Imperial College
London

M4/5A43

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)

May-June 2017

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

Inference, Control and Driving in Natural Systems

Date: Thursday 01 June 2017

Time: 10:00 - 12:30

Time Allowed: 2.5 Hours

This paper has 5 Questions.

Candidates should use ONE main answer book.

Supplementary books may only be used after the relevant main book(s) are full.

All required additional material will be provided.

- DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.
- Affix one of the labels provided to each answer book that you use, but DO NOT USE THE LABEL WITH YOUR NAME ON IT.
- Credit will be given for all questions attempted, but extra credit will be given for complete or nearly complete answers to each question as per the table below.

Raw Mark	Up to 12	13	14	15	16	17	18	19	20
Extra Credit	0	½	1	1 ½	2	2 ½	3	3 ½	4

- Each question carries equal weight.
- Calculators may not be used.

1. (a) The presence of a predator increases the concentration of a toxin, T, from its fluctuating background level to a higher, fluctuating, level. Both levels are Gaussian distributed. A primitive organism can move to avoid its predator and knows, with certainty, the parameters of the two possible Gaussian distributions.

   (i) By accounting for the primitive organism's prior beliefs appropriately, find the organism's assessment of the probability of the presence of a predator as a function of toxin level, T. By considering particular parameter choices for the Gaussian distributions, sketch an example of this function of T. [5]

   (ii) What influences the T at which it moves to avoid the predator? [1]

(b) (i) What is the Szilárd Engine? [4]

   (ii) Show that $\lim_{x \to 0} x \log_2 x = 0$ [1]

   (iii) A complex biomolecule can be found in two forms: Left and Right. At equilibrium the two forms occur equiprobably. A feeding organism knows that, in fact, a particular biomolecule is in state Left. How much free energy is available to it? You may use the result that if a system is in state i with probability u_i and has an equilibrium distribution with probability of state i being π_i then the free energy is $k_B T \sum_i u_i \log \frac{u_i}{\pi_i}$. [3]

   (iv) Multiple copies of the same biomolecule as considered above are all placed in the same non-equilibrium state: Left with a probability b=0.6. The same feeding organism knows the bias (towards Left) that these systems have away from equilibrium and decides to extract work from them before they equilibrate. Assuming that an incorrect guess has zero energy penalty, give the optimal strategy that the organism should consistently apply to all biomolecules to extract the most energy, and state if this strategy is deterministic. [5]

   (v) We assumed the organism knows b. How many bits are needed to encode the *relevant* information about b? [1]

2. Consider the dynamical system $\dot{y} = -\omega_0 y + G_0\omega_0 u$ with scalar $y(t)$ and $u(t)$ and initial condition $y(0) = 0$.

(a) (i) Find the response function, $y(s)/u(s)$, by taking a Laplace transform. [2]

(ii) By considering $s = i\omega$ sketch the Bode magnitude and phase plots for the response function. [3]

(iii) How does the system respond to different input frequencies and is there a characteristic frequency? [2]

(b) For the rest of the question let $y = r$ (with r =constant) be the desired system state.

(i) Define open loop control and briefly explain one of its shortcomings. [2]

(ii) Identify the most appropriate feedforward control. [1]

(iii) Suppose u is now a feedback proportional to the difference between r and y. Write down the differential equation for the system dynamics and find the steady state of y for large gains G_0. [2]

(iv) Find the form of $y(s)/r(s)$ and compare its characteristic frequency (for increasing gain G_0) to the open loop response function you found above. [3]

(v) Compare the open loop and closed loop steady-state sensitivities of y/r to small changes in the gain G_0. [3]

(c) Explain the relevance of Bode's integral formula for control system design. [2]

3. (a) A system has a scalar state x_t at time t and a scalar control u_t applied to it at time t. It has dynamics $x_{t+dt} = x_t + f(x_t, u_t, t)dt$, the (additive) cost of a control u_t at time t for duration dt is $w(x_t, u_t, t)dt$, and the final cost at time T is $W(x_t, T)$.

