BE/APh161: Physical Biology of the Cell
Homework 5
Due Date: Wednesday, February 11, 2015

"Failure is not fatal. Failure to change might be." - John Wooden

Useful reading can be found in chaps. 5 and 13 of PBoC.

1. Counting Proteins with Partitioning Statistics.

(a) Begin by reading the paper by Rosenfeld *et al.* entitled "Gene Regulation at the Single-Cell Level" (posted on the website with the homework) and write a one paragraph commentary on the paper with special reference to how they used the binomial partitioning as a way to count repressor proteins. What is the experiment they did and what were they trying to learn?

In the rest of the problem we work out for ourselves the ideas about binomial partitioning introduced in the Rosenfeld *et al.* paper in order to consider the concentration of mRNA or proteins as a function of time in dividing cells. In particular, the point of this problem is to work out the concentration of mRNA or protein given that we start with a single parental cell that has N copies of this mRNA or protein (in the experiments of Golding *et al.* they watch the mRNA dilution effect while in the experiments of Rosenfeld *et al.* this is a fluorescently-labeled transcription factor). In the Rosenfeld experiment, at some point while the culture is growing, the production of the protein is stopped by providing a chemical in the medium and then the number of copies per cell is reduced as a result of dilution as the cells divide.

Interestingly, this problem opens the door to one of the most important themes in physics, namely, that of fluctuations. In particular, as the cells divide from one generation to the next, each daughter does not really get $N/2$ copies of the protein since the dilution effect is a stochastic process. Rather the partitioning of the N proteins into daughter cells during division follows the binomial distribution. Analyzing these fluctuations can actually lead to a quantification of the number of copies of a protein in a cell.

(b) If we think of the N copies of the protein as being divided between the two daughters with N_1 going to daughter 1 and $N-N_1$ going to daughter 2, write the probability distribution $p(N_1, N)$. Next, work out the expected fluctuations in the partitioning process after each division by noting that the fluctuations can be written as $\sqrt{<(N_1-N_2)^2>}$, where N_1 and N_2 are the number of proteins that end up in daughter cells 1 and 2, respectively. Show that $\sqrt{<(N_1-N_2)^2>} = \sqrt{N}$.

(c) Next, look at the Rosenfeld paper and explain how measuring fluorescence variations can be used to calibrate the exact number of copies of the fluorescent protein in a cell. Specifically, assume that the fluorescence intensity in each cell can be written as $I = \alpha N$, where α is an as-yet unknown calibration factor and N the number of proteins in the cell. Explain what this equation means and why you think it is justified. Derive an expression relating I_1, I_2 and I_{tot} using the result of part (b). Make a plot of $\sqrt{<(I_1-I_2)^2>}$ versus I_{tot} and explain how to get the calibration factor α from this plot.

(d) Now we are going to repeat the Rosenfeld experiment numerically in order to *fit* the calibration factor. Consider a fluorescent protein such that the calibration factor between the intensity and the number of fluorophores is 50, that is $I = 50N$. Generate intensity data by choosing $N_1 + N_2 = 10, 50, 100, 1000$ and 5000 and for each case, "partition" the proteins from the mother cell to the two daughters 100 times (i.e. as if you are looking at 100 mother cells divide for each choice of the protein copy number). Then, make a plot of the resulting $\sqrt{<(I_1-I_2)^2>}$ vs I_{tot} just as we did analytically in the previous problem. What I mean is that you need to make a plot of all of your simulation results. Then, do a fit to your "data" and see how well you recover the calibration factor that you actually put in by hand. Plot the fit on the same graph as all of the "data".

2. Fly problem.

In class we examined how using a steady-state solution to the reaction diffusion equation for Bicoid, we could understand how the exponential profile is set up.

(a) Give a brief description (a paragraph or less) of the Bicoid gradient in *Drosophila* and how it is relevant to fly development.

(b) Describe the observed concentration profile of Bicoid along the anterior-posterior axis of the fly mathematically. What is the functional form?

(c) Repeat a derivation of the reaction-diffusion equation and justify the form

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2} - \frac{c(x,t)}{\tau}. \tag{1}$$

Make sure you explain carefully where all of these terms come from.

(d) Now solve this equation in steady-state by finding the general solution subject to the boundary condition that $J(0,t) = j_0$ and $J(L,t) = 0$. Make sure you explain what these boundary conditions mean relative to the biology of the problem. Suggest approximations that can be made to simplify the result to that I presented in class.