STAT 489-01: Bayesian Methods of Data Analysis

Problem Set 1

Assigned 2017 January 24
Due 2017 January 31

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 Assigning and Calculating Probabilities

Suppose I have a bag containing three fair dice with four, six, and eight sides, respectively.

(a) If I tell you that I've drawn a die at random from the bag, what probability should you assign to the proposition that it was the six-sided die? (This is not supposed to be a trick question, but you could make it one if you like.)

(b) Suppose I draw a die at random and roll it six times. What is the probability that the sequence of rolls will be 3, 1, 5, 1, 2, 4, if the die I chose was: **(i)** the four-sided one? **(ii)** the six-sided one? **(iij)** the eight-sided one?

(c) What probablity should you assign to the sequence of rolls above occurring, if you don't know which die I chose?

(d) Given that I chose a die at random and rolled the sequence above, what is the probability that it was: **(i)** the four-sided one? **(ii)** the six-sided one? **(iij)** the eight-sided one?

2 The Monty Hall Problem

This is a classic problem in logic and probability, usually described in terms of the game show *Let's Make a Deal*. You are given a choice of three doors; behind one there is a valuable prize (a new car), and behind the other two are booby prizes (goats). The car and goats were randomly placed before the game, and there is nothing special about any of the doors. You choose door #1. Before you open it, the host, Monty Hall, opens one of the other two doors, reveals that there is a goat behind it. You are then given the opportunity to switch from door #1 to the other unopened door. Monty was obligated to open one of the doors, knows which door has the car behind it, and deliberately chose a door with a goat.

(a) What is the probability that you will win the car if you switch? What is the probability that you will win if you don't switch?

(b) Suppose now that Monty did not know where the car was, and chose one of the two unopened doors at random. Given that that door happened to contain a goat, what is the probability that you will win if you switch? If you don't switch?

3 Marginalized Sampling Distribution

Consider the scenario described in class, where $\mathbf{y} = y_1, \ldots, y_n$ is the outcome of a series of Bernoulli trials with unknown probability parameter θ which has a uniform prior probability distribution. Suppose there are a total of $y_{\text{tot}} = \sum_{i=1}^{n} y_i$ successes in the n trials.

(a) Explicitly calculate the marginalized sampling distribution $p(\mathbf{y}|I) = \int_0^1 p(\mathbf{y}|\theta, I)\, p(\theta|I)\, d\theta$ and verify that $p(\theta|\mathbf{y}, I) = \frac{p(\mathbf{y}|\theta,I)p(\theta|I)}{p(\mathbf{y}|I)}$

(b) Explicitly calculate the marginalized sampling distribution $p(y_{\text{tot}}|n, I) = \int_0^1 p(y_{\text{tot}}|n, \theta, I)\, p(\theta|I)\, d\theta$ for the corresponding binomial experiment and verify that $p(\theta|y_{\text{tot}}, n, I) = \frac{p(y_{\text{tot}}|\theta,n,I)p(\theta|I)}{p(y_{\text{tot}}|I)}$.

You may find it useful to use the identity (valid for positive α and β)

$$\int_0^1 \theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{3.1}$$

where $\Gamma(k) = (k-1)!$ when k is a positive integer.

4 Numerical Parameter Estimation

Recall the Bernoulli trial problem considered in class, with the parameter θ representing the probability of success in each trial. Use R to produce plots of the normalized prior and posterior distributions in the following cases:

(a) Start with a uniform prior probability distribution $p(\theta|I)$ and suppose we have seen a total of **(i)** 3 successes in 4 trials; **(ii)** 12 successes in 16 trials; **(iij)** 45 successes in 60 trials.

(b) Start with a prior probability distribution $p(\theta|I) = 1 - \cos(2\pi\theta)$ and suppose we have seen a total of **(i)** 3 successes in 4 trials; **(ii)** 12 successes in 16 trials; **(iij)** 45 successes in 60 trials.

(c) Start with a prior probability distribution $p(\theta|I) = 2$ for $\theta < \frac{1}{2}$, zero elsewhere, and suppose we have seen a total of **(i)** 3 successes in 4 trials; **(ii)** 12 successes in 16 trials; **(iij)** 45 successes in 60 trials. Comment on the appropriateness of this prior distribution in light of the observed data.

5 Conjugate Prior Distribution

Consider the Bernoulli trial experiment where the prior distribution is the Beta distribution

$$p(\theta|I_{\alpha,\beta}) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \qquad 0 < \theta < 1 \tag{5.1}$$

for some positive (not necessarily integer) α and β.

(a) Construct the normalized posterior $p(\theta|\mathbf{y}, I_{\alpha,\beta})$ where $\mathbf{y}$ includes, as before, y_{tot} successes in n trials.

(b) Calculate the posterior expectation $E(\theta|\mathbf{y}, I_{\alpha,\beta}) = \int_0^1 \theta\, p(\theta|\mathbf{y}, I_{\alpha,\beta})\, d\theta$. (Hint: use the fact that the beta distribution has prior expectation $E(\theta|I_{\alpha,\beta}) = \frac{\alpha}{\alpha+\beta}$.)