(i) Define the optimal cost to go. [2]

(ii) Give a recursive definition for the optimal cost to go at time t in terms of position x_t and control u_t and its value a time dt later (dt small). [2]

(iii) Derive the Hamilton-Jacobi-Bellman Equation. [2]

(iv) Provide a brief interpretation of the form of the Hamilton-Jacobi-Bellman Equation. [1]

(b) (i) Give one reason why, in Bayesian Inference, we sample from probability distributions. [1]

(ii) Outline the Metropolis Algorithm and briefly discuss why it works. [4]

(iii) How can optimization (and so optimal control) be connected to sampling? [1]

(iv) Given an example of when Importance Sampling is used. [1]

(v) Explain Importance Sampling as a method and outline a proof of its validity. [4]

(vi) Give one shortcoming of Importance Sampling and briefly explain its origin. [2]

4. (a) Call the stoichiometry matrix, S, flux vector, v, and constraint vector, c.

   (i) Outline the elements of flux balance analysis and account for its relevance. You might like to explain the meaning of the stoichiometry matrix, flux vector, and constraint vector. [3]

   (ii) Briefly explain the principles behind the simplex algorithm and explain its relevance to flux balance analysis. [3]

(b) Consider the reaction system $\phi \to A$; $A \to B$; $A \to C$; $B \to \phi$; $C \to \phi$; $C \to D$; $D \to \phi$.

   (i) Write down S for this system. [1]

   (ii) Write down a choice of c that corresponds to maximizing flux through D. Write down this maximal flux. [2]

   (iii) A large amount of A is instantaneously introduced into the system. What is the change in B, C, D a very short time later? [2]

   (iv) How does releasing small amounts of different species and investigating the immediate responses of other allow the reaction graph to be mapped? [1]

   (v) State how the subsequent steady state following release of A differs from the preceding steady state. [1]

(c) Consider the loop: $A \underset{\epsilon}{\overset{K}{\rightleftharpoons}} B \underset{\epsilon}{\overset{K}{\rightleftharpoons}} C \underset{\epsilon}{\overset{K}{\rightleftharpoons}} A$

   (i) Consider the chemicals to have concentrations $[A], [B], [C]$. Write differential equations for the dynamics of these concentrations. Solve for steady state if $\epsilon = 0$ and all concentrations start equal at $[A_0]$. [2]

   (ii) How would you simulate the full chemical system ($\epsilon, K \neq 0$) above with the Gillespie algorithm? How might there be numerical issues as $\epsilon \to 0$ when K is constant? [4]

   (iii) Write down a pair of K, ϵ in which the system would be in detailed balance. How much external work is required to drive the system in this case? [1]

5. By drawing on specific examples, and providing equations where appropriate, discuss how physical constraints can limit, and create trade-offs in, the actions of living systems. [20]

[M45A43 (2017) Solutions. Setter: Jones; Checker: Kantas]

1. (a) The presence of a predator increases the concentration of a toxin, T, from its fluctuating background level to a higher, fluctuating, level. Both levels are Gaussian distributed. A primitive organism can move to avoid its predator and knows, with certainty, the parameters of the two possible Gaussian distributions.

(i) By accounting for the primitive organism's prior beliefs appropriately, find the organism's assessment of the probability of the presence of a predator as a function of toxin level, T. By considering particular parameter choices for the Gaussian distributions, sketch an example of this function of T. [5]

** We seek $P(Pr|T)$ and know $P(T|Pr)$ and $P(T|nPr)$ (Pr meaning predator present and nPr the converse). $P(T|Pr)$ and $P(T|nPr)$ are Gaussians with means and variances $\mu_{Pr} > \mu_{nPr}$ and $\sigma_{Pr}, \sigma_{nPr}$. $P(Pr)$ is unspecified in the problem and is the organism's prior on the proportion of time a predator is present. Thus $P(Pr|T) = \frac{P(T|Pr)P(Pr)}{P(T)} = \frac{P(T|Pr)P(Pr)}{P(Pr)P(T|Pr)+(1-P(Pr))P(T|nPr)}$. The sketch is an important part of the answer helping stronger candidates point out subtleties like the importance of the relative scales of $\sigma_{Pr}, \sigma_{nPr}$ and $\mu_{Pr} - \mu_{nPr}$. A plot of a sigmoidal response of $P(Pr|T)$ vs T is likely. U

(ii) What influences the T at which it moves to avoid the predator? [1]

