Math 775: Homework 5

Alex Dewey

May 11, 2016

1 Notes

1. Start
    2. Start
    3. Done
    5.11. Start
    5.13. Start

2 Exercises

Exercise 1.

a. If our posterior distribution is $p(\theta|y) = Beta(y + 1/2, n - y + 1/2)$ then the true posterior mean is just $E(\theta|y) = \frac{y+1/2}{(y+1/2)+(n-y+1/2)} = \frac{y+1/2}{n+1}$.

b. In order to use the Laplace approximation for our beta-binomial model, we have to write the expectation of the posterior mean in an exponential form:

$$E[\theta|y] = \frac{\int \theta \times \theta^{y-1/2}(1-\theta)^{n-y-1/2}d\theta}{\theta^{y-1/2}(1-\theta)^{n-y-1/2}d\theta} = \frac{\int \theta^{y+1/2}(1-\theta)^{n-y-1/2}d\theta}{\theta^{y-1/2}(1-\theta)^{n-y-1/2}d\theta} =$$

$$\frac{\int e^{(y+1/2)\log\theta}e^{(n-y-1/2)\log 1-\theta}d\theta}{\int e^{(y-1/2)\log\theta}e^{(n-y-1/2)\log 1-\theta}d\theta} =$$

$$\frac{\int e^{(y+1/2)\log\theta+(n-y-1/2)\log(1-\theta)}d\theta}{\int e^{(y-1/2)\log\theta+(n-y-1/2)\log(1-\theta)}d\theta} =$$

$$\frac{\int e^{(y+1/2)\log\theta+((y+1/2)-n)\log\frac{1}{(1-\theta)}}d\theta}{\int \theta^{-1}e^{(y+1/2)\log\theta+((y+1/2)-n)\log\frac{1}{(1-\theta)}}d\theta} =$$

$$\frac{\int e^{((y+1/2)-n)\log\frac{\theta}{(1-\theta)}}d\theta}{\int \theta^{-1}e^{((y+1/2)-n)\log\frac{\theta}{(1-\theta)}}d\theta} =$$

$$\frac{\int \frac{\theta}{(1-\theta)}^{y+1/2} e^{-n \log \frac{\theta}{(1-\theta)}} d\theta}{\int \frac{1}{(1-\theta)}^{y+1/2} e^{-n \log \frac{\theta}{(1-\theta)}} d\theta} =$$

.

$$\frac{\int \frac{\theta}{(1-\theta)}^{y+1/2} e^{-n logit(\theta)} d\theta}{\int \frac{1}{(1-\theta)}^{y+1/2} e^{-n logit(\theta)} d\theta} =$$

$$\frac{\int \frac{1}{(1-\theta)}^{y+1/2} e^{-n(logit(\theta) - \frac{log\theta}{n})} d\theta}{\int \frac{1}{(1-\theta)}^{y+1/2} e^{-n logit(\theta)} d\theta}$$

This is in "fully exponential form", with $h_N(\theta) = logit(\theta) - \frac{log\theta}{n}, h_D(\theta) = logit(\theta)$,

$b_N(\theta) = b_D(\theta) = \frac{1}{(1-\theta)}^{y+1/2}$.

By Corollary 6.2.5 in (C. P. Robert 1999, pp. 298-301),

$$E[g(\theta)|y] = \frac{\hat{b}_N}{\hat{b}_D} \frac{\sigma_N^2}{\sigma_D^2} e^{-n(\hat{h}_N - \hat{h}_D)} + O(n^{-2})$$

,

where $\sigma_D^2 = h_D''(\hat{\theta})^{-1}$, $\hat{\theta} = \min_\theta h(\theta) = 0$, and the other $\hat{h}$ are functions evaluated at $\hat{\theta} = 0$.

$h_N(\theta)$ has a minimum at $0 = \frac{1}{\theta} + \frac{1}{1-\theta} - \frac{1}{n\theta} = \frac{n-1}{n\theta} + \frac{1}{1-\theta}$, so $-\frac{1}{1-\theta} = \frac{n-1}{n\theta}$ and $-n\theta = (n-1)(1-\theta)$

$$E[g(\theta)|y] = \hat{g} + \frac{\sigma_D^2 \hat{b}'_D \hat{g}'}{n\hat{b}_D} + O(n^{-2}) = 0 + \frac{h''(0)^{-1} b'_D(0) g'(0)}{n b_D(0)} + O(n^{-2})$$

.

