Statistical Analysis of Data
Problem sheet #5

Please send by the announced due date to Glen Cowan, Physics Dept., Royal Holloway, University of London, Egham, Surrey, TW20 0EX, or e-mail to `g.cowan@rhbnc.ac.uk`.

Exercise 5.1: Consider a single observation of a Poisson distributed variable n. What is the maximum-likelihood estimator of the mean ν? Show that the estimator is unbiased and find its variance. Show that the variance of $\hat{\nu}$ is equal to the minimum variance bound.

Exercise 5.2: Early evidence supporting the Standard Model of particle physics was provided by the observation of a difference in the cross sections $\sigma_{\rm R}$ and $\sigma_{\rm L}$ for inelastic scattering of right (R) or left (L) hand polarized electrons on a deuterium target. For a given integrated luminosity L (proportional to the electron beam intensity and time of data taking), the numbers of scattering events of each type are Poisson variables, $n_{\rm R}$ and $n_{\rm L}$, with means $\nu_{\rm R}$ and $\nu_{\rm L}$. The means are related to the cross sections by $\nu_{\rm R} = \sigma_{\rm R} L$ and $\nu_{\rm L} = \sigma_{\rm L} L$, and the experiment is set up such that the luminosity L is equal for both cases. Using the result from Exercise 5.1, construct an estimator $\hat{\alpha}$ for the polarization asymmetry,

$$\alpha = \frac{\sigma_{\rm R} - \sigma_{\rm L}}{\sigma_{\rm R} + \sigma_{\rm L}} \,. \tag{1}$$

Using error propagation, find the standard deviation $\sigma_{\hat{\alpha}}$ as a function of α and $\nu_{\rm tot} = \nu_{\rm R} + \nu_{\rm L}$. The asymmetry was expected to be at the level of 10^{-4}. How many scattering events must be observed so that $\sigma_{\hat{\alpha}}$ is a factor of ten smaller than this? (The number of is so large that the events could not be recorded individually, but rather the output current of the detector was measured. See C.Y. Prescott et al., Parity non-conservation in inelastic electron scattering, Phys. Lett. B77 (1978) 347.)

Exercise 5.3: One of the earliest determinations of Avogadro's number was based on Brownian motion. The experimental set-up shown in Fig. 1 was used by Jean Perrin [1] to observe particles of mastic (a substance used in varnish) suspended in water.

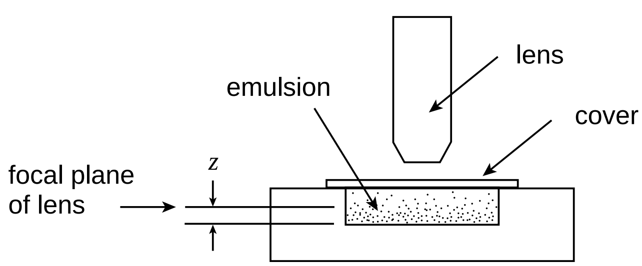

Figure 1: Experimental set-up of Jean Perrin for observing the number of particles suspended in water as a function of height.

The particles were spheres of radius $r = 0.52\ \mu$m and had a density of 1.063 g/cm^3, i.e. 0.063 g/cm^3 greater than that of water. By viewing the particles through the microscope, only those in a layer approximately 1 μm thick were in focus; particles outside this layer were not visible.

[1] Jean Perrin, Mouvement brownien et réalité moléculaire, *Ann. Chimie et Physique, 8e série*, **18** (1909) 1-114; *Les Atomes*, Flammarion, Paris, 1991 (first edition, 1913); *Brownian Movement and Molecular Reality*, in Mary-Jo Nye, ed., *The Question of the Atom*, Tomash, Los Angeles, 1984.

By adjusting the microscope lens, the focal plane could be moved vertically. Photographs were taken at 4 different heights z, (the lowest height is arbitrarily assigned a value $z = 0$) and the number of particles $n(z)$ counted. The data are shown in Table 1.

Table 1: Perrin's data on the number of mastic particles observed at different heights z in an emulsion.

height z (μm)	number of particles n
0	1880
6	940
12	530
18	305

The gravitational potential energy of a spherical particle of mastic in water is given by

$$E = \frac{4}{3}\pi r^3 \,\Delta\rho\, g z, \tag{2}$$

where $\Delta\rho = \rho_{\rm mastic} - \rho_{\rm water} = 0.063$ g/cm^3 is the difference in densities and $g = 980$ cm/s^2 is the acceleration of gravity. Statistical mechanics predicts that the probability for a particle to be in a state of energy E is proportional to

$$P(E) \propto e^{-E/kT}, \tag{3}$$

where k is Boltzmann's constant and T the absolute temperature. The particles should therefore be distributed in height according to an exponential law, where the number n observed at z can be treated as a Poisson variable with a mean $\nu(z)$. By combining (2) and (3), this is found to be

$$\nu(z) = \nu_0 \exp\left(-\frac{4\pi r^3 \,\Delta\rho\, g z}{3kT}\right), \tag{4}$$

where ν_0 is the expected number of particles at $z = 0$.