** An answer that seeks to evaluate the cost of moving (or not) will score. E.g. A given trade off between being near the predator and the cost of acting will yield a threshold for $P(Pr|T)$, and equivalently a range of values for T. Above this threshold the organism should move. U

(b) (i) What is the Szilárd Engine? [4]

** A particle in a box is treated as an ideal gas. The particle could be in either half of the box. If an observer knows which half the particle is in, it can insert a barrier half-way and allow the particle to do work on the barrier and so energy can be extracted of amount $W = k_B T \log 2$. If the observer is wrong then work is not extracted. There is thus an interplay between knowledge of system state and how much work can be extracted. Accept sensible variants. S

(ii) Show that $\lim_{x\to 0} x \log_2 x = 0$ [1]

** L'Hôpital's rule and setting $t = 1/x$ to obtain $-\frac{\log t}{t}$. U

(iii) A complex biomolecule can be found in two forms: Left and Right. At equilibrium the two forms occur equiprobably. A feeding organism knows that, in fact, a particular biomolecule is in state Left. How much free energy is available to it? You may use the result that if a system is in state i with probability u_i and has an equilibrium distribution with probability of state i being π_i then the free energy is $k_B T \sum_i u_i \log \frac{u_i}{\pi_i}$. [3]

** $\pi_L = \pi_R = \frac{1}{2}$. $u_L = 1$. Free energy is thus $k_B T(1 \times \log 2 + 0)$ using the result from the previous part. U

(iv) Multiple copies of the same biomolecule as considered above are all placed in the same non-equilibrium state: Left with a probability b=0.6. The same feeding organism knows the bias (towards Left) that these systems have away from equilibrium and decides to extract work from them before they equilibrate. Assuming that an incorrect guess has zero energy penalty, give the optimal strategy that the organism should consistently apply to all biomolecules to extract the most energy, and state if this strategy is deterministic.

[5]

** Payoff is $k_B T \log 2$ for correct guess and zero for incorrect. No penalty for getting the size of the payoff wrong – it only needs to be treated as a constant per molecule. Adopt strategy of guessing L with probability p. Expected pay-off is thus $k_B T(pb + (1-p)(1-b)) + 0 \times (p(1-b) + (1-p)b)$ where the first two terms are guessing correctly left or right and the second are for incorrect guesses. Maximizing with respect to p for fixed $b > \frac{1}{2}$ yields a deterministic strategy of always picking L $(p = 1)$. U

(v) We assumed the organism knows b. How many bits are needed to encode the *relevant* information about b? [1]

** 1 bit of information (since the strategy only depends on whether $b > \frac{1}{2}$). U

2. Consider the dynamical system $\dot{y} = -\omega_0 y + G_0\omega_0 u$ with scalar $y(t)$ and $u(t)$ and initial condition $y(0) = 0$.

(a) (i) Find the response function, $y(s)/u(s)$, by taking a Laplace transform. [2]

** $sy(s) = -\omega_0 y(s) + \omega_0 G_0 u(s)$ so $G(s) = y/u = G_0/(1 + s/\omega_0)$. S/U

(ii) By considering $s = i\omega$ sketch the Bode magnitude and phase plots for the response function. [3]

** Sketch of $|G(i\omega)| = G_0/\sqrt{1 + \frac{\omega^2}{\omega_0^2}}$; Sketch of $\varphi = \arctan[-\frac{\omega}{\omega_0}]$. S/U

(iii) How does the system respond to different input frequencies and is there a characteristic frequency? [2]

** System response: E.g. Acts like a low-pass filter or creates a $-\pi/2$ phase shift for large frequencies. $\omega = \omega_0$ characteristic frequency. S/U

(b) For the rest of the question let $y = r$ (with r =constant) be the desired system state.

