Assignment “Introduction to Statistical Thermodynamics”
of the course Bioinformatics Data Analysis and Tools
Anton Feenstra (feenstra@few.vu.nl)

Problem 1

In the first lecture we learned that the total number of states $\Omega$ of two isolated systems A and B is equal to the product of the number of states of each of them: $\Omega = \Omega_A \Omega_B$. We considered an example of a box separated into 3 parts with one particle inside. Two such boxes have 9 different states.

(1) How many states have this system of two boxes, if we allow particles (the particles are identical) to penetrate through the boxes’ boundaries? What conclusions can you derive from this example? What can you say in general about the total entropy of the system in these two cases (of two isolated and two non-isolated boxes)?

(2) Consider now the next example given in the lecture: A square lattice (Fig. 1) with two types of molecules A and B.

![Fig. 1](image)

Compare numerically the entropy per molecule computed using the Stirling formula (see the lecture) and the exact formula. You can assume for simplicity $N_A = N_B$. Use your favourite computing software. Plot the entropy per molecule $S/N$ as a function of the number of molecules in the system. What do you observe upon increasing of the number of molecules? What conclusion can you derive?
(1) Consider now two boxes as one system (Fig. 2). What is the total entropy of this system (per molecule) if the boxes are isolated? What is it if the boxes are non-isolated? (These two questions are analogous to the questions you already answered in part (1) and (2) of Problem 1.) Answer these questions first using the Stirling formula, and after that, using the exact formula for the entropy. Plot the total entropy per molecule for the system of two isolated and two non-isolated boxes as a function of the number of molecules in the system (you can assume again \( N_A = N_B = \frac{1}{2}N \) where \( N \) is the total number of molecules in one box; you can also assume that the boxes are identical). What conclusions can you derive?

**What is the most important conclusion?**

*Technical hint:* You should plot the entropy, and not the number of states. If you are really curious, try to plot the number of states, and see what happens.
Problem 2

In Lecture 2 we introduced the Boltzmann probability distribution. We considered the example of a system with $M$ discrete states [for example, a protein molecule with several possible conformations]. The free energy of the system can be written as follows:

$$F = \sum_{i=1}^{M} p_i E_i + k_B T \sum_{i=1}^{M} p_i \ln p_i \quad (2.1)$$

where $p_i$ is the probability of the system to be in state $i$. Using the material of Lecture 2, derive the Boltzmann probability distribution by minimizing this free energy $F$. Substitute the Boltzmann probability distribution into Equation (2.1), and derive the answer obtained in Lecture 2. What qualitative conclusions can you make about the free energy of this $M$-state system?
Problem 3

In Lecture 2 we introduced the notion of Probability distribution. We considered two examples of probability distributions – the Gaussian probability distribution, and the uniform probability distribution. Consider now the example very similar to the one considered in the Lecture 2: a spring (Fig. 3) with equilibrium length \( x_0 \) connected to a thermal bath (thermal reservoir) with temperature \( T \).

![Spring diagram](image)

**Fig. 3**

The energy of the spring is \( E_a = \frac{1}{2} \gamma (x - x_0)^2 \), where \( \gamma \) is the spring rigidity.

**Compute the fluctuation (variance) of the spring position** \( < (x - x_0)^2 > \)

Hint: (1) Write down the probability distribution of the spring fluctuations. (2) Normalize this probability distribution. (3) Compute the fluctuations using the normalized probability distribution.

You will need the following integrals:

\[
\int_{-\infty}^{\infty} e^{-Ay^2} \, dy = \sqrt{\frac{\pi}{A}}
\]

\[
\int_{-\infty}^{\infty} y^2 e^{-Ay^2} \, dy = \frac{1}{2A} \sqrt{\frac{\pi}{A}}
\]

**Note:** This problem is of great practical importance because in many bio-molecular systems the fluctuations (for example, small fluctuations of protein chains) are described by the Gaussian \( E \sim x^2 \) (which means quadratic) law.
**Problem 4**  (Unfolding of single RNA molecule).

The aim of this exercise is to demonstrate that the very simple, two-state model (Lecture 2) can correctly describe the physics of unfolding of single RNA molecule.


(2) Using the material of Lecture 2, try to reproduce the experimental plot [fraction of unfolded RNA as a function of applied force]. You have to figure out what physical quantity corresponds to the fraction of unfolded RNA [this was briefly discussed in the Lecture].

(3) What can you say about the kinetics of the unfolding process?
Problem 5 (Diffusion)

The (one dimensional) diffusion equation has the following form:
\[ \frac{dc}{dt} = D \frac{d^2c}{dx^2} \]

where \( c(x,t) \) is the concentration of molecules in the point \( x \) at time \( t \).

(1) Verify that the solution of this equation is:
\[ c(x, t) = \frac{A}{\sqrt{t}} e^{-x^2/(4Dt)} \]

(2) The Probability distribution \( P(x, t) \) for the diffusion process is described by the same law:
\[ P(x, t) = \frac{B}{\sqrt{t}} e^{-x^2/(4Dt)}, \text{ where } B \text{ is the normalization constant.} \]

- What is the fluctuation of the position of a diffusing particle \( \langle x(t)^2 \rangle \)?
- From this result you can find the normalization constant \( B \). Alternatively, find \( B \) from the normalization condition on \( P(x, t) \).
- Explain what is the physical meaning of \( P(x, t) \).