# MATH 495: Homework #7

Spring 2019

**Due: Monday, May 6, 2019**

---

Solve the following related to stochastic models of carcinogenesis. Note that this is the final homework assignment of the semester. Please hand in a **physical copy** to my office by **6 pm on Monday, May 6**$^{\text{th}}$.

1. Assuming a two-stage disease (i.e. two mutations are required for malignancy), a system of three differential equations describing the probability of a cell being in each stage ($i = 0, 1, 2$) at time $t$ can be derived (in class) as below:

$$\begin{aligned} \dot{p}_0 &= -\lambda_0 p_0, \\ \dot{p}_1 &= \lambda_0 p_0 - \lambda_1 p_1, \\ \dot{p}_2 &= \lambda_1 p_1. \end{aligned} \tag{1}$$

   More precisely, $p_i(t)$ denotes the probability that a **single** cell is in stage $i$ at time $t$. Here stage 0 denotes a normal (healthy) cell, and stage 2 a fully malignant cell. Assume, for simplicity, that the $\lambda_i$ are **constant and distinct.**

   (a) Assuming the cell is healthy at time $t = 0$, write down a set of initial conditions for the system (1).

   (b) Solve the IVP described by equations (1) with initial conditions as in part (a).
   *Hint:* Start with the $p_0$ equation, and work down from there.

   (c) The hazard function $h(t)$ represents the instantaneous rate of developing the disease, i.e. the incidence rate for a random event. If we assume a population of $N$ cells acting independently, it can be shown (again, in class), that $h(t)$ is given by

$$h(t) = N \frac{\dot{p}_2(t)}{1 - p_2(t)}.$$

   Assuming that $p_2(t) \approx 0$, show that $h(t)$ is, to first order, a line. Find its slope. For more details, see Resource [14].

2. Consider the two-stage Armitage-Doll model as in Problem #1. Suppose that the rate parameters are as follows (the units are per year):

$$\lambda_0 = 0.01$$
$$\lambda_1 = 0.15$$

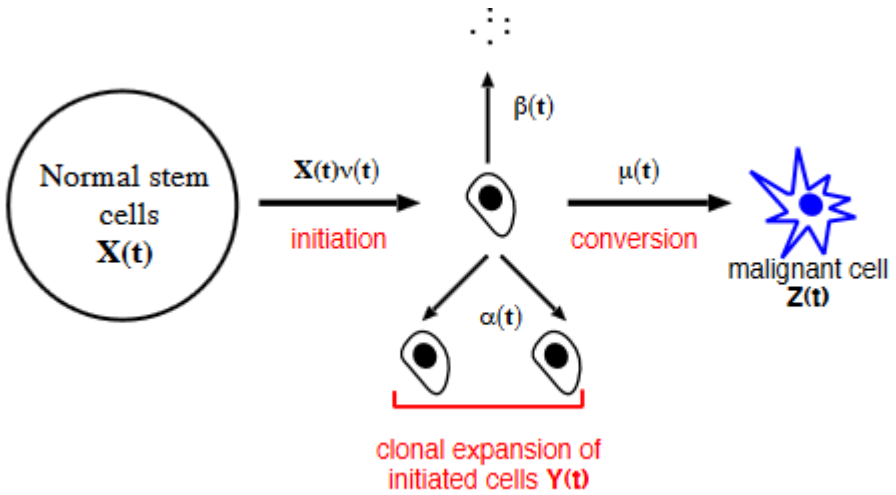
A (very basic) MATLAB function (*armitage_doll_two_stage.m*) is provided which simulates one realization of the two-stage model for parameters $\lambda_1$ and $\lambda_2$.

(a) Create a histogram to understand the distribution of **10000** times to cancer progression. That is, use *armitage_doll_two_stage.m* to create 10000 times which represent the time to a cell first reaching stage $E_2$. Provide a histogram of your computed distribution. Make sure each bin of the histogram spans 10 years.

(b) Compute the expected time to reach stage $E_2$ using your numerical results from part (a).

(c) What is the theoretical value for this mean? Does it agree with your result from part (b)?

3. Consider the two-stage clonal expansion model (TSCE) represented schematically below:




Here,

$$X(t) = \text{Number of healthy cells at time } t$$
$$Y(t) = \text{Number of initiated cells at time } t$$
$$Z(t) = \text{Number of malignant cells at time } t$$

That is, initiation is modeled as a non-homogeneous Poisson process with rate $\nu(t)X(t)$, and initiated cells undergo a Markovian birth/death/mutation (conversion) process. Understanding the dynamics is then equivalent to understanding the distributions of $Y(t)$ and $Z(t)$ ($X(t)$ is assumed given by some growth law):

$$p_{i,j}(t) = \mathbb{P}(Y(t) = i, Z(t) = j),$$

for non-negative integers $i, j$.

(a) Fix a small amount of time $\Delta t$. By conditioning the system on the state at time $t$, derive a difference equation (the so called Chapman-Kolmogorov equations) for $p_{i,j}(t+\Delta t)$ in terms of $p_{i,j}(t), p_{i-1,j}(t), p_{i+1,j}(t)$, and $p_{i,j-1}(t)$.

(b) Take the limit as $\Delta t \to 0$ in your equation in part (a) to obtain a system of differential equation describing $\frac{d}{dt}p_{i,j}(t)$.