Lomb-Scargle Periodogram

- Lomb (1976) - Scargle (1982) Periodogram:

$$P_{\mathrm{LS}}(\nu) = \frac{1}{2} \frac{\left\{ \sum_{i=1}^{N} x(t_i) \cos[2\pi\nu(t_i - \tau)] \right\}^2}{\sum_{i=1}^{N} \cos^2[2\pi\nu(t_i - \tau)]} + \frac{\left\{ \sum_{i=1}^{N} x(t_i) \sin[2\pi\nu(t_i - \tau)] \right\}^2}{\sum_{i=1}^{N} \sin^2[2\pi\nu(t_i - \tau)]}.$$

  - Good for general uneven sampling
  - Equivalent to linear least-square fit to sin+cos
  - Statistically robust

Lomb-Scargle Periodogram

- In this expression, the reference epoch τ is chosen in such a way that:

$$\sum_{i=1}^{N} \cos[2\pi\nu(t_i - \tau)] \sin[2\pi\nu(t_i - \tau)] = 0,$$

- Or, equivalently

$$\tan(4\pi\nu\tau) = \frac{\sum_{i=1}^{N} \sin(4\pi\nu t_i)}{\sum_{i=1}^{N} \cos(4\pi\nu t_i)}.$$

- It looks complicated, but it's basically the regular periodogram adapted to handle unevenly spaced data. In the limit of equal spacing, it actually reduces to the classical result.

Lomb-Scargle Periodogram

- One of the reasons to have introduced the Lomb-Scargle periodogram is that its value does not change when all time values t_i are replaced by $t_i + T$ because of the definition of τ.

Properties of the Lomb-Scargle periodogram

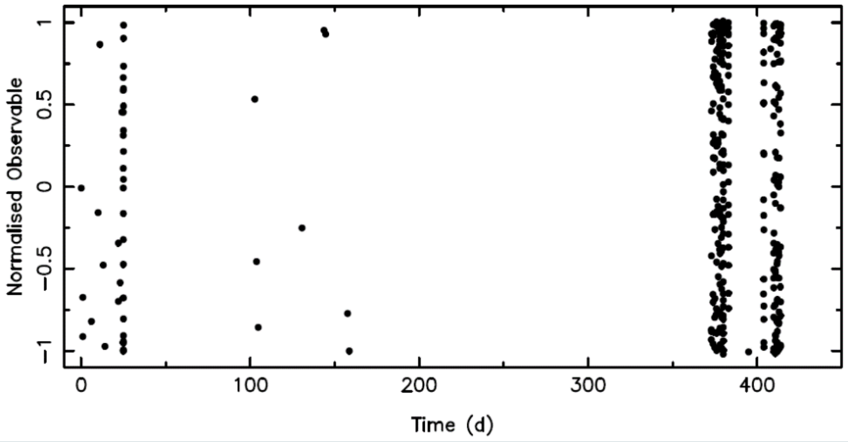
- The most important feature of the Lomb-Scargle periodogram is the significance of the power at an individual frequency:

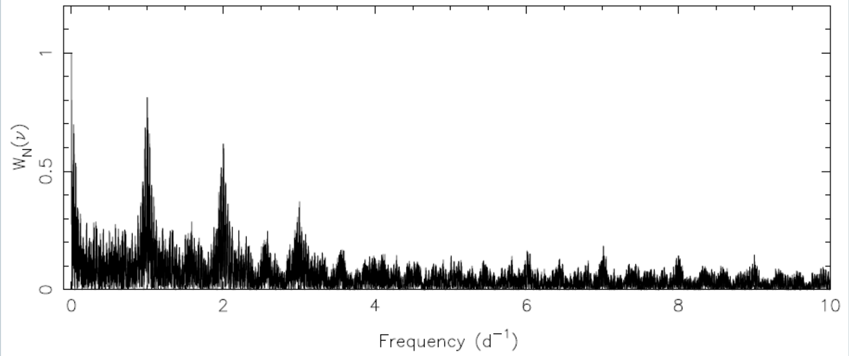
$$\mathrm{Prob}(P(\nu)>P_{det})=\exp(-P_{det})$$

