Chem 163, Problem Set 1
Due 9/8/2022 in class

September 1, 2022

1 Daily ATP consumption

1. How many Calories do you eat in a day? Give an estimate and convert this to joules. (Recall that 1 Cal = 1 kcal = 1,000 cal.)
2. Approximately how many Watts does your metabolism consume?
3. Approximately 50% of the calories you eat end up being used via ATP hydrolysis, which provides about 50 kJ/mol ATP. How many moles of ATP does your body use in a day? How many kilograms is this? Express this as a fraction of your estimated body weight. Does this number seem strange? Explain. (The molecular weight of ATP is 507 g/mol.)
4. Typical ATP concentration in the body is 1-10 mM. Estimate the volume of the body and the number of moles of ATP in the body.
5. How many times per day does each molecule of ATP get used? If you suddenly stopped producing ATP, how long would you have before your body ran out?

2 Number of bacteria in your body

Estimate the volume of your body. A useful approach is to start with your mass, and then calculate volume making a reasonable assumption about your density (what happens when you go swimming?). Estimate the number of human cells in your body.

Your intestines have about 1 L of bacteria in them. Estimate the number of bacteria in your intestines, assuming the bacteria are close-packed next to each other. Compare to the number of cells in your body. Are you surprised?

3 Pair-wise diffusion

Suppose you have two particles, A and B, with diffusion coefficients D_A and D_B respectively, diffusing on a 1-D line. In a time-step Δt, what is the distribution of displacements for each particle? What is the distribution of changes in the distance *between* the particles, x_{AB}? What is the mean-square change in distance between the particles as a function of time? Can you assign an effective diffusion coefficient to the inter-particle distance?

4 Log-normal distribution

The generality of the Central Limit Theorem might make you think that *all* iterated random processes lead to a Gaussian (a.k.a. Normal) distribution. This is not the case, as we will discover in this problem! Consider an exponentially growing bacterial population in a fluctuating environment (e.g. the grad student sometimes forgets to feed the poor creatures). The bacterial population can be modeled as a *multiplicative* random walk, starting at $x = x_0$. x evolves in time according to the following rules

$$\begin{aligned} x(t+1) &= x(t) \times (1+\epsilon)\text{: probability } p \\ x(t+1) &= x(t) \times (1+\delta)\text{: probability } q, \end{aligned}$$

where $|\epsilon| \ll 1$ and $|\delta| \ll 1$, but ϵ and δ can be positive or negative.

Let $y = \ln x$. For large t, what is the probability distribution $P(y)$? What are the mean, μ, and variance, σ^2, of this distribution? Now we want the probability distribution, mean, and variance of x. This scenario comes up quite often in probability theory: we have some random variable y, with a known probability distribution, $P(y)dy$ (you just calculated it!). Then we have another variable, x, which we can express as some function $x = f(y)$. We want to know the probability distribution of x, $Q(x)dx$. If x is a monotonically varying function of y, then

$$Q(x)dx = P(y)\left|\frac{dy}{dx}\right|dx.$$

What is the probability distribution of x for large t? This distribution, $Q(x)$ is called the "log-normal" distribution. Do you understand why?

Calculate the mean of this distribution. Feel free to express your answers in terms of the mean, μ, and variance, σ^2, of $P(y)$ (this is not a totally trivial integral to evaluate; feel free to use Mathematica or to look it up). On average, which grows faster: the bacteria of the sloppy grad student who sometimes forgets to feed the bacteria, and then overcompensates by giving them extra food at later times, or the bacteria of the fastidious grad student who maintains the same average growth rate, μ, at all times?

The log-normal also describes the distribution of the intensity of radiation (e.g. light) transmitted through a medium composed of a random distribution of many small absorbing or scattering objects. For instance, it applies to sunlight propagation through clouds, cellphone signals traveling through cities, and light penetrating the forest canopy.

5 Working with Matlab

The purpose of this problem is to get you conversant with Matlab, and to demonstrate the Central Limit Theorem numerically.

If you don't have Matlab already installed on your computer, download and install it from:

http://downloads.fas.harvard.edu/download

Alternatively, Matlab is installed on the computers in the computer lab in the Science Center Basement.

If you haven't used Matlab before, download and read the "Getting Started" guide, accessible from the Help menu when you open Matlab. The most important parts of this guide to look at are: "Matrices and Arrays," "Graphics," and "Programming."

For the purpose of these exercises we ask that you put all of your commands in a script, so that you can submit a printout which, if run, would generate the results you will turn in.

Create a row vector, `P1`, of dimension (1,101), filled with zeros (look up the `zeros` command by typing `help zeros` at the command line; other commands can be looked up in this way too). Now fill some elements near the middle of this array with positive nonzero values, normalized so that `sum(P1) = 1`. This array will be the probability distribution for the displacements of each step of a random walk.

