ASTP 611-01: Statistical Methods for Astrophysics

Problem Set 8

Assigned 2014 April 10
Due 2014 April 17

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 Bayes Factor

Consider measurements $\{x_i\}$ taken at times $\{t_i\}$, which are assumed differ from values $\{\mu_i\}$ predicted by the "correct" model by uncorrelated Gaussian errors with standard deviations $\{\sigma_i\}$, so that the likelihood function for a model predicting $\boldsymbol{\mu}$ is

$$f(\mathbf{x}|\boldsymbol{\mu}) = \prod_i \frac{1}{\sigma_i\sqrt{2\pi}}\, e^{-(x_i-\mu_i)^2/2\sigma_i^2} = \frac{1}{\sqrt{\det 2\pi\boldsymbol{\sigma}^2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\sigma}^{-2}(\mathbf{x}-\boldsymbol{\mu})\right) \tag{1.1}$$

The measured values and corresponding uncertainties are in the data file which can be downloaded from `http://ccrg.rit.edu/~whelan/courses/2014_1sp_ASTP_611/data/ps08.dat`

a) Consider a model $\mathcal{H}_0$ in which $\mu_i = 0$. Calculate the likelihood $f(\mathbf{x}|\mathcal{H}_0)$ for the data above.

b) Consider an alternative model $\mathcal{H}_1$ in which $\mu_i = \theta$ where θ is a single parameter obeying a (normalized) uniform prior $f(\theta|\mathcal{H}_1)$ ranging over values $-5 < \theta < 5$. Calculate the marginalized likelihood

$$f(\mathbf{x}|\mathcal{H}_1) = \int_{-5}^{5} d\theta\, f(\mathbf{x}|\theta,\mathcal{H}_1)\, f(\theta|\mathcal{H}_1) \tag{1.2}$$

(You may do this by numerical integration, or a semi-analytic calculation involving the error function.)

c) Calculate the Bayes factor $\mathcal{B}_{10} = f(\mathbf{x}|\mathcal{H}_1)/f(\mathbf{x}|\mathcal{H}_0)$ relating the two models.

d) Find (either analytically or numerically) the value $\hat{\theta}$ which maximizes the likelihood $f(\mathbf{x}|\theta,\mathcal{H}_1)$ and calculate the ratio

$$\frac{f(\mathbf{x}|\hat{\theta},\mathcal{H}_1)}{f(\mathbf{x}|\mathcal{H}_0)} \tag{1.3}$$

Compare this to the results of part c) and comment on the effect of the Occam factor in this problem.

e) Calculate the χ^2 statistic for each of the two models and comment on the corresponding frequentist comparison between the models.

2 Optional Stopping

Note: this problem is closely related to section 7.4 of Gregory, which I encourage you to refer to for additional insight and commentary.

Consider the ESP experiment described in class, where the null hypothesis $\mathcal{H}_0$ is that the psychic has a 25% chance to guess each card's suit and the alternative hypothesis $\mathcal{H}_1$ is that the psychic has some imperfect ESP which a probability $0.25 < \theta < 1$ to guess each card's suit correctly. Suppose that we observe a test in which 7 cards are guessed correctly and 13 incorrectly.

a) Considering this to be a binomial experiment in which the observable is K, the number of correct guesses out of 20, find the p-value

$$p_b = P(K \geq 7|\mathcal{H}_0) \tag{2.1}$$

Express this as an explicit sum over k (including appropriate limits for the sum) and evaluate the sum numerically using the software package of your choice. (Doing it by hand is in principle possible, but not recommended!)

b) It so happens that the last card was guessed correctly, and afterwards the experimenter tells you that rather than trying 20 cards, she had decided to stop after 7 correct guesses, and therefore the observed quantity is actually the number of incorrect guesses, which we'll call X.

   i) The probability distribution which describes this situation is known as the negative binomial distribution. Recall that, if the probability of success on each trial is θ, the probability of any specific sequence of k successes and x failures is $\theta^k(1-\theta)^x$. The number of possible sequences of k successes and x failures which ends with a success is the same as the number of sequences of $k-1$ successes and x failures. Use this information to construct the total probability of getting a sequence of k successes and x failures which ends with a success, which is the pmf $p(x|k,\theta)$ for the negative binomial distribution. Which values of x are possible?

   ii) Since the experiment was designed to end after $k = 7$ correct guesses, the p-value is defined as

$$p_{nb} = P(X \leq 13|\mathcal{H}_0) \tag{2.2}$$

   Again, express this as an explicit sum over x and then evaluate it with a computer. Compare this p-value to the p_b calculated in part a).

c) The Bayesian version of this problem finds the Bayes factor

$$\mathcal{B} = \frac{P(D|\mathcal{H}_1)}{P(D|\mathcal{H}_0)} \tag{2.3}$$

which is the ratio of the evidence (marginalized likelihood) for two models. As before $\mathcal{H}_0$ is the model where the probability of success is 0.25. Our definition above of $\mathcal{H}_1$

as having $0.25 < \theta < 1$ isn't specific enough to let us do the marginalization, but in the absence of other information, we'll assume a prior

$$f(\theta|\mathcal{H}_1) = \begin{cases} \frac{4}{3} & \frac{1}{4} < \theta < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.4}$$

Work out expressions for the evidences $P(D|\mathcal{H}_0)$ and $P(D|\mathcal{H}_1)$, and the Bayes factor $\mathcal{B}$ in each of the following cases. You should leave the expressions in terms of integrals, powers, binomial coëfficients, etc, and don't worry about evaluating them yet.

i) The binomial experiment where $n = 20$ and $D \equiv D_b \equiv \{K = 7\}$.
ii) The negative binomial experiment where $k = 7$ and $D \equiv D_{nb} \equiv \{X = 13\}$.
iii) An interpretation of the experiment where $D \equiv D_{\text{spec}}$ is the specific sequence of 7 successes and 13 failures which was seen in the experiment.

d) Evaluate the Bayes factor found in each of the parts of part c), to get a numerical value. (The marginalization integral over θ is impractical to evaluate by hand, since it involes a big multinomial expansion, but its value is related to the incomplete beta function (q.v.), which can be found in your favorite numerical software package.)

3 Critical Region

Using the ipython notebook `http://ccrg.rit.edu/~whelan/courses/2014_1sp_ASTP_611/data/ps08.ipynb` consider the critical regions for chi-square tests of two null hypotheses:

a) $\mathcal{H}_0$, where the data we collect, z_1 and z_2, are a sample drawn from a standard normal ($N(0,1)$) distribution.
b) $\mathcal{H}_0(\theta)$, where the data we collect, z_1 and z_2, are a sample drawn from a $N(\theta, 1)$ distributions, i.e., Gaussian with unit variance and unknown mean.

Submit your completed notebook (preferably as a hardcopy).