BST 701: Bayesian Modeling in Biostatistics
Breheny

Assignment 1

Due: Tuesday, January 29

1. A common objection to Bayesian analysis is the belief that the frequentist approach is more objective (objectivity, of course, being an important goal in science). Read the article "Statistical Analysis and the Illusion of Objectivity", by Berger and Berry (1998, `Berger1988.pdf`), which argues that the supposed objectivity of frequentist statistics is an illusion. Provide at least one of Berger and Berry's arguments.

2. Write an `R` function, `binom.bayes`, that implements a beta-binomial analysis for simple binomially distributed data. The function should be able to run with, say, `binom.bayes(20, 40)`, where a reference prior will be used (whether the function uses a Jeffreys or uniform prior by default is up to you). The function should also be able to be called with options `a` and `b` for the parameters of the beta prior on θ. In addition, the function should allow for a `plot=TRUE` option, which will produce a plot of the posterior distribution, and an option called `conf` which describes the desired probability for the posterior intervals. The function should return a list with the following components:

   - `mean.post`: The posterior mean
   - `mode.post`: The posterior mode
   - `var.post`: The posterior variance
   - `ci.central`: The central `100*conf`% posterior interval
   - `ci.hpd`: The `100*conf`% highest posterior density interval

   The function can return additional components if you would like, but must contain the above components. As a check, I highly recommend running through the code in `1-15.R` to ensure that your results agree with those reported in lecture. (Hint: To construct the HPD intervals, consider using `uniroot`, `R`'s built-in function to solve for the roots of an equation).

3. Your friend Andy claims to have ESP (extra-sensory perception). To test this claim, you propose the following experiment: You select one of four cards and Andy tries to identify it. Let θ denote the probability that Andy is correct in identifying the card. You believe that Andy has no ESP ability ($\theta = 1/4$), but there is a small chance that θ is either larger or smaller than 1/4. After some thought, you place the following prior distribution on θ:

θ	0	.125	.250	.375	.500	.625	.750	.875	1
$p(\theta)$	.001	.001	.950	.008	.008	.008	.008	.008	.008

   Suppose that the experiment is repeated ten times and Andy is correct six times and incorrect four times. Find the posterior probabilities of these values of p (you may wish to look at Section 3.4 of our textbook). In particular, what is your posterior probability that Andy has no ability?

4. For a Poisson likelihood with conjugate prior,

   (a) Show that the posterior mean can be written as a weighted average of the prior and sample means.

   (b) Show that, regardless of the choices of α and β for the prior, the posterior of $\theta|\mathbf{y}$ converges in distribution to θ_0, where θ_0 is the true Poisson rate.

5. For the Poisson likelihood, show that $p(\theta) = \theta^{-1/2}$ is the Jeffreys prior.

6. In a 2006 study published in *The New England Journal of Medicine*, 78 pairs of patients with Parkinson's disease were randomly assigned to receive treatment (which consisted of deep-brain stimulation of a region of the brain affected by the disease) or control (which consisted of taking a prescription drug). The researchers found that in 50 of 78 pairs, the patients who received deep-brain stimulation had improved more than their partner in the control group. The parameter of interest is θ, the probability of doing better on treatment than control.

   (a) Consider the description of the study (but not its results) and decide on a prior for θ (you may use any prior you want, conjugate or not, but do not use the Jeffreys prior). Briefly (one sentence), explain why you chose this prior.

   (b) Derive (if you chose a conjugate prior) or draw from the posterior distribution of θ. Report the posterior mean, standard deviation, a 95% posterior interval, and plot the posterior distribution.

   (c) What is the posterior probability that $\theta > 0.5$?

   (d) Re-analyze the data using a Jeffreys prior; do your conclusions from (b) or (c) change?

   (e) Suppose 10 new pairs of patients are enrolled in the study. Based on your posterior, what is the (expected) probability that 6 or more will do better on treatment than control?

   (f) Provide a 90% posterior interval for the quantity in (e).

   (g) Does your posterior represent strong belief that the treatment works? In the paper, the authors claim that deep-brain stimulation is "more effective than medical management." Based on your posterior, do you agree?