Statistical Analysis of Data
1998/99 Exam questions

For this question, answer all of (a)–(d) and either (e) or (f).

Consider N independent Poisson variables $n_1, \ldots, n_N$ with mean values $\nu_1, \ldots, \nu_N$. Suppose the mean values are related to a controlled variable x by

$$\nu(x) = \theta x,$$

where θ is an unknown parameter. The values of ν_i for $i = 1, \ldots, N$ are thus given by $\nu_i = \nu(x_i) = \theta x_i$, where the values $x_1, \ldots, x_N$ are known. We are given a data sample consisting of $n_1, \ldots, n_N$. (An example of this situation is where we measure the numbers of neutrino scattering events produced at N different values of the neutrino energy. The variable x is proportional to the energy and θx is the predicted number of events.)

Answer all of (a)–(d):

(a) (4 points) Show that the log-likelihood function is

$$\log L(\theta) = \sum_{i=1}^{N} (n_i \log \nu_i - \nu_i)$$

and hence that the maximum-likelihood estimator for θ is given by

$$\hat{\theta} = \frac{\sum_{i=1}^{N} n_i}{\sum_{i=1}^{N} x_i}.$$

(b) (3 points) Show that $\hat{\theta}$ is an unbiased estimator for θ.

(c) (4 points) Find the variance of $\hat{\theta}$ and show that it is equal to the RCF bound.

(d) (4 points) Use the method of least squares to find an estimator for θ. For the denominators in the χ^2 use the Poisson variances $\sigma_i^2 = \nu_i$.

Answer one of (e) or (f) (5 points):

(e) Explain how to evaluate the goodness-of-fit for both the maximum-likelihood and least-squares cases. Comment on what must be done if some of the n_i are small or zero.

(f) Suppose n_i represents the number of events found with neutrino beam energy E_i and integrated luminosity L_i. The expected number of events can be written as

$$\nu_i = \theta E_i L_i\,.$$

Suppose data can be taken at two different energies, E_1 and E_2, and the total luminosity $L = L_1 + L_2$ is fixed by budget constraints. How should the luminosities L_1 and L_2 be shared between the two energies so as to minimize the error in the estimate of θ?