(i) Define open loop control and briefly explain one of its shortcomings. [2]

** In open loop control the signal $u(t)$ is selected independent of $y(t)$. E.g. it is thus sensitive to errors in the model of the plant or sources of noise. S

(ii) Identify the most appropriate feedforward control. [1]

** $\dot{y} = 0$ and $y = r$ solve $y = uG_0\omega_0/\omega_0 = r$ so $u = r/K$ where K is set to G_0. S

(iii) Suppose u is now a feedback proportional to the difference between r and y. Write down the differential equation for the system dynamics and find the steady state of y for large gains G_0. [2]

** $\dot{y} = -\omega_0 y + G_0\omega_0 K(r - y)$. $\dot{y} = 0$ so $y = G_0 K(r - y)$ thus $y = \frac{rG_0}{1+G_0K}$ which goes to $y = r$ as G_0 goes large. S/U

(iv) Find the form of $y(s)/r(s)$ and compare its characteristic frequency (for increasing gain G_0) to the open loop response function you found above. [3]

** By LT $y/r = \frac{G_0\omega_0 K}{s+\omega_0(1+G_0K)}$. A term in s/ω_0 in the open loop case is swapped for one in $s/[\omega_0(1 + G_0K)]$ so if $G_0K \gg 1$ the closed-loop response frequency is faster. S/U

(v) Compare the open loop and closed loop steady-state sensitivities of y/r to small changes in the gain G_0. [3]

** Sensitivity of Q to P is $\frac{P}{Q}\frac{dQ}{dP}$. This is $S = \frac{G_0}{y/r}\frac{d(y/r)}{dG_0}$. For open loop $y/r = G_0 K$ so $S = 1$. For closed loop $y/r = G_0K/[1 + G_0K]$ so $S = [G_0 r/y] \times [(K + G_0K^2) - G_0K^2]/(1 + G_0K)^2$ simplifying to $S = 1/(1 + G_0K)$ which tends to zero as $G_0 \to \infty$. S/U

(c) Explain the relevance of Bode's integral formula for control system design. [2]

** Accept a range of qualitative answers like: Bode's integral formula suggests that the integrated frequency response of the log-magnitude of the sensitivity is a constant. This means that tailoring a feedback to make the system less sensitive at one frequency inevitably means making it more sensitive at another. S

3. (a) A system has a scalar state x_t at time t and a scalar control u_t applied to it at time t. It has dynamics $x_{t+dt} = x_t + f(x_t, u_t, t)dt$, the (additive) cost of a control u_t at time t for duration dt is $w(x_t, u_t, t)dt$, and the final cost at time T is $W(x_t, T)$.

(i) Define the optimal cost to go. [2]

** The cost to go from time t_i is the cost associated with a particular control function $u(t)$ on the range $t_i \leq t \leq T$. The optimal cost to go at time t_i is the cost associated with the optimal control u_o on this range (where u_o is chosen to minimize the cost to go). Or accept equations as an answer. S

(ii) Give a recursive definition for the optimal cost to go at time t in terms of position x_t and control u_t and its value a time dt later (dt small). [2]

** $J(t, x_t) = min_{u_t}[w(t, x_t, u_t)dt + J(t + dt, x_{t+dt})]$ so $J(t + dt, x_{t+dt}) = J(t + dt, f(x_t, u_t, t)dt + x_t)$ S

(iii) Derive the Hamilton-Jacobi-Bellman Equation. [2]

** Taylor expanding the above yields $J(x, t) \approx min_u[w(t, x, u)dt + J(t, x) + \partial_t J(t, x)dt + \partial_x J(t, x) f(x, u, t)dt]$ simplifying yields $-\partial_t J(t, x) = min_u[w(t, x, u) + f(x, u, t)\partial_x J(t, x)]$.

(iv) Provide a brief interpretation of the form of the Hamilton-Jacobi-Bellman Equation. [1]

** The optimal control u_{ot} at time t is such that the cost increase of applying it for dt exactly corresponds to the decrement in the optimal cost to go: $w(t, x, u_{ot}) = -\partial_t J(t, x) - f(x, u_{ot}, t)\partial_x J(t, x)]$ U

(b) (i) Give one reason why, in Bayesian Inference, we sample from probability distributions. [1]

** Accept other plausible answers but a good answer is: Because it is very hard to normalize many probability distributions of interest. S/U

(ii) Outline the Metropolis Algorithm and briefly discuss why it works. [4]

