ASTP 611-01: Statistical Methods for Astrophysics

Problem Set 7

Assigned 2014 April 1
Due 2014 April 8

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 Fisher Matrix for a Counting Experiment

a) Consider an experiment where the observed quantity is the number of events x in a particular interval of a Poisson process, where the parameter θ is the expected number of events (rate times interval duration) for that process.

   i) Write the likelihood function $L(\theta) = f(x|\theta)$ for a given observed value of x.

   ii) Find the maximum likelihood estimate $\widehat{\theta}$ by maximizing the log-likelihood $\ell(\theta) = \ln L(\theta)$.

   iii) Find the one-dimensional "Fisher matrix" $-\ell''(\widehat{\theta})$. What is the corresponding one-sigma uncertainty in the parameter θ?

b) Now consider the case where we have n counting experiments with numbers of counts $\{x_i|i = 1, \ldots, n\}$. Suppose that each x_i is drawn from a Poisson distribution with its own parameter μ_i. In particular, suppose we're counting the number of photons collected in a series of bins with frequencies $\{\nu_i|i = 1, \ldots, n\}$, and the model we're trying to fit is a background which is approximately linear over the band of interest, so $\mu_i = \theta_0 + \theta_1\nu_i$, and the are parameters θ_0 and θ_1.

   i) Write the likelihood function $L(\theta_0, \theta_1) = f(\{x_i\}|\theta_0, \theta_1)$.

   ii) Work out the equations satisfied by the best-fit parameters $\widehat{\theta}_0$ and $\widehat{\theta}_1$ which maximize the log-likelihood $\ell(\theta_0, \theta_1) = \ln L(\theta_0, \theta_1)$. For the case $n = 3$, $\nu_1 = -1$, $\nu_2 = 0$, $\nu_3 = 1$, solve the equations to get $\widehat{\theta}_0$ and $\widehat{\theta}_1$ as functions of the $\{x_i\}$.

   iii) Work out the second derivatives $\frac{\partial^2\ell}{\partial\theta_\alpha\partial\theta_\beta}$, where $\alpha, \beta \in \{0, 1\}$. For the $n = 3$ case described above, evaluate these at the maximum likelihood point to get the elements $F_{\alpha\beta} = -\left.\frac{\partial^2\ell}{\partial\theta_\alpha\partial\theta_\beta}\right|_{\theta_0=\widehat{\theta}_0, \theta_1=\widehat{\theta}_1}$ of the Fisher matrix.

c) In a more realistic situation, we'd be trying to fit something like a line plus a continuum, so for instance $\mu_i = a + b\nu_i + \frac{I}{1+([\nu_i-\nu_0]/\gamma)^2}$ and then the parameters would be a, b, I, ν_0, and γ. Perhaps some would be known, perhaps not, and we might marginalize the likelihood or the posterior over parameters we didn't care about. (No question, just food for thought.)

2 Plausible Intervals

Consider a posterior probability distribution of the form

$$f(\theta|D) = \frac{\theta^2}{2} e^{-\theta} \qquad 0 < \theta < \infty \tag{2.1}$$

a) Use the ipython notebook `http://ccrg.rit.edu/~whelan/courses/2014_1sp_ASTP_611/data/ps07.ipynb` to determine the lower and upper bounds of the following 90% plausible intervals:

   i) An upper limit (so the lower bound is 0)

   ii) A lower limit (so the upper bound is ∞)

   iii) A symmetric plausible interval, so $P(\theta < \theta_\ell) = 5\% = P(\theta > \theta_u)$

   iv) The narrowest 90% plausible interval which can be constructed

   In each case, plot the pdf $f(\theta|D)$ with the area under the curve between θ_ℓ and θ_u shaded. Turn in the notebook with the relevant plots and calculations, and include the bounds of the plausible intervals in your homework solution.

b) Find the mode $\widehat{\theta}$ of the distribution, i.e., the θ which maximizes $f(\theta|D)$. Why does it not make sense to construct a confidence interval centered on $\widehat{\theta}$?

3 Upper Limits

Consider an experiment designed to measure an unknown physical quantity θ, which returns a value X whose pdf is defined by the likelihood function

$$f(x|\theta) = \frac{e^{-(x-\theta)^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \tag{3.1}$$

a) Suppose the experiment has been performed and the result x has been found. Calculate the frequentist upper limit $\theta_{\text{UL}}^{\text{freq}}$ at confidence level α, defined by

$$\int_x^\infty f(x'|\theta_{\text{UL}}^{\text{freq}})\, dx' = \alpha \ . \tag{3.2}$$

   You should be able to write this with the help of the inverse complementary error function $\operatorname{erfc}^{-1}(\xi)$, defined as the inverse of $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\, dt$. Note that $\operatorname{erfc}^{-1}(\xi)$ is positive if $0 < \xi < 1$ and negative if $1 < \xi < 2$, and that $\operatorname{erfc}^{-1}(2-\xi) = -\operatorname{erfc}^{-1}(\xi)$

b) Consider a Bayesian analysis with a uniform prior on θ, so that by Bayes's theorem, the posterior is

$$f(\theta|x) = \frac{f(\theta)}{f(x)}\, f(x|\theta) = \mathcal{A}\, f(x|\theta) \ . \tag{3.3}$$

   Using the explicit form of the likelihood (3.1) and the normalization requirement

$$\int_{-\infty}^\infty f(\theta|x)\, d\theta = 1 \tag{3.4}$$

   find the value of $\mathcal{A}$ and therefore the explicit form of the posterior $f(\theta|x)$.

c) Supposing again that we've performed the experiment and found a result x, find the Bayesian upper limit $\theta_{\text{UL}}^{\text{Bayes}}$ at confidence level α, defined by

$$\int_{-\infty}^{\theta_{\text{UL}}^{\text{Bayes}}} f(\theta|x)\, d\theta = \alpha \tag{3.5}$$

d) For the case where $\alpha = 0.9$, write $\theta_{\text{UL}}^{\text{freq}}$ and $\theta_{\text{UL}}^{\text{Bayes}}$ explicitly in terms of x and σ, with any constants evaluated to three significant figures. (You'll need to refer to the explicit value of $\operatorname{erfc}^{-1}(\xi)$ for a particular ξ; in matplotlib you can get access to the inverse complementary error function via `from scipy.special import erfcinv`.)

e) Suppose now that θ is physically constrained to be positive and let the prior be uniform for positive θ, so that the posterior is

$$f(\theta|x) = \frac{f(\theta)}{f(x)}\, f(x|\theta) = \begin{cases} \mathcal{B}\, H(\theta)\, f(x|\theta) & \theta > 0 \\ 0 & \theta < 0 \end{cases} . \tag{3.6}$$

Use the normalization condition

$$1 = \int_0^\infty f(\theta|x)\, d\theta = \mathcal{B} \int_0^\infty f(x|\theta)\, d\theta \tag{3.7}$$

to find the value of $\mathcal{B}$ and therefore the explicit form of $f(\theta|x)$.

f) Supposing again that we've performed the experiment and found a result x, calculate the Bayesian upper limit $\theta_{\text{UL}}^{\text{Bayes+}}$ associated with the posterior (3.6), defined by

$$\int_0^{\theta_{\text{UL}}^{\text{Bayes+}}} f(\theta|x)\, d\theta = \alpha \tag{3.8}$$