

Problem Set 2 Written Answers

Modeling The Living Cell - Spring 2021

Benjamin Fry - bfry2

Problem 1:

Completed in MATLAB script.

Problem 2:

2a.

The frequency of the residues in the 300 residue sequence is captured in the histogram in figure 1. We would expect the frequency of each residue in the histogram to be at about $0.05 = 1/20$ since each amino acid is equally likely to appear. What we actually see is some noise around this value as by nature of random number sampling, some counts will be slightly above and some will be slightly below the expected frequency.

2b.

Plotted in Figure 2 using stairs plot and residue labels.

2c.

The frequency of the residues generated is shown in figure 4. I plotted a bar chart of the data from ExPASy in figure 3 for comparison. The actual frequency of the residues that occur in the sequence is similar to that in the known frequencies, however, as expected there is noise that keeps the random sampling from equaling the expected frequencies exactly. If you increase the number of residues to see more of the limiting case by replacing the 300 random numbers with 30,000 random numbers, we can easily see that the frequencies approach the frequencies in the CDF being sampled from. Even still, the 300 residue sequence does a good job of capturing the expected number of residues with the given frequencies. For example, we see that residues like Leucine and Alanine are relatively common while Cysteine and Tryptophan are less common.

2d.

In reality, residues definitely do not occur completely randomly. Most proteins start with a methionine residue, their secondary structure segments have residues that facilitate stability, and when they interact with membranes, they have regions of hydrophobicity and hydrophilicity that correspond to transmembrane and solvent accessed regions respectively. We would expect biological residues to have similar features to those that neighbor them in the 3-dimensional structure of the protein as we know that proteins have hydrophobic cores and hydrophilic exteriors.

Problem 3:

Completed in MATLAB. Calculations logged to terminal.

Problem 4:

a) How many different arrangements?

There are

$$\binom{8}{4} = 70$$

possible arrangements for the particles.

b) Probabilities

The probability of having a given number of black particles on the left side of the membrane is given by the hypergeometric distribution as the number of 'successes' to draw from decreases with each black particle drawn if we label them such.

If we let $8 = N$ = the total number of objects to draw from (white + black particles)

$4 = m$ = the number of black particles

$4 = n$ = the sample size

and k = the number of 'successes' or black particles observed, the probability of a given k is given by

$$P(k) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}}$$

Therefore, the probabilities for each number of black particles on the left side are as follows:

$$P(4) = \frac{\binom{4}{4} \binom{4}{0}}{\binom{8}{4}} = \frac{1}{70}$$

$$P(3) = \frac{\binom{4}{3} \binom{4}{1}}{\binom{8}{4}} = \frac{16}{70}$$

$$P(2) = \frac{\binom{4}{2} \binom{4}{2}}{\binom{8}{4}} = \frac{36}{70}$$

$$P(1) = P(3)$$

$$P(0) = P(4)$$

The most likely arrangement is therefore the one with two black and two white on each side of the membrane.

c) Random Transit

Starting with 3 black and 1 white on the left of the membrane, the probability of the randomly selected particle on the left being white is $\frac{1}{4}$ and the probability of the randomly selected particle on the right being black is $\frac{1}{4}$. Since the selection events are independent, the probability of both of these events happening (which corresponds to an end state of having 4 black particles on the left) is

$$P(4 \text{ black}) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

On the other hand, the probability of selecting a black particle on the left from the same initial state is $\frac{3}{4}$ and the probability of selecting a white particle from the right is also $\frac{3}{4}$. By the same logic

$$P(2 \text{ black}) = \frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$$

which is the probability of ending with two black and two white particles on the left after the same time instant.

Therefore, ending with two black and two white on each side is the more likely scenario.

Problem 5:

i) Equation of Motion for Bond

$$\begin{aligned} U(x) &= \frac{1}{2}k(x - x_0)^2 \\ F &= \frac{-dU}{dx} = -k(x - x_0) = ma \\ a &= \frac{d^2x}{dt^2} \\ \frac{d^2x}{dt^2} &= \frac{-k(x - x_0)}{m} \end{aligned}$$

ii) Plot $x(t)$, dx/dt , and $F(t)$

Plotting done in MATLAB main script.

Since we are given $x(t) = x_0 \cos(\sqrt{\frac{k}{m}}t) + x_0$, the first time derivative of this equation gives the velocity of the bond as a function of time which is:

$$\frac{d}{dt}x(t) = -x_0\sqrt{\frac{k}{m}}\sin(\sqrt{\frac{k}{m}}t)$$

Since $F(t) = ma$ and $a = \frac{d^2}{dt^2}x(t)$, force is given by the product of mass and the second derivative of the $x(t)$ equation.

$$\frac{d^2}{dt^2}x(t) = -x_0\frac{k}{m}\cos(\sqrt{\frac{k}{m}}t)$$

multiplying this equation by m gives the equation for force as a function of time:

$$F(t) = -x_0k\cos(\sqrt{\frac{k}{m}}t)$$

These functions are plotted in my MATLAB script with the given parameter values.