STAT 489-01: Bayesian Methods of Data Analysis

Problem Set 9

Assigned 2017 April 20
Due 2017 April 27

Show your work on all problems! Be sure to give credit to any collaborators, or outside sources used in solving the problems. Note that if using an outside source to do a calculation, you should use it as a reference for the method, and actually carry out the calculation yourself; it's not sufficient to quote the results of a calculation contained in an outside source.

1 Maximum Entropy

(a) Consider a discrete distribution $p(y|n, I)$ where $y = 0, 1, \ldots, n$ and $E(y|n, I) = \mu$ where $0 \leq \mu \leq n$. Suppose the measure is $m(y|n, I) = \binom{n}{y}$ and derive the pmf which maximizes the entropy $S = -\sum_{y=0}^{n} p(y|n, I) \ln \frac{p(y|n,I)}{m(y|n,I)}$ subject to the constraints of normalization and $E(y|n, I) = \mu$. Show that by appropriate choice of the Lagrange mulitpliers, it can be written as the binomial distribution

$$p(y|n, I) = \binom{n}{y} \left(\frac{\mu}{n}\right)^y \left(1 - \frac{\mu}{n}\right)^{n-y} \tag{1.1}$$

(b) Consider a continuous distribution $p(x|\mu, \sigma, I)$ where $-\infty < x < \infty$. Show that, assuming a uniform measure $m(x|\mu, \sigma, I) = 1$, the distribution which maximizes the entropy $S = -\int_{-\infty}^{\infty} p(x|\mu, \sigma, I) \ln p(x|\mu, \sigma, I)\, dx$ subject to the constraints

$$\int_{-\infty}^{\infty} p(x|\mu, \sigma, I)\, dx = 1\,; \quad \int_{-\infty}^{\infty} x\, p(x|\mu, \sigma, I)\, dx = \mu\,; \quad \int_{-\infty}^{\infty} (x-\mu)^2\, p(x|\mu, \sigma, I)\, dx = \sigma^2 \tag{1.2}$$

is the normal distribution. (For the purposes of computing the derivatives, it's convenient to approximate the integrals as sums and compute $p_x = p(x|\mu, \sigma, I)\, dx$ for each value of x.)

2 Hamiltonian Monte Carlo

Repeat problem 2 of problem set 8, using Hamiltonian Monte Carlo (coded in Stan) rather than the Gibbs Sampler (coded in JAGS).

3 Project Proposals (one per team)

Submit a proposal for your computational project, in conference abstract format, i.e., with a title, authors, and a paragraph explaining the proposed project. Describe how you'd simulate or acquire the data, what analysis you'd perform, and what properties you'd test.