DUBLIN CITY UNIVERSITY

SEMESTER ONE REPEAT EXAMINATION 2005

MODULE:	CA540 Computational Biology
COURSE:	M. Sc. in BioInformatics
YEAR:	1
EXAMINERS:	Mr. P. Cunningham, Dr. Martin Crane, (ext. 8974)
TIME ALLOWED:	3 hours
INSTRUCTIONS:	Attempt **four** questions. If you answer more than four, please indicate which questions you wish to be marked. All questions carry equal marks.
REQUIREMENTS:	Mathematical Tables, Graph Paper

THE USE OF PROGRAMMABLE OR TEXT STORING CALCULATORS IS EXPRESSLY FORBIDDEN

DO NOT TURN OVER THIS PAGE UNTIL INSTRUCTED TO DO SO.

QUESTION 1

Consider the general Leslie matrix:

$$L = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & 0 & 0 \\ 0 & b_2 & 0 \end{pmatrix}$$

where the terms in the matrix have their usual meanings.
(a) Show that the characteristic equation is given by:

$$\lambda^3 - a_1\lambda^2 - a_2 b_1 \lambda - a_3 b_2 b_1 = 0$$

[8 marks]

(b) If λ_1 is the first eigenvalue and

$$\mathbf{u} = \begin{pmatrix} 1 \\ b_1/\lambda_1 \\ b_1 b_2/\lambda_1^2 \end{pmatrix},$$

show that $\mathbf{u}$ is the corresponding eigenvector to λ_1.

[3 marks]

(c) For a certain population, the Leslie matrix is given by:

$$L = \begin{pmatrix} 0 & 7 & 6 \\ \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

Show that this matrix has an eigenvalue of $\frac{3}{2}$. Using the above formula determine the corresponding eigenvector.

[8 marks]

(d) If $\lambda_1, \mathbf{u}$ are the dominant eigenpair, show that the population will be stable with distribution 18:3:1.

[6 marks]

QUESTION 2

A factor that can affect the distribution of genes in a population is *mutation*. This happens when external factors (e.g. background radiation) cause an allele A to change to an allele α (say). The rate at which mutation occurs is usually very small (at fraction 10^{-5} per generation) and more commonly acts to change a dominant gene into a recessive gene. A big question is to find the growth of proportion of the mutants in the population. We assume random mating, equal survival and, in addition, we assume that A-alleles mutate to α-alleles at a rate μ per generation.

(a) Show, that, without mutation but with random mating and equal survival, at the end of the generation $k+1$, the numbers with each genotype, $\mathcal{N}_{k+1}()$ (in terms of the total number of fertilized eggs at the start of that generation N^*_{k+1}) is given by:

$$\begin{aligned} \mathcal{N}_{k+1}(AA) &= r\left[P_k(A)\right]^2 N^*_{k+1} \\ \mathcal{N}_{k+1}(A\alpha) &= 2rP_k(A)P_k(\alpha)N^*_{k+1} \\ \mathcal{N}_{k+1}(\alpha\alpha) &= r\left[P_k(\alpha)\right]^2 N^*_{k+1} \end{aligned}$$

where r is the survival fraction from the beginning to the end of the generation and $P()$ is the proportion of a particular allele in the population.

[10 marks]

(b) If μ is the mutation rate and

$$P_{k+1}(\alpha) = \frac{no.\ of\ \alpha - alleles\ +\ \mu \times (\ no.\ of\ A - alleles)}{total\ no.\ of\ alleles\ in\ gene\ pool}$$

show that the proportion of α-alleles at the end of year $k+1$ is given by:

$$P_{k+1}(\alpha) = (1-\mu)P_k(\alpha) + \mu$$

[10 marks]

(c) Hence derive an expression for $P_{k+1}(\alpha)$ in terms of the initial proportion of α-alleles P_0

[5 marks]

QUESTION 3

(a) Using (if you like) the logistic growth model as an example, describe the key difference between discrete and continous models of change.

[4 marks]

(b) Starting with the continuous Logistic Growth model:

$$\frac{dN}{dt} = kN\left(1 - \frac{N}{K}\right)$$

(where N is the number in the population), Derive, using partial fractions or otherwise, the general solution:

$$N(t) = \frac{K}{\left(\frac{K}{N_0} - 1\right)e^{-kt} + 1}$$

where N_0 is the initial value of N)

[12 marks]

(c) Sketch the curve of $N(t)$ against t for the following initial conditions:

- $N_0 > K/2$
- $N_0 < K/2$

[5 marks]

and describe qualitatively the shape of the curves.

[4 marks]

QUESTION 4

(a) Write down the equation for unconstrained (Malthusian) growth. What are the assumptions behind this model?

[8 marks]

(b) Starting with the equation you have written down, derive an expression for the doubling time of a species undergoing such growth.

[7 marks]

(c) According to Bruckman's book on cancer:

> with most cancers, you could only begin to detect a lump when the number of cancer cells reached approximately one billion (10^9).

Assuming the initial cancer cell starts off on January 1 2000 and reproduces once every month, what is the earliest that it would become detectable?

[5 marks]

Assuming that 100 billion cells weigh about 125g and that a human body can tolerate at most 2.5 kg of tumour, how long roughly before such a tumour becomes lethal?

[5 marks]

QUESTION 5

The Guerrilla Combat model is a greatly simplified model of combat on a battlefield. Using the following assumptions

1. the number of 'friendly' soldiers is x and 'enemy' (i.e. guerrilla forces) is y at time t.
2. the rate at which each soldier shoots is R_x, R_y (assumed constant) for respective armies, the subscript denoting the army firing,
3. the probability that a single shot hits the target is P_x, P_y for respective armies, the subscript denoting the army firing,

4. the number of soldiers is so large as to permit approximation by continuous variables.

(a) derive the following equations for guerrilla combat:

$$\begin{aligned} \dot{x} &= -ay \\ \dot{y} &= -bxy \end{aligned} \tag{1}$$

(where $\dot{x}, \dot{y}$ denote the change in x, y respectively with time)

[12 marks]

The linear Lanchester equations (as described in the Guerrilla Combat model above eqn(1)) do not take into account the fact that an increasing number of combatants of one army in a battlefield implies an easier target for the enemy! Thus the attrition rate for x should depend not only on the size of y and the effectiveness of y's weapons but also on the size of x itself. The nonlinear Lanchester equations, including further the possibility of reinforcements and logistic constraints are

$$\begin{aligned} \dot{x} &= -k_2 xy + x\,(r_1 - a_1 x) \\ \dot{y} &= -k_1 xy + y\,(r_2 - a_2 y) \end{aligned}$$

(b) Interpret each of the terms in the above equations.

[7 marks]

(c) Are the Lanchester equations a realistic description of modern battlefields?

[6 marks]

QUESTION 6

(a) Give two reasons why it is not feasible to model infectious diseases as predator-prey systems.

[4 marks]

(b) Give the assumptions behind the Susceptibles-Infectives- Removed (SIR) model of infectious diseases

[4 marks]

(c) Starting with the equations for the SIR model:

$$\begin{aligned} \dot{S} &= -\beta IS \\ \dot{I} &= \beta IS \quad -\nu I \\ \dot{R} &= \nu I \end{aligned}$$

(where a dot over a letter indicates differentiation w.r.t. time and β, ν are the so-called infection and removal rates), derive an expression for I in terms of S.

[8 marks]

(d) Roughly graph a phase-plane plot of I against S (the position of $\rho = \nu/\beta$ should be clearly visible)

[5 marks]

(e) Describe *qualitatively* what happens for $S < \rho$ and $S > \rho$.

[4 marks]