** Accept a range of answers demonstrating understanding: If we want to sample from the unnormalized distribution $P^*(x)$. Start the algorithm at location $x^{t=0}$. Use the proposal distribution $Q(x'|x^t)$ to generate x'. If $r = P^*(x')/P^*(x^t) > 1$ then $x^{t=1} = x'$. If $r = P^*(x')/P^*(x^t) \leq 1$ then let $x^{t+1} = x'$ with probability r and otherwise let $x^{t+1} = x'$. Then repeat the cycle. This exploits the fact that even though we can't normalize $P^*(x)$ we can calculate its relative sizes at two points. Or accept answers emphasising the fact that the algorithm is detailed balanced. S/U

(iii) How can optimization (and so optimal control) be connected to sampling? [1]

** A range of answers that show engagement with the ideas at play accepted. E.g. In optimization one seeks to e.g. maximize something. In sampling strategies one often seeks to sample in proportion to the relative size: a sampler would thus identify the maximum of a function in proportion to its relative height. One might thus e.g. modify samplers to enhance the amount of time spent at maxima. U

(iv) Given an example of when Importance Sampling is used. [1]

** To calculate expectations like $F = \int f(x)P(x)dx$ S

(v) Explain Importance Sampling as a method and outline a proof of its validity. [4]

** When seeking to calculate F above, when we only have $P^*(x)$, we sample instead from $Q(x)$ (normalized) many times, N. Reweight each draw from $Q(x)$ by $w_i =$

$P^*(x_i)/Q(x_i)$ and then use the approximation $F = \sum_{i=1}^{N} f(x_i)w_i/Z$ where $Z = \sum_{i=1}^{N} w_i$. Needs w_i finite.

This works since we can approximate the expectation above as $F \approx \sum_{i=1}^{N} f(x_i)/N$ where the x_i are drawn from $P(x)$. However we can't sample from $P(x)$. Instead write: $F = \int f(x)(P^*(x)/Z)dx = \int f(x)(P^*(x)/Z)(Q(x)/Q(x))dx = \int [\frac{f(x)P^*(x)}{ZQ(x)}]Q(x)dx$ with the normalizer Z. By the above $F \approx \sum_{i=1}^{N} [\frac{f(x_i)P^*(x_i)}{NZQ(x_i)}] = \sum_{i=1}^{N} f(x_i)w_i/Z$ where $Z = \int [P^*(x)/Q(x)]Q(x)dx \approx \sum_{i=1}^{N} w_i/N$. S/U

(vi) Give one shortcoming of Importance Sampling and briefly explain its origin. [2]

** Importance sampling can end up assigning very small weights to most samples from Q if Q is not close to P. As such, many samples from Q are required for good estimates of properties of P. This problem becomes particularly acute in higher dimensions where it is hard to match P and Q. U

4. (a) Call the stoichiometry matrix, S, flux vector, v, and constraint vector, c.

(i) Outline the elements of flux balance analysis and account for its relevance. You might like to explain the meaning of the stoichiometry matrix, flux vector, and constraint vector. [3]

** The stoichiometry matrix, S, has each row as a different species and each reaction as a different column. S_{ij} tells us the number of species i in reaction j. v is the amount of flux through each reaction. $\frac{dX}{dt} = S \cdot v$ gives the dynamics of X, the vector of amounts of each chemical species. Flux balance analysis assumes $\frac{dX}{dt} = 0$. The fluxes at steady state are thus $S \cdot v = 0$. A constraint, e.g. find v to maximize $c \cdot v$, encodes the desirable fluxes through the system. This is relevant as it takes a hard problem which has dynamics depending on (hard to find) kinetic constants and, by considering steady state, turns it into a linear programming problem that relies only on an (easy to find) matrix of stoichiometries. S/U

(ii) Briefly explain the principles behind the simplex algorithm and explain its relevance to flux balance analysis. [3]

** Optimization of the fluxes v given c and S is a linear programming problem: $v \geq 0$ and often one defines the largest element of v to be v_{max}. Thus the condition $c \cdot v$ is a linear function on a convex polytope given by $S \cdot v = 0$, $v \geq 0$, $v \leq v_{max}$; $\exists i$ s.t. $v_i = v_{max}$.

