# 1060-710
# Mathematical and Statistical Methods for Astrophysics

Problem Set 6

Assigned 2010 October 19
Due 2010 October 26

**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 Upper Limits

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

$$f(y|x) = \frac{e^{-(y-x)^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \tag{1.1}$$

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

$$\int_{\hat{y}}^{\infty} f(y|x_{\text{UL}}^{\text{freq}})\, dy = \alpha \ . \tag{1.2}$$

You should be able to write this with the help of the inverse complementary error function $\operatorname{erfc}^{-1}(\xi)$. 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 $x$, so that by Bayes's theorem, the posterior is

$$f(x|y) = \frac{f(x)}{f(y)}\, f(y|x) = \mathcal{A}\, f(y|x) \ . \tag{1.3}$$

Using the explicit form of the likelihood (1.1) and the normalization requirement

$$\int_{-\infty}^{\infty} f(x|y)\, dx = 1 \tag{1.4}$$

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

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

$$\int_{-\infty}^{x_{\text{UL}}^{\text{Bayes}}} f(x|\hat{y})\, dx = \alpha \tag{1.5}$$

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

e) Suppose now that $x$ is physically constrained to be positive and let the prior be uniform for positive $x$, so that the posterior can be written in terms of the Heaviside step function

$$\Theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases} \tag{1.6}$$

as

$$f(x|y) = \frac{f(x)}{f(y)}\, f(y|x) = \mathcal{B}\, \Theta(x)\, f(y|x) \ . \tag{1.7}$$

Use the normalization condition

$$1 = \int_0^\infty f(x|y)\, dx = \mathcal{B} \int_0^\infty f(y|x)\, dx \tag{1.8}$$

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

f) Supposing again that we've performed the experiment and found a result $\hat{y}$, calculate the Bayesian upper limit $x_{\text{UL}}^{\text{Bayes+}}$ associated with the posterior (1.7), defined by

$$\int_0^{x_{\text{UL}}^{\text{Bayes+}}} f(x|\hat{y})\, dx = \alpha \tag{1.9}$$

# 2 Marginalization and the Inverse Fisher Matrix

Consider two variables $X_1$ and $X_2$ whose joint pdf is a Gaussian with zero mean:

$$f(\mathbf{x}) = \frac{\sqrt{\det \mathbf{F}}}{2\pi} \exp\left[-\frac{1}{2}\mathbf{x}^{\text{T}}\,\mathbf{F}\,\mathbf{x}\right] = \frac{\sqrt{F_{11}F_{22} - F_{12}^2}}{2\pi} \exp\left[-\frac{F_{11}}{2}(x_1)^2 - F_{12}x_1x_2 - \frac{F_{22}}{2}(x_2)^2\right] \tag{2.1}$$

where $\mathbf{F}$ is some symmetric, positive definite matrix.

a) Show that $\mathbf{F}$ is indeed the Fisher matrix.

b) Marginalize over $x_2$ and show that the resulting pdf for $x_1$ is a Gaussian whose variance is the $1,1$ component of the inverse Fisher matrix $\mathbf{F}^{-1}$:

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 = \frac{1}{\sqrt{2\pi\,(F^{-1})_{11}}} \exp\left(-\frac{x_1^2}{2(F^{-1})_{11}}\right) \tag{2.2}$$

# 3 Least Squares and Chi-Squared

Consider measurements $\{y_i\}$ taken at times $\{t_i\} = \{-1, 0, 1, 2\}$. We wish to fit these measurements with a straight-line model with predicted expectation values $\mu_i = \lambda_1 + \lambda_2 t_i$. The model predicts measurments which differ from $\mu_i$ by uncorrelated Gaussian errors with standard deviations $\{\sigma_i\} = \{\sqrt{2}, 1, \sqrt{2}, \sqrt{3}\}$.

a) Find the matrix $\mathbf{A}$ describing the linear relationship $\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\lambda}$, i.e.,

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \mathbf{A} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \tag{3.1}$$

b) Since the errors are uncorrelated, the standard deviations are described by a matrix

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{pmatrix} . \tag{3.2}$$

Construct the matrix $\mathbf{A}^{\mathrm{T}}\boldsymbol{\sigma}^{-2}\mathbf{A}$ and find its inverse $\left[\mathbf{A}^{\mathrm{T}}\boldsymbol{\sigma}^{-2}\mathbf{A}\right]^{-1}$. (Since it is a $2 \times 2$ matrix, you should actually be able to invert it by hand.)

c) In class we showed that if the measured values are $\mathbf{y}$, the maximum likelihood estimates of the parameters will be $\hat{\boldsymbol{\lambda}}(\mathbf{y}) = \left[\mathbf{A}^{\mathrm{T}}\boldsymbol{\sigma}^{-2}\mathbf{A}\right]^{-1} \mathbf{A}^{\mathrm{T}}\boldsymbol{\sigma}^{-2}\mathbf{y}$. Work out the elements of the matrix appearing for this problem in

$$\begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix} = \left[\mathbf{A}^{\mathrm{T}}\boldsymbol{\sigma}^{-2}\mathbf{A}\right]^{-1} \mathbf{A}^{\mathrm{T}}\boldsymbol{\sigma}^{-2} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \tag{3.3}$$

d) Suppose we measure $\{y_i\} = \{1.07241020, 0.40438919, 2.89906726, 8.98526374\}$. Calculate, to three significant figures,

   i) The best-fit parameters $\hat{\lambda}_1$ and $\hat{\lambda}_2$

   ii) The $\chi^2$ value relating the data to the best-fit model,

$$\chi^2 = (\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\lambda}})^{\mathrm{T}}\boldsymbol{\sigma}^{-2}(\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\lambda}}) \tag{3.4}$$

   iii) The $p$ value, i.e., probability that data generated according to the model would have a $\chi^2$ equal to or higher than the one observed.