ASTP 611-01: Statistical Methods for Astrophysics

Problem Set 9

Assigned 2014 April 22
Due 2014 April 29

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 Monte Carlo

Consider an observation where transient events are arriving from all points of the visible sky (above the horizon). An experimenter wishes to test the hypothesis that the locations of these events are isotropic by dividing the sky into approximately (but not quite) equal-area patches and constructing a chi-square statistic out of the expected and observed numbers of events occurring in each patch. Suppose that she chooses six patches, using as coördinates the angle θ down from her zenith and an azimuth ϕ measured around her zenith direction:

- $S_1 \equiv \{0 \le \theta < \frac{\pi}{4}, 0 \le \phi < \pi\}$
- $S_2 \equiv \{0 \le \theta < \frac{\pi}{4}, \pi \le \phi < 2\pi\}$
- $S_3 \equiv \{\frac{\pi}{4} \le \theta < \frac{\pi}{2}, 0 \le \phi < \frac{\pi}{2}\}$
- $S_4 \equiv \{\frac{\pi}{4} \le \theta < \frac{\pi}{2}, \frac{\pi}{2} \le \phi < \pi\}$
- $S_5 \equiv \{\frac{\pi}{4} \le \theta < \frac{\pi}{2}, \pi \le \phi < \frac{3\pi}{2}\}$
- $S_6 \equiv \{\frac{\pi}{4} \le \theta < \frac{\pi}{2}, \frac{3\pi}{2} \le \phi < 2\pi\}$

and observes a total of $n = 100$ events. Then the expected number of events in patch S_i is $E_i = n\Omega_i/(2\pi)$ where $\Omega_i = \int_{\phi_i^{\min}}^{\phi_i^{\max}} \int_{\theta_i^{\min}}^{\theta_i^{\max}} \sin\theta \, d\theta \, d\phi$ is the solid angle of patch S_i, and the chi-squared statistic is

$$W = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i} \tag{1.1}$$

where O_i is the actual number of events observed in patch S_i.

a) Work out the expected numbers of events $E_1 = E_2$ and $E_3 = E_4 = E_5 = E_6$.

b) Perform a Monte Carlo simulation with 100,000 iterations as follows. For each iteration:

   i) Generate $n = 100$ (θ, ϕ) pairs each drawn from the distribution

   $$f(\theta, \phi) = \frac{\sin\theta}{2\pi} \qquad 0 \le \theta \le \pi/2;\ 0 \le \phi < 2\pi \tag{1.2}$$

   (Hint: factor this into independent pdfs for θ and ϕ separately, and use the pdf for each along with a uniform random number generator.)

ii) Count how many of the sky points fall into each patch.

iii) Construct the chi-squared statistic and note its value.

iv) Generate a p-value from this statistic using the chi-squared distribution with the appropriate number of degrees of freedom.

v) As you go along, keep track of how often $p < 0.05$, $p < 0.01$ and $p < 0.001$. Compare these results to your theoretical expectations.

2 Markov Chain Monte Carlo

Use the ipython notebook `http://ccrg.rit.edu/~whelan/courses/2014_1sp_ASTP_611/data/ps09.ipynb` to perform and evaluate MCMC simulations for the Bradley-Terry model. Wherever there is an **EXERCISE**, add code at that point in the notebook to perform the requested calculations.

3 Project Design

Propose a simple monte carlo experiment to test a statistical method decribed in class or in the textbooks in a quasi-realistic scenario. Describe how you'd simulate the data, what analysis you'd perform, and what properties you'd test. Try not to make any more simplifying assumptions than necessary in the simulation stage; one of the things the monte carlo can test is the validity of any approximations in the analysis stage.