Math 142 - Homework 6
Due Wednesday, May 13th

(1) **Stochastic simulations of population growth**

Modify the Matlab code stochasticBirth.m (posted on Piazza) to simulate cell population growth, with a birth rate $\lambda = 0.4$.

(a) Starting with a step size $\Delta t = 0.001$ and an initial population size of $N(0) = 1$, modify the code to run 10,000 stochastic simulations (realizations). Calculate the distribution of population sizes (i.e., the number of realizations with $N = 1$, $N = 2$, ... individuals) at time $t = 5$.

In class, we derived that the probability that there are exactly n organisms in the population at time t is

$$P_n(t) = e^{-\lambda t}(1 - e^{-\lambda t})^{n-1}.$$

Confirm this result by making a plot comparing $P_n(5)$ as a function of n with the estimates coming from your simulations.

(b) We argued in class that the simulations should be run with a small step size. However, provided that the step size is small enough, it does not affect the results. Confirm this by re-running the simulations from part (a) using a step size of $\Delta t = 0.005$ and plotting the distribution of population sizes. What happens if you run the simulation with a step size of $\Delta t = 1$? Do your simulations still reproduce the predicted distribution of population sizes? If not, why? (An answer of "the step size is too big" is not enough.)

(c) Now consider what happens if, in addition to cells dividing, the cells in the population may also die, at rate μ. That is, in each interval $[t, t + \Delta t]$, the likelihood of a given cell dying is $\mu\Delta t$. Explain in words how you would need to modify your stochastic simulation algorithm to include cell death. (Hint: Consider the number of cells that are born and the number that die separately. Explain why you don't need to worry, so long as Δt is small, about the probability of a cell both dividing and dying in the same time interval).

(d) Implement your stochastic simulation from part (c) for the same step size and total time as in part (a), with $\lambda = 0.4$ and $\mu = 0.2$. By running 10,000 simulations, make a plot of the average population size as a function of time. Characterize the average growth. Is it exponential? If so, what is the rate of growth? (You can calculate the growth rate by finding the slope of the log of the average cell population).

(2) **Stochastic simulations of a death process**

We will now build a stochastic model for how a population of bacteria is depleted by an antibiotic. Assume that at time $t = 0$, there are precisely N bacteria in the population. Once the antibiotic is added, the bacteria stop dividing (that is, $\lambda = 0$), and start to die with a mortality rate μ. That is, in time Δt, the probability of a given cell dying is $\mu\Delta t$. We want to recalculate the probability mass function $P_n(t)$, which represents the probability that there are n cells at time t.

(a) Explain what the initial conditions on $P_n(t)$ are.

(b) Show that

$$\frac{dP_N}{dt} = -N\mu P_N \quad \text{and}$$
$$\frac{dP_n}{dt} = (n+1)\mu P_{n+1} - n\mu P_n, \quad n = 0, 1, 2, \ldots, N-1.$$

(c) Using the equations from part (b), find $P_N(t)$, $P_{N-1}(t)$ and $P_{N-2}(t)$. Can you see any pattern that would help you to guess $P_n(t)$? (Reading Chapter 36 in Haberman and your class notes might help you with this. You do not need to prove that your answer is correct.)

(d) Verify that your solution for $P_n(t)$ is correct by comparing it with stochastic simulations. Specifically, implement a stochastic simulation for the death of bacteria. Use at least 1,000 simulations, and run them for time $t = 10$, with $\Delta t = 0.001$, $\mu = 0.1$, and $N = 20$.

(e) It is hard to see how, on average, the population will change over time, using the formulas for $P_n(t)$. But we can get this average directly. Suppose that at time t, the average number of cells in the population is $\bar{n}(t)$. Between time t and $t + \Delta t$, how many cells, on average, will die? Use this information to derive the following differential equation:

$$\frac{d\bar{n}}{dt} = -\mu\bar{n}\ .$$

(f) Solve the differential equation from part (e) and show that it agrees with the average population size according to stochastic simulations. Use the same parameter values as in part (d).

(g) Run a stochastic simulation for a long enough time for the bacteria population to reach zero (print out a plot showing this outcome). Does your solution from part (f) reach zero in finite time? Which model (stochastic or deterministic) do you think is more realistic when the amount of bacteria is very small?

(3) **Population genetics**

The Moran process is a model that describes how the diversity of populations changes with time. Consider a population of $N = 10$ cells. These cells have some form of diversity (i.e., different genes): to keep the biology simple, let's imagine that initially half of the cells are colored red and the other half are colored green. At each time step, exactly one cell divides. We pick this cell at random, and make a copy of it (if the original cell is red, then its child will also be red, etc.). Because of overcrowding, the overall number of cells must remain constant (i.e., equal to N). Each time a cell divides, one must be pushed out of the population. So pick one of the original N cells and remove it from the population (in particular, the cell that is removed could be the same or different as the one that divides).

(a) Do you expect the average number of red cells (initially 5) to change over time? Explain why or why not.

(b) Use the Matlab code Moran.m to simulate a single Moran process for a population of $N = 10$ cells. Show that after some finite time, your simulated population becomes only red or only green (homogeneous).

(c) Simulate 100 populations of $N = 10$ cells. Show that they all eventually become homogeneous, in the same way as in part (b).

(d) How do you reconcile your answer from part (c) with your answer from part (a)? Can you describe in words why the populations eventually homogenize?

(e) Use simulations with different values of N to show that the average time for the population to become homogeneous is proportional to N^2. Calculate the average time for the population to become homogeneous for each N by running 100 stochastic simulations. Plot these data in a way that makes it clear that the time to become homogeneous is proportional to N^2.