- You still have to worry about the number of independent frequencies you test to account for trials factors

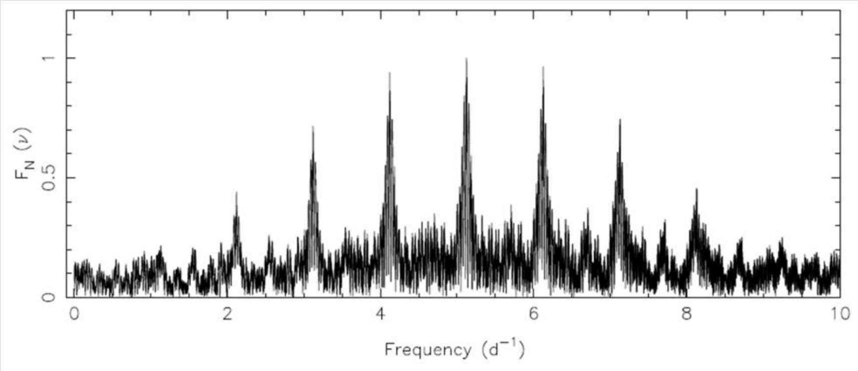
Spectral Analysis with Unevenly-Spaced Data

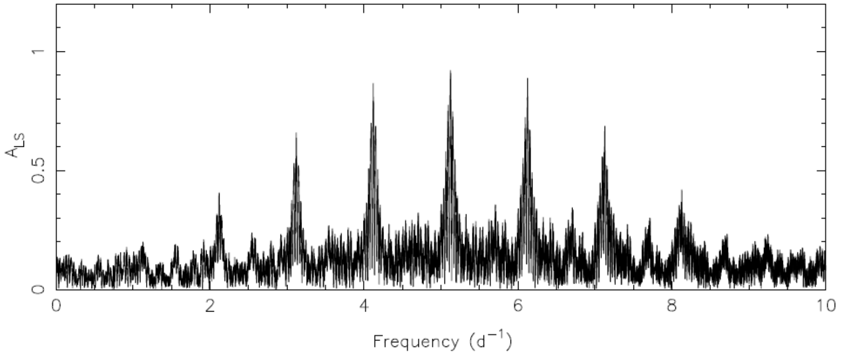
The gapped data and its Spectral window





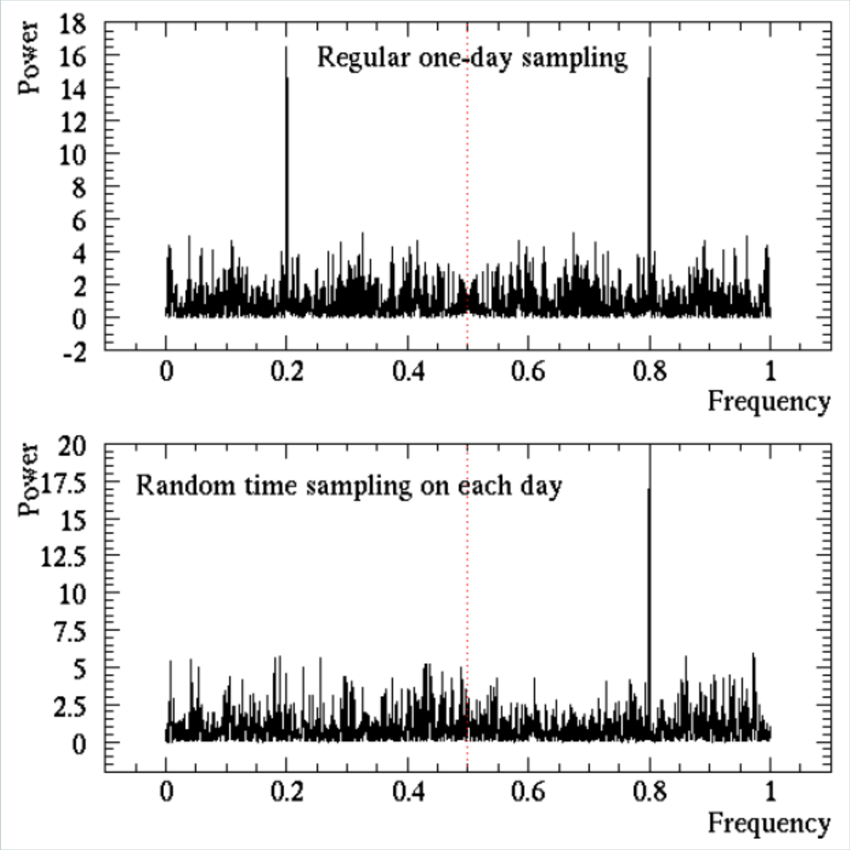
DFT (top) and LS periodogram (bottom)





