Assignment 4: The Luria-Delbruck Experiment *in silico*

BIOL224H - Mathematics of Life - Spring 2023

The objective of this exercise is to compare the results from simulating the competing hypotheses from Luria and Delbrück's classic 1943 paper on the origin of mutations

For the exercise, please create Matlab files with the code blocks below (i.e. invoke the filename of your new files on the Matlab command line, rather than pasting the code below into the command line). You will, as usual, need to Matlab which directory to find the files in, and to save your output.

Given the wellness days next week, this assignment will be due on Thursday, Feb 16th, at 9:30am (rather than Tuesday).

Mutation hypothesis

The following MATLAB code simulates a population of cells undergoing some number of cell divisions, given a certain probability that a resistance mutation will occur in a bacterium during each generation. Once a mutation occurs, all of that cell's daughters will carry the mutation in subsequent cell divisions. A number of replicate cultures are simulated. You can modify that and the other parameters at the top of the program. After this code block runs, the vector `mut` will contain the number of resistant bacteria in each culture (equivalent to how many resistant colonies would be seen by plating the full culture).

```
% set variables
prob_mutation = 1e-8;
cultures = 30;
generations = 20;
initial_inoculum = 100;

mut=zeros(cultures,1);
for i=1:cultures
   n=initial_inoculum;
     for t=1:generations
         n=2*n;
         new_mutations = poissrnd(n*prob_mutation);
         mut(i) = 2*mut(i)+new_mutations;
         n = n-new_mutations;
      end
    end
```

Problem 1

Explain in your own words how this algorithm produces the distribution under the random 'mutation' hypothesis of Luria and Delbrück (1943), with reference to *each line* in the code.

Problem 2

To visualize the resulting numbers of resistant bacteria in each culture, we can ask Matlab to make a histogram of the distribution of resistant bacteria in each culture.

```
hist(mut);
xlabel('resistant bacteria');
ylabel('number of cultures')
title('mutation hypothesis')
```

We can aksi use `mean(mut)` and `var(mut)` to compute the Fano factor, defined as the variance divided by the mean, for a set of replicate cultures.

(a) Submit the histogram you obtain.

(b) What is the Fano factor for the distribution of resistant bacteria that you found among your cultures?

Acquired immunity hypothesis

Problem 3

(a) Under the 'acquired immunity' hypothesis, what mutation probability would give the same mean number of resistant bacteria as you saw for the 'mutation' hypothesis above?

The following slow, but simple, MATLAB code will return a vector `acq` with the number of resistant bacteria under the hypothesis that each bacterium mutates with the probability above when challenged during the final generation of the culture. Use the mutation rate from above, but otherwise us the parameters from the mutation hypothesis code.

```
acq = zeros(cultures,1);
n = initial_inoculum*2^generations;
for i = 1:cultures
  acq(i) = binornd(n,prob_mutation);
end
```

(a) Explain in your own words how this algorithm produces the distribution under the 'acquired resistance' hypothesis, with reference to *each line* in the

code.

(a) Generate a histogram as in Problem 2. Make sure that it has labeled axes and that the title distinguishes it from the first histogram. Submit the histogram you obtain.

(b) How do the Fano factors for the 'mutation' and 'acquired resistance' simulations compare to one another and to what you would expect?

Problem 4

The algorithm in Problem 3 samples from a Binomial probability distribution (with large N, which is why it is slow).

(a) Wait, didn't Luria and Delbrück tell us this would be Poisson distributed? Why does it work at all to sample from the Binomial?

(b) How would you rewrite the 'acquired immunity' simulator to sample from a Poisson (describe what code you would replace and what you would replace it with)?

(c) Run the simulation with the code from (b), generate and submit the histogram (with a new title!), and calculate the Fano factor. How do the results compare to those from from Problem 2?

Problem 5

The mutation rates in the two models have different biological meanings, even if they are set to the same value. Let's estimate them as if we only had the experimental results.

(a) Estimate the rate of mutation λ in the 'acquired resistance' simulation using just the cultures that have zero resistant bacteria (see Luria and Delbrück's Equation 5). How does it compare to to the true value from your parameter settings?

(b) Estimate the rate of mutation a for the results from the 'mutation' simulation in Problem 2 following the first method of Luria and Delbrück (the second to last paragraph on p507). How does that compare to the true value from your parameter settings? Hint: see Equation 4 in the paper.

Problem 6

Let's see how lucky, or perhaps strategic, Luria and Delbrück were in their experimental design.

(a) In what region of parameter space would you hypothesize the two models would be difficult to quantitatively distinguish, and why?

(b) Test your hypothesis by using the 'mutation' and 'acquired resistance' simulators with altered parameter values. Present your results (using histograms and/or Fano factors).

(c) Was your hypothesis was supported or not, and why?

Problem 7 - Extra credit

Now imagine that resistant bacteria reverted to sensitive at the same rate as sensitive mutations become resistant.

(a) How would you modify the 'mutation' simulation to include this process?

(b) How do you predict the distribution and Fano factors will change?

(b) What does the distribution of resistant bacteria per culture look like if the two mutation rates are equal? What is the Fano factor? How do these compare with your predictions?

Reference

Luria SE, Delbrück M (1943) Mutations of Bacteria from Virus Sensitivity to Virus Resistance, Genetics 28, 491.