Vignette: Poisson Statistics

1. Poisson Statistics describes distributions of discrete and small numbers. It is often useful in biology and in thinking about the set-up and interpretation of experiments. You probably will see a problem similar to these on Midterm I.

2. P = probability of event: Q = probability of non-event.

$$P+Q = 1$$

A. Remember also that probabilities multiply:

$$P(x,y) = P(x) \cdot P(y)$$

3. For small (discrete) numbers:

$$P(x) = \frac{m^x}{x!} e^{-m}$$

Where: x = number of events
m = mean number of events
P(x) = probability of x events

4. This expression is useful for all sorts of experimental considerations and problems. e.g.:

Problem 1: You are interested in determining the "burst size" of a bacteriophage growing on *E. coli*. In order to do that, you mix suspensions of phage and *E. coli* as follows:

1 ml E.coli at $5\text{x}10^{\mathbf{8}}$/ml.

20 μl phage at $2.5\text{x}10^{\mathbf{10}}$/ml

- What portion of the bacteria are infected with exactly one phage? Two phage? Three phage?

- What portion of the bacteria are uninfected? (I.e., what is P(0). Remember here that, by convention, $n^0 = 1$ and 0! = 1.), so

$$P(0) = e^{-m}$$

Continuing with the experiment, you dilute the infected culture so that when you put 100 μl into different wells of a multiple-well plate you are putting, on average (m) , one bacterial cell per well; you will incubate to lysis and then assay the number of phage in the indvidual wells to extract burst-size. In the distribution into wells:

- What fraction of the wells gets exactly one cell?
- What fraction of the wells gets one infected cell?

Problem 2. What would the phage/bacterium ratio have to be so that exactly 95% of the bacteria are infected?

Problem 3. You are engaged in a genome project, and are dancing between costs and maximum coverage. You choose a random approach. If you sequence single-pass (one sequence run on a one-strand genome equivalent) on your library, what would be the coverage (fraction nucleotides sequenced single-pass)?

- How many passes for 95% double-strand coverage (a common goal)?

Problem 4: Make-up and solve a problem using Poisson statistics.

Poisson Problems Solved

1. One ml Eco at $5x10^8$/ml = $5x10^8$ cells
   20μ page at $2.5x10^{10}$/ml = $5x10^8$ phage

   m = 1 phage per cell

 a. Cells with exactly one phage?

   $P(1) = [1^1/1!]e^{-m} = e^{-1} = 0.37$ (remember that e^{-1} = ~0.37)

 b. Cells with exactly two phage?

   $P(2) = [1^2/2!]e^{-1} = e^{-1}/2 = 0.18$

 c. Cells that are uninfected?

   $P(0) = [1^0/o!]e^{-1} = e^{-1} = 0.37$ (remember: $N^0 = 1$, 0! = 1)

 d. Distribute into wells: m=1

   P(one cell/well) = 0.37

   P(one infected cell) = 0.37x0.37 = ~0.14

2. m=? for P(infection) = 0.95

   Flip the question: $P(0) = 0.05 = e^{-m}$, m=~3

3. Single-pass means m=1

   a. P(sequence for any nt) = e^{-m} = 0.37

   b. Double vs. single-strand coverage hinges only on the number of nucleotides covered.

      For 95% coverage: as above, $P(0) = 0.05 = e^{-m}$: m = ~3

      For 95% double strand coverage: m = twice single pass = ~6