prove the limit in b)ii), but the other questions were answered correctly by most students.	
MATH97016 MATH97094	4	The students did very well in the theoretical part a). In b), most students found the correct expression for the maximum likelihood estimator, but few students obtained the correct expression for the associated mean-square error.	
MATH97016 MATH97094	5	This question was very well answered overall, and I was pleased that most students obtained the correct expression for the transition probability in b). A common error in part a) was to write $Y = Z - x$ instead of $Y = Z + x$; remember that, if $x > 0$, then the graph of $g(. - x)$ is obtained from that of $g(.)$ by a translation to the right, not to the left. Part d) was quite challenging, and only a few students gave a complete answer, with a careful examination of the case $x = -y$.	