The simplex algorithm can be used to optimize linear cost functions on convex polytopes. The optimal solution will always be on an extreme of the polytope (vertex or edge) and there will always be only one such extreme. The objective function will always increase, decrease or stay the same along an edge. So from any vertex moving along an adjacent edge that increases the objective function, and repeating until this is not possible, will, by convexity of the search space, lead to the extreme. S/U

(b) Consider the reaction system $\phi \to A$; $A \to B$; $A \to C$; $B \to \phi$; $C \to \phi$; $C \to D$; $D \to \phi$.

(i) Write down S for this system. [1]

** Reaction by reaction with elements (A, B, C, D): $(1,0,0,0)$; $(-1,1,0,0)$; $(-1,0,1,0)$; $(0,0,-1,1)$; $(0,-1,0,0)$; $(0,0,-1,0)$; $(0,0,0,-1)$. U

(ii) Write down a choice of c that corresponds to maximizing flux through D. Write down this maximal flux. [2]

** $c^T = (0,0,0,1)$ and $v^T = (1,0,1,1,0,0,1)$. U

(iii) A large amount of A is instantaneously introduced into the system. What is the change in B, C, D a very short time later? [2]

** An ODE representation would be $[\dot{B}] = K_{AB}[A]$; $[\dot{C}] = K_{AC}[A]$; $[\dot{D}] = K_{CD}[C]$. So in a time increment Δt as no direct connection to A $[\dot{D}] = 0$ but others are $\Delta B = K_{AB}[A]\Delta t$ and $\Delta C = K_{AC}[A]\Delta t$. Accept a variety of similar answers. U

(iv) How does releasing small amounts of different species and investigating the immediate responses of other allow the reaction graph to be mapped? [1]

** Immediate, downstream, neighbours of the selected node can be identified as they change their state shortly after the chemical is introduced. U

(v) State how the subsequent steady state following release of A differs from the preceding steady state. [1]

** It is exactly the same. U

(c) Consider the loop: $A \underset{\epsilon}{\overset{K}{\rightleftharpoons}} B \underset{\epsilon}{\overset{K}{\rightleftharpoons}} C \underset{\epsilon}{\overset{K}{\rightleftharpoons}} A$

(i) Consider the chemicals to have concentrations $[A], [B], [C]$. Write differential equations for the dynamics of these concentrations. Solve for steady state if $\epsilon = 0$ and all concentrations start equal at $[A_0]$. [2]

** $[\dot{A}] = (-K-\epsilon)[A] + K[C] + \epsilon[B]$; $[\dot{B}] = (-K-\epsilon)[B] + K[A] + \epsilon[C]$; $[\dot{C}] = (-K-\epsilon)[C]+K[B]+\epsilon[A]$. All rates of change zero at steady state and all concentrations equal to $[A_0]$. U

(ii) How would you simulate the full chemical system ($\epsilon, K \neq 0$) above with the Gillespie algorithm? How might there be numerical issues as $\epsilon \to 0$ when K is constant? [4]

** Using the next reaction method. Assume an initial species number A, B, C and calculate the event rates $AK, BK, CK, A\epsilon, B\epsilon, C\epsilon$. Total rate tr is sum of these rates. Sample from exponential with this total rate to give the time of the next reaction. Identify which reaction occurs with a probability of K/tr for the reactions with rate K and probability of ϵ/tr for the remainder. Increment the species numbers consequent on which reaction occurred.

If $\epsilon \ll K$ then there could be issues with e.g. simulating a coin with bias ϵ/tr as this will be very small. U

(iii) Write down a pair of K, ϵ in which the system would be in detailed balance. How much external work is required to drive the system in this case? [1]

** $K = \epsilon$. No work. U

5. By drawing on specific examples, and providing equations where appropriate, discuss how physical constraints can limit, and create trade-offs in, the actions of living systems. [20]

** Strong candidates will draw from the auxiliary literature provided and beyond. Good answers might develop:

– A good model of the world requires a rich specification in bits. This has a run cost, at least, in error correction of those bits. This run cost must be smaller than the advantage that having the rich model provides. Bacteria have simple models of the world as having more complicated ones is, among other things, not energy efficient.

– Sampling from distributions is time consuming and might have an energetic expense. As such it might be better to make trade off accuracy with speed by accepting very approximate estimates based on a very few samples. E.g. some distributions have a high dimensional character where most of their mass is localized in a particular part of the space: a single sample gives a huge amount of information about that distribution.