Exercise 3.

In his short paper "A Problem in Forensic Science" (Biometrika, Aug. 1977), D. V. Lindley explores a Bayesian statistical approach to a common problem of forensics.

Lindley's motivating example is a case where fragments of glass are found in two separate places: at a crime scene and, separately, on a suspect's clothing. Both samples of broken glass has a certain refractive index, and we are interested in whether the two samples of glass were from the same source. The paper's plot proceeds briskly around this forensic problem and Lindley's solution:

First, Lindley describes the problem of broken glass, makes some weak assumptions about the distribution of reflactive indices, and assumes normality at first for the pedagogical purpose of describing his method. Next, Lindley derives and the describes a Bayes-factor test to decide whether the refractive index of the glass fragments are most likely from the same source or from two different sources. Finally, Lindley presents distributional quantiles and some tables and

notes to help the reader interpret his method, ending with some technical notes and an extension to the non-normal case.

The larger problem, of course, is that there are two competing models of reality–one in which the suspect is found to have damning forensic evidence on their person and one in which their pleas of innocence are far more credible. Lindley addresses this specifically near the end, describing even the "type-one and type-two errors" in the explicit terms of a guilty person going free or an innocent person being convicted. How can we make a probability model that is credible enough to have sway in determining someone's ultimate innocence or guilt.

The test derived in the first section of the paper looks superficially like a classical frequentist test. In a classical null-hypothesis significance test, we make an assumption and measure the probability of our observed data under that assumption, and if it's improbable enough we reject the assumption. But appearances are deceiving: While Lindley's test certainly *looks* like an NHST, it differs starkly in at least one critical way: Lindley uses a Bayes factor, the ratio of two complex-looking posterior statistics–one statistic from each hypothesis. This Bayes factor (essentially an odds ratio) allows Lindley to conduct one test which assesses the two probability models *simultaneously*.

Lindley claims this simultaneity is necessary and points to an example where two singleton samples have refractive indices that are far away from each other by their own distributions (each source has its own standard deviation under the "guilty" hypothesis), but are extremely close together when one considers the extraordinarily closeness and rarity on the general distribution of glass. In other words, the samples may not come from the exact same *fragment* of jade, but they're still both jade, so that probably has to mean something.

The point of this example is that if we assess the two models independently, or compare our two hypotheses with a standard NHST, we might not ever grasp the actual probability (or improbability) of what we're seeing. After all, with a sample size of just two or three, improbability at the level of significance may be hard to come by, and the two samples may be several standard deviations apart, which might point in the direction of innocence.

If we're working on these two tests independently, we might be very diligent and note that as the evidence mounts,the numerator (working against the suspect) has moved up and the denominator (working for the suspect) might go down. But if we never piece these two moving parts together, we may have no idea what the machine is telling us. Conversely, if the numerator and denominator go in the *same* direction, mounting evidence in the favor of the prevailing hypothesis may obscure an even greater (though subtler) mountain of countermanding evidence.

Lindley's test allows for simultaneous movement of both the numerator and denominator of the Bayes factor to allow for nonintuitive-but-common situations where an insufficient consideration of both hypotheses makes inference in either direction more difficult. Because it's a solution which "fits" particularly well to its problem, Lindley's test not only uncovers hidden improbabilities but allows for much more robust analysis and decision-making by the statistician or

forensic scientist to effectively balance type-one and type-two errors as well as to incorporate prior domain knowledge.

The science-fiction author Orson Scott Card wrote a sequel to his famous and controversial novel "Ender's Game", called "Ender's Shadow". Unlike most sequels that may only *feel* formulaic, "Ender's Shadow" is explicitly a retelling of the events of the original novel but from a different character's perspective. In the preface, Card explicitly calls the second work "a parallax novel" and hopes that the term will catch on. Parallax is an apt term for Card's narrative approach, a borrowed term from astronomy. The principle of parallax is that we can calculate how far away an object in space is by how much or little it appears to shift its position as we the observers shift position to different telescopes or to different points in the Earth's orbit. The principle of parallax is interesting in that once it's understood, it's difficult to imagine how else one would even make astronomical measurements with anything close to the level of precision and modelling options that this principle affords.

As for the problem this paper describes, Lindley achieves the same effect with his solution.