1060-710
Mathematical and Statistical Methods for Astrophysics

Problem Set 5

Assigned 2009 October 15
Due 2009 October 22

**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 Colored Noise

This problem should be done in your favorite numerical analysis environment. Please turn in a printout of the final set of commands you used as well as the plots themselves.

Consider a $T = 16$ sec of data sampled at $\delta t = 1024$ Hz, so that $N = 16384$.

a) Generate an $N$-point series $\{x_j\}$ of random samples drawn from a distribution with zero mean and unit variance, i.e., $\langle x_j \rangle = 0$ and $\langle x_j x_\ell \rangle = \delta_{j\ell}$. (This can be done in matplotlib with `randn(N)`.) Plot $x_j$ versus $t_j - t_0 = j \delta t$ for $t_j - t_0$ from 0 to 16 sec.

b) Take the discrete Fourier transform $\hat{x}_k$ and plot $|\hat{x}_k|$ versus $f_k = k \delta f = k/T$ for $f_k$ from $-512$ Hz to 512 Hz.

c) Calculate the average

$$\frac{1}{N} \sum_{k=-N/2}^{N/2-1} |\hat{x}_k|^2$$

(1.1)

How does it compare to your theoretical expectations for $\langle |\hat{x}_k|^2 \rangle$?

d) Construct $\hat{y}_k = \hat{x}_k e^{-f_k^2/2\sigma_f^2}$, where $\sigma_f = 16$ Hz, for $k \in [-N/2, N/2 - 1]$ and plot $|\hat{y}_k|$ versus $f_k$ for $f_k$ from $-512$ Hz to 512 Hz.

e) Take the inverse discrete Fourier Transform to get $y_j$ and plot $y_j$ versus $t_j - t_0$ for $t_j - t_0$ from 0 to 16 sec.


2 Power Spectral Density

Consider a single random variable \( \psi \) which is equally likely to fall anywhere between 0 and \( 2\pi \), so that expectation values of random variables whose only randomness comes from \( \psi \) can be calculated as

\[
\langle F(\psi) \rangle = \frac{1}{2\pi} \int_0^{2\pi} F(\psi) \, d\psi .
\] (2.1)

Let \( x(t) = A \cos(2\pi f_0 t + \psi) \) for some fixed \( A \) and \( f_0 \).

a) Find \( \langle x(t) \rangle \) and \( \langle x(t) x(t') \rangle \) and show that \( x(t) \) is wide-sense stationary.

b) Find the power spectral density \( P_x(f) \).

3 Averages and Expectation Values

Let \( \{x_k|k = 0 \ldots N-1\} \) be a set of \( N \) uncorrelated random variables all drawn from the same distribution with (possibly unknown) mean \( \mu \) and variance \( \sigma^2 \) so that the expectation values are

\[
\langle x_k \rangle = \mu \tag{3.1a}
\]
\[
\langle (x_k - \mu)(x_\ell - \mu) \rangle = \delta_{k\ell} \sigma^2 \tag{3.1b}
\]

a) Consider the average of the \( N \) instantiations

\[
\bar{x} = \frac{1}{N} \sum_{k=0}^{N-1} x_k \tag{3.2}
\]

and show that its expectation value \( \langle \bar{x} \rangle \) is equal to \( \mu \). (This is pretty easy to show if you remember that the expectation value is a linear operation, so that \( \langle \alpha + \beta \rangle = \langle \alpha \rangle + \langle \beta \rangle \).)

This means that \( \bar{x} \) is an unbiased estimator of the mean of the underlying distribution, even though it’s constructed from a finite number of samples from that distribution.

b) Calculate the expected variance of \( \bar{x} \), i.e.,

\[
\langle (\bar{x} - \langle \bar{x} \rangle)^2 \rangle = \langle (\bar{x} - \mu)^2 \rangle = \text{what?} \tag{3.3}
\]

c) Suppose we know the exact value of \( \mu \) (via some sort of physical principle or something); we could use

\[
\frac{(x - \mu)^2}{N} = \frac{1}{N} \sum_{k=0}^{N-1} (x_k - \mu)^2 \tag{3.4}
\]
as an estimator of the underlying variance \( \sigma^2 \). Show that this is an unbiased estimator, i.e., that its expectation value is indeed \( \sigma^2 \):

\[
\langle (\bar{x} - \mu)^2 \rangle = \sigma^2 \tag{3.5}
\]
d) Now suppose the true mean $\mu = \langle x \rangle$ is not known, and must also be estimated from the same $N$ data points. We can consider

$$\overline{(x - \bar{x})^2} = \frac{1}{N} \sum_{k=0}^{N-1} (x_k - \bar{x})^2$$

(3.6)

as a potential estimator of $\sigma^2$. Calculate its expectation value

$$\langle \overline{(x - \bar{x})^2} \rangle,$$

(3.7)

keeping in mind that both $x_k$ and $\bar{x}$ are random variables. This will not be equal to $\sigma^2$ which means (3.6) is a biased estimator.

e) Modify (3.6) to produce an unbiased estimator of $\sigma$ which can be calculated from only the samples $\{x_k\}$. (This can’t include the unknown actual value of $\mu$ nor any expectation values, only averages calculated from the actual samples $\{x_k\}$.)