(a) Write a computer program to determine the parameters k and ν_0 with the method of maximum likelihood. Use the data given in Table 1 to construct the log-likelihood function based on Poisson probabilities (cf. SDA Section 6.10),

$$\log L(\nu_0, k) = \sum_{i=1}^{N} (n_i \log \nu_i - \nu_i), \tag{5}$$

where $N = 4$ is the number of measurements. For the temperature use $T = 293$ K.

(b) From the value you obtain for k, determine Avogadro's number using the relation

$$N_{\rm A} = R/k, \tag{6}$$

where R is the gas constant. The value used by Perrin was $R = 8.32 \times 10^7$ erg/mol K.

(c) Instead of maximizing the log-likelihood function (5), estimate ν_0 and k by minimizing

$$\chi^2_{\rm P}(\nu_0, k) = 2 \sum_{i=1}^{N} \left(n_i \log \frac{n_i}{\nu_i} + \nu_i - n_i \right) , \tag{7}$$

where $\nu_i = \nu(z_i)$ depends on ν_0 and k through equation (4). Use the value of $\chi^2_{\rm P}$ to evaluate the goodness-of-fit (cf. SDA Section 6.11). Comment on possible systematic errors in Perrin's determination of $N_{\rm A}$.

Exercise 5.4 (optional): A random variable x follows a p.d.f. $f(x;\theta)$ where θ is an unknown parameter. Consider a sample $\mathbf{x} = (x_1, \ldots, x_n)$ used to construct an estimator $\hat{\theta}(\mathbf{x})$ for θ (not necessarily the ML estimator). Prove the Rao-Cramér-Frechet (RCF) inequality,

$$V[\hat{\theta}] \geq \frac{\left(1 + \frac{\partial b}{\partial \theta}\right)^2}{-E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right]} , \tag{8}$$

where $b = E[\hat{\theta}] - \theta$ is the bias of the estimator. This will require several steps:

(a) First, prove the Cauchy–Schwarz inequality, which states that for any two random variables u and v,

$$V[u]V[v] \geq (\mathrm{cov}[u,v])^2, \tag{9}$$

where $V[u]$ and $V[v]$ are the variances and $\mathrm{cov}[u,v]$ the covariance. Use that fact that the variance of $\alpha u + v$ must be greater than or equal to zero for any value of α. Then consider the special case $\alpha = (V[v]/V[u])^{1/2}$.

(b) Use the Cauchy–Schwarz inequality with

$$\begin{aligned} u &= \hat{\theta}, \\ v &= \frac{\partial}{\partial \theta} \log L, \end{aligned} \tag{10}$$

where $L = f_{\rm joint}(\mathbf{x};\theta)$ is the likelihood function, which is also the joint p.d.f. for $\mathbf{x}$. Write (9) so as to express a lower bound on $V[\hat{\theta}]$. Note that here we are treating the likelihood function as a function of $\mathbf{x}$, i.e. it is regarded as a random variable.

(c) Assume that differentiation with respect to θ can be brought outside the integral to show that

$$E\left[\frac{\partial}{\partial \theta} \log L\right] = \int \ldots \int f_{\rm joint}(\mathbf{x};\theta) \frac{\partial}{\partial \theta} \log f_{\rm joint}(\mathbf{x};\theta)\, dx_1 \ldots dx_n = 0. \tag{11}$$

The form of the RCF inequality that we will derive depends on this assumption, which is true in most cases of interest. (It is fulfilled as long as the limits of integration do not depend on θ.) Use (11) with (9) and (10) to show that

$$V[\hat{\theta}] \geq \frac{\left(E\left[\hat{\theta}\frac{\partial \log L}{\partial \theta}\right]\right)^2}{E\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right]}. \tag{12}$$

(d) Show that the numerator of (12) can be expressed as

$$E\left[\hat{\theta}\frac{\partial \log L}{\partial \theta}\right] = 1 + \frac{\partial b}{\partial \theta}, \tag{13}$$

and that in a similar way the denominator is

$$E\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right] = -E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right]. \tag{14}$$

Again assume that the order of differentiation with respect to θ and integration over $\mathbf{x}$ can be reversed. Prove (8) by putting together the ingredients from (c) and (d).