Exercises on Probability Theory and Bayesian Statistics

Luc Demortier[1]

Problem 1: Eliminating nuisance parameters by conditioning.

In the frequentist paradigm, handling nuisance parameters can be a thorny problem. A method that sometimes works is based on the idea of conditioning. To illustrate this approach, suppose we measure an event count N that is Poisson distributed with mean $\mu\nu$, where μ is the parameter of interest and ν a nuisance parameter. Assume that ν is constrained by the auxiliary measurement of a Poisson variate K with mean $\tau\nu$, where τ is a known constant:

$$N \sim \text{Poisson}(\mu\nu), \tag{1}$$
$$K \sim \text{Poisson}(\tau\nu). \tag{2}$$

In high energy physics one could think of μ as the production cross section for some process of interest and ν as a product of efficiencies, acceptances, and integrated luminosity. One can argue that the *sum* $M \equiv N + K$ provides no information about the *ratio* μ/τ of the above two Poisson means, or about μ itself. It is therefore interesting to seek inferences that condition on M.

1. Compute the conditional distribution of N given M.

2. Next, assume that the expectation value of N is the sum of μ and ν instead of their product, so we have:

$$N \sim \text{Poisson}(\mu + \nu), \tag{3}$$
$$K \sim \text{Poisson}(\tau\nu). \tag{4}$$

   What is the conditional distribution of N given M here?

Problem 2: Eliminating nuisance parameters by Bayesian marginalization.

Here we take the first of the above problems ($N \sim \text{Poisson}(\mu\nu)$) and eliminate the nuisance parameter ν by Bayesian marginalization. One can proceed as follows:

1. Consider the auxiliary measurement of ν, via $K \sim \text{Poisson}(\tau\nu)$.

2. Compute Jeffreys' prior for ν for that auxiliary measurement.

3. Compute the posterior for ν for that auxiliary measurement.

4. Use the auxiliary posterior for ν as a prior for ν in the measurement of μ.

[1]luc.demortier@mail.rockefeller.edu

5. Now we still need a prior for μ. In fact, since the problem involves two parameters, μ and ν, what we really need is a conditional prior for μ given ν. Reference analysis provides a method for calculating this conditional prior, and the result is identical to Jeffreys' prior calculated for a fixed value of ν.

6. Write out the joint posterior for μ and ν, without trying to normalize it.

7. Integrate out the ν dependence.

8. Compare the resulting posterior for μ with the μ dependence of the conditional pdf for N obtained in Problem 1.

Problem 3: Sampling to a foregone conclusion.

This is an exercise to illustrate the Law of the Iterated Logarithm. Write a little Monte Carlo program to do the following:

- Generate random numbers X_i from a Gaussian distribution with zero mean and unit standard deviation.

- As you generate them, compute a "running significance" Z_n:

$$Z_n \equiv \frac{1}{\sqrt{n}} \left| \sum_{i=1}^{n} X_i \right|, \tag{5}$$

  where n is the number of Gaussian variates generated so far.

- Make a plot of Z_n versus n. At what value of n does the first crossing of $Z_n = 2$ occur? What about $Z_n = 3$? Compare with the curve $\sqrt{2 \ln \ln n}$. What happens if, instead of checking the significance after each new data point, we only check it after every 100 new data points?

Problem 4: Bayesian intervals for an exponential lifetime.

Consider an exponential decay time t with probability density $f(t \mid \tau) = e^{-t/\tau}/\tau$. Derive Jeffreys' prior for this problem and compute the corresponding posterior. Construct equal-tailed intervals from this posterior and compute their frequentist coverage. Repeat with a flat prior.