Create another row vector, `x`, that goes from -50 to 50. Calculate the expectation value $\langle x \rangle$ by summing `x(i)*P1(i)` for `i = 1:101`. There are at least three ways to do this sum:

- write a `for` loop that tallies the sum;
- look up the `.*` operation. Evaluate `sum(x.*P1)`
- perform a matrix multiplication by evaluating `x*P1'`.

Approach 3 is by far the most efficient, both to type into the computer and for the computer to evaluate. Do you understand how it works?

Make sure that $|\langle x \rangle| < 1$. If not, change your single-step distribution.

Make a plot of `P1` and print it out.

Now create a second vector, `P2`, that is the convolution of `P1` with itself (look up the `conv` command). The vector `P2` will be longer than `P1` (can you explain why?). Resize `P2` to keep only the middle values, elements `51:151`. Plot `P2` and make a printout. `P2` is the probability distribution after two steps of a random walk with single-step distribution given by `P1`.

Now write a loop to calculate the probability distribution `PN` after `N` steps. Make a printout for a few values of `N`, say 5, 10, 20, 50. Does the behavior make sense? Calculate the mean and variance of the distribution after each of these times. How do they both change with `N`?

6 Simulating a random walk

In this problem you will re-create the distributions you calculated in the first problem, by two other means. Let the variable x undergo a random walk according to the following rules:

$$\begin{aligned} x(t+1) &= x(t) - 1: \text{ probability } p \\ x(t+1) &= x(t) + 1: \text{ probability } q, \end{aligned}$$

with $p + q = 1$.

1. For the case $p = 0.6$ and $q = 0.4$, use the code from problem 1 to simulate the probability distribution of where you will be after 100 steps, and make a plot of this distribution.

2. Now you will re-create this distribution by direct simulation. Create a variable x (initialized to 0) and write a `for` loop to evolve x according to the rules above (Hint: look up the `rand` command. How would you generate an event with probability p, given a random variable uniformly distributed between 0 and 1?). Have this loop run for 100 steps. Make a plot showing an overlay of 5 trajectories, all starting at $x = 0$ (Hint: look up the `hold all` command)

3. Now write a program to repeat this 100-step simulation 10,000 times, storing each of the final values in the variable `xfinal(k)`. You don't need to keep the intermediate steps of each trajectory.

4. On the same graph as in part 1, plot a probability distribution, `P`, of `xfinal` (use the `hold` command to keep one plot while you overlay another on top). You can use the `hist` command to generate the probability distribution. Choose a bin size Δx and normalization so that the probability distribution is normalized (i.e. $(\sum \texttt{P})\, \Delta x = 1$)! Report the mean and variance of `xfinal` (look up the commands `mean` and `var`).

5. Finally, re-create the distribution by application of the central limit theorem. Analytically calculate what you expect for the mean and variance, and write the formula for the expected probability distribution, $P(x)$. Plot this probability distribution on the same plot as the other two probability distributions. Does the result make sense?

7 Diffusion to capture

In this problem you will simulate 1D diffusion toward an absorbing boundary. This is a simple model of nutrient transport and uptake, e.g. of sugar diffusing toward a bacterium. Consider a particle with a diffusion coefficient of $D = 100 \mu\text{m}^2/\text{s}$.

1. In a time step $\delta t = 1$ ms, what is the distribution of 1D displacements of this particle?

2. Consider a 1D channel, 10 μm long, with a reflecting boundary at $x = 0$, and an absorbing boundary at $x = 10\mu$m. Suppose the particle above is released at $x = 1\mu$m. Simulate its trajectory in time steps of 1 ms, using the step size distribution you calculated above (Hint: the `randn` command will give you a Gaussian distributed random variable with mean 0 and variance 1.). If a step takes the particle to a negative coordinate, implement the reflection by taking the absolute value of the position. End the simulation when the particle crosses $x = 10\mu$m for the first time. Make a plot of the trajectory.

3. Repeat the above simulation 1,000 times. For each iteration, record (a) the time elapsed until the particle is absorbed, and (b) the histogram of particle positions over the entire trajectory, dividing the interval [0, 10] into bins of 0.5 μm. What is the mean time elapsed before the particle is absorbed? This is called the 'mean first-passage time'. Make a plot of the mean distribution of particle locations. Does this plot make sense? If particles were released at $x = 1\mu$m a constant rate, this distribution would be proportional to the steady-state concentration profile. Is this concentration profile a steady-state solution to the diffusion equation? Why or why not? Does this concentration profile obey the no-flux boundary condition ($\partial c/\partial x = 0$ at $x = 0$)?

4. Repeat the above simulation 1,000 times, but with absorbing boundary conditions at $x = 0$ and $x = 10\mu$m. Plot the steady-state concentration profile. What is the ratio of absorption on the right (at $x = 10\mu$m) vs. absorption at the left (at $x = 0$)? Now do this again for particles released at $x = 2\mu$m, $x = 3\mu$m, ..., $x = 9\mu$m. Can you use the diffusion equation to derive a formula for the branching ratio as a function of the position where the particles are released? Hint: The rate of absorption at a boundary is proportional to the flux into the boundary, which is proportional to the gradient in concentration at the boundary.