Université de Genève Autumn term 2022

# Statistical Methods in Physics (14P058)

Prof. Federico Sánchez (federico.sancheznieto@unige.ch)
Dr. Hepeng Yao (hepeng.yao@unige.ch)
Dr. Knut Zoch (knut.zoch@unige.ch)

# Exercise VI – Maximum likelihood principle

13 December 2022, 15:15, room: SCI–222

In this exercise, we will use the maximum likelihood principle to perform parameter estimation. We will first generate the energy for a large number of atoms which follows the Boltzmann distribution. Then, we assume this to be the experimental data from an actual measurement and estimate the temperature of the system based on this data.

Similar to the previous exercise, we consider a box system containing a large number of $^{85}$Rb atoms with mass $m = 85\, m_0$ at room temperature $T$, where $m_0$ is the atomic mass unit. The kinetic energy of the system $E$ follows the probability density function

$$f(E) = 2\left(\frac{1}{k_B T}\right)^{3/2} \sqrt{\frac{E_k}{\pi}} e^{-\frac{E_k}{k_B T}} \tag{1}$$

As a reminder: the expectation value of this distribution is $1.5\, k_B T$.

1. Write a Monte Carlo process which generates the energy distribution for $N_{par} = 400$ particles at temperature $T = 300\,\mathrm{K}$. In the following, we will assume this is the data from a single measurement in an actual thermal gas experiment.

2. Write a function which computes the likelihood $L(E_1, E_2, ..., E_{N_{par}} \,|\, T)$ of your data at a given temperature $T$. As a reminder: the definition of the likelihood follows

$$L(\vec{X}|\theta) = \prod_i f(x_i|\theta). \tag{2}$$

3. In actual experiment, we may know in advance a wide range of temperatures where the system's temperature locates. In our case, assume that we know the system has a temperature in the range $[250\,\mathrm{K}, 350\,\mathrm{K}]$. Taking a list of $T$ in the range $[250\,\mathrm{K}, 350\,\mathrm{K}]$ with a distance of $1\,\mathrm{K}$, plot $-\ln L$ as a function of $T$ and locate the minimum of $-\ln L$. This will be your estimation of $T_e$.

4. Now, we would like to perform two different methods to estimate the uncertainty of the obtained $T_e$. We start with repeating the same estimation process and study the distribution of $T_e$. To do so, you should repeat the same process of question 1-3 with $N_{exp} = 200$. (In an actual experiment, this means you perform the same measurement 200 times on the same system.) Then, check that the distribution of $T_e$ follows a Gaussian. Find the standard variance $\sigma_{T_e}$ of this Gaussian.

5. Another estimation of the uncertainty focuses only on a single set of measurement. Taking your data from question 3, find the two values $T_1$ and $T_2$ which satisfy: (a) the difference between its corresponding $-\ln L$ and the one of $T_e$ is less than 0.5, (b) the value $T_2 - T_1$

is maximized. Then, the obtained value $(T_2 - T_1)/2$ is the estimation of the uncertainty. Compare it with the result from question 4.

6. Fix $N_{exp} = 200$, estimate the Fisher information $I(\hat{T})$ with particle numbers $N_{par} = 100, 200, 400$. Comment on your results.

7. Fix $N_{par} = 200$, estimate the Fisher information $I(\hat{T})$ with experiment numbers $N_{exp} = 100, 200, 400$. Comment on your results.