Homework 3

ISyE 6420
Spring 2020

Due February 16, 2020, 11:55pm. HW3 is not time limited except the due date. Late submissions will not be accepted.

Use of all available electronic and printed resources is allowed except direct communication that violates Georgia Tech Academic Integrity Rules.

1. Traffic. Marietta Traffic Authority is concerned about the repeated accidents at the intersection of Canton and Piedmont Roads. Bayes-inclined city-engineer would like to estimate the accident rate, even better, find a credible set.

A well known model for modeling the number of road accidents in a particular location/time window is the Poisson distribution. Assume that X represents the number of accidents in a 3 month period at the intersection od Canton and Piedmont Roads.

Figure 1: The Intersection od Canton and Piedmont Roads, Marietta, East Cobb, GA

Assume that $[X|\theta] \sim \mathcal{P}oi(\theta)$. Nothing is known a priori about θ, so it is reasonable to assume the Jeffreys' prior

$$\pi(\theta) = \frac{1}{\sqrt{\theta}}\,\mathbf{1}(0 < \theta < \infty).$$

In the four most recent three-month periods the following realizations for X are observed: 1, 2, 0, and 1.

(a) Compare the Bayes estimator for θ with the MLE (For Poisson, recall, $\hat{\theta}_{MLE} = \bar{X}$).

(b) Compute (numerically) a 95% equitailed credible set.

(c) Compute (numerically) a 95% HPD credible set.

(d) Numerically find the mode of the posterior, that is, MAP estimator of θ.

(e) If you test the hypotheses

$$H_0 : \theta \geq 1 \qquad vs \qquad H_1 : \theta < 1,$$

based on the posterior, which hypothesis will be favored?

2. Lady Guessing Coin Flips. There is a famous story involving statistician (Sir Ronald Fisher) and a lady (Muriel Bristol). The lady claimed that, when tasting tea with milk, she can positively determine whether tea or milk had been poured in a cup first. Fisher designs an experiment with 8 cups, 4 in which milk was poured first and 4 in which tea was poured first. Muriel correctly guesses in all 8 cases. This has come to be known as the *Lady Tasting Tea Experiment* and the example is often used to illustrate Fisher's exact test.[1]

Now, imagine another lady who claims ESP and the ability to predict results in subsequent flips of a fair coin. A fair coin is flipped 16 times and the lady correctly predicts the outcome 15 times. Let p be the probability of the lady guessing correctly.

Test a precise hypothesis $H_0 : p = 0.5$ (guessing) versus the alternative $H_1 : p > 0.5$ (presence of ESP), in Bayesian fashion.

To test a precise hypothesis in Bayesian fashion, a prior with a point-mass is needed. Assume that the prior on p is

$$\pi(p) = \pi_0\, \delta_{0.5} + \pi_1\, \mathcal{U}(0.5, 1) = 0.95\, \delta_{0.5} + 0.05 \cdot 2 \cdot \mathbf{1}(0.5 < p < 1),$$

where $\delta_{0.5}$ is point-mass at 0.5, and $\mathcal{U}(a, b)$ is uniform distribution on (a, b). From the form of the prior we see that prior probabilities of the null and alternative hypotheses are $P(H_0) = \pi_0 = 0.95$ and $P(H_1) = \pi_1 = 1 - \pi_0 = 0.05$. Thus, apriori we are sceptic of lady's ESP, and favor H_0.

(a) Find the posterior probabilities of hypotheses, p_0 and p_1, and Bayes Factor.

(b) What is the assessment of H_1 according to Jeffreys scale, i.e., is the experiment convincing that the lady possesses ESP?

Hint. Note that the likelihood is binomial and classical p-value of the test is $P(X \geq 15 | H_0) = \binom{16}{15} 0.5^{15}\, 0.5^1 + \binom{16}{16} 0.5^{16} = 17 \times 0.5^{16} = 0.0002594$. Thus, given the evidence, a frequentist statistician is notoriously biased against H_0, and is strongly convinced of ESP.

Here

$$p_0 = P(H_0|X) = \left[1 + \frac{\pi_1}{\pi_0} \cdot \frac{m_1(x)}{f(x|0.5)}\right]^{-1} \text{ and } B_{01} = \frac{f(x|0.5)}{m_1(x)},$$

[1] Fisher exact test is frequentist testing for homogeneity of 2 x 2 contingency tables with fixed marginals, see page 602 in the text at `http://statbook.gatech.edu`.

where

$$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x},$$

and

$$m_1(x) = \int_{0.5}^{1} f(x|p)\, 2\, dp.$$

To find $m_1(x)$ you will need to numerically evaluate the incomplete beta function.