– In Kullback-Leibler control the cost function can naturally be related to the free energy between the distribution under the passive dynamics and the distribution under the control. The energetic cost of a control can then be linked to the energetic return (in terms of feeding) of that action.

– In neuroscience, in probabilistic population coding there can be a coincidence between the dimensionality of the distribution being represented and the number of neurons required: with the number scaling exponentially in the dimensionality. This makes probabilistic population coding appropriate for encoding low dimensional distributions but sampling-based approaches are more natural in higher dimensions.

Imperial College London
Department of Mathematics

Examiner's Comments

Exam: M4/5 A43 Session: 2016-2107

Question 1

Please use the space below to comment on the candidates' overall performance in the exam. A brief paragraph highlighting common mistakes and parts of questions done badly (or well) is sufficient. Do not refer to individual candidates. The purpose of this exercise is to provide guidance to the external examiners, and to the candidates themselves, on how you feel the cohort faired. Your comments will be available to students online.

Many candidates found the first part of this question challenging. This was odd as similar content had been covered in a practical session. Performance on the second part, b, was much stronger.

Marker: N Jones

Signature: [signature] Date: 8/6/17

Please return with exam marks (one report per marker)

Imperial College London
Department of Mathematics

Examiner's Comments

Exam: M4/5 A43 Session: 2016-2107

Question 2

Please use the space below to comment on the candidates' overall performance in the exam. A brief paragraph highlighting common mistakes and parts of questions done badly (or well) is sufficient. Do not refer to individual candidates. The purpose of this exercise is to provide guidance to the external examiners, and to the candidates themselves, on how you feel the cohort faired. Your comments will be available to students online.

This was a relatively straightforward question and the moderate average performance can be attributed to particular candidates not knowing core elements / time etc.

Marker: N Jones

Signature: N Date: 8/6/17

Please return with exam marks (one report per marker)

Imperial College London
Department of Mathematics

Examiner's Comments

Exam: M4/5 A43 **Session: 2016-2107**

Question 3

Please use the space below to comment on the candidates' overall performance in the exam. A brief paragraph highlighting common mistakes and parts of questions done badly (or well) is sufficient. Do not refer to individual candidates. The purpose of this exercise is to provide guidance to the external examiners, and to the candidates themselves, on how you feel the cohort faired. Your comments will be available to students online.

Responses to the first part were buggy - in fact poor handwriting was a significant factor among the students for many of the questions. Though the second part of the question was straightforward some candidates struggled to give a clear account of importance sampling.

Marker: N Jones

Signature: NJ **Date:** 8/6/17

Please return with exam marks (one report per marker)

Imperial College London
Department of Mathematics

Examiner's Comments

Exam: M4/5 A43 Session: 2016-2107

Question 4

Please use the space below to comment on the candidates' overall performance in the exam. A brief paragraph highlighting common mistakes and parts of questions done badly (or well) is sufficient. Do not refer to individual candidates. The purpose of this exercise is to provide guidance to the external examiners, and to the candidates themselves, on how you feel the cohort faired. Your comments will be available to students online.

Answers to the first part were mostly sound though opportunities to show careful understanding, given the number of marks assigned, were missed. Answers to the second part were good though, as occurred throughout, the really challenging parts (typically assigned few marks) with opportunities to quickly show deep understanding were neglected.

Marker: N Jones

Signature: [signature] Date: 8/6/17

Please return with exam marks (one report per marker)

Imperial College London
Department of Mathematics

Examiner's Comments

Exam: M4/5 A43 Session: 2016-2107

Question 5

Please use the space below to comment on the candidates' overall performance in the exam. A brief paragraph highlighting common mistakes and parts of questions done badly (or well) is sufficient. Do not refer to individual candidates. The purpose of this exercise is to provide guidance to the external examiners, and to the candidates themselves, on how you feel the cohort faired. Your comments will be available to students online.

I was pleased with the overall quality of the responses – in the past candidates have returned incomprehensible responses and this was not the case here. The responses showed clear grasp of the relevant course content with the better responses having a clearer connection to physical constants.

Marker: N Jones

Signature: [signature] Date: 8/6/17

Please return with exam marks (one report per marker)