Advantages of non-uniform sampling

- Consider the following Lomb-Scargle periodograms of an f=0.8 signal.
  Top: sampling exactly once per day at noon
  Bottom: sampling once per day at a random time within the 24-hour period.
- Regular sampling gives strong alias peak at f=0.2: in fact Nyquist frequency is f<0.5, so you'd conclude there was really a signal at f=0.2
- Random sampling gets rid of alias peak! And it gives sensitivity to higher frequencies - since random times can be close to each other, Nyquist cutoff is not a hard limit anymore!



Combined analysis of power spectra

- In astronomy, it is often necessary to compare the power spectra of two or more time series, e.g.:
  - X-ray binaries are often observed simultaneously in X-rays, UV and optical wavelengths (and γ-rays).
  - the Sun has been observed more or less routinely for many years and in a variety of modes (sunspots, radio, UV, X-ray, irradiance, etc.), so one may need to compare two or more solar data sets.
- One might also wish to estimate the significance of a particular peak that shows up in two or more power spectra.

Combined analysis of power spectra

- Assuming, that each power spectrum is distributed **exponentially** (e.g., Lomb-Scargle periodogram), Sturrock et al. (2005) proposed three such statistics, that are useful for the combined study of two or more time-series:
  - Combined Power Statistic
  - Minimum Power Statistic
  - Joint Power Statistic

The paper is on the course web-page

Combined analysis of power spectra

- **Combined Power Statistic**
  - If we wish to combine information from n independent power spectra, the combination that would correspond to the chi-square statistic is the sum of the powers, which we write as

$$\boldsymbol{Z = S_1 + S_2 + \cdots + S_n}.$$

  - The following function of Z ("**combined power statistic**") is distributed exponentially:

$$G_n(Z) = Z - \ln\left(1 + Z + \frac{1}{2}Z^2 + \cdots + \frac{1}{(n-1)!}Z^{n-1}\right).$$

Combined analysis of power spectra

- **Minimum Power Statistic**
  - We may wish to determine the frequency for which the minimum power among two or more power spectra has the maximum value. Let's consider the following quantity, formed from the independent variables $x_1, x_2, \ldots, x_n$, each of which is distributed exponentially:

$$\boldsymbol{U(x_1, x_2, \ldots, x_n) = \mathrm{Min}(x_1, x_2, \ldots, x_n)}$$

  - It can be shown that the following function of U ("**minimum power statistic**") is distributed exponentially:

$$\boldsymbol{K_n(U) = nU}$$

Combined analysis of power spectra

- **Joint Power Statistic**
  - Let’s now consider the need to compare spectra from two quite different times series. If one of the time-series has very strong peaks and the other has comparatively weak peaks, then simply adding the powers would not be very revealing, since the sum would be dominated by the stronger spectrum.
  - In this situation, it is more useful to form something resembling a “correlation function" by forming the product of the two powers. It proves convenient to work with the square root of the product (the geometric mean):

$$X = (S_1 S_2)^{1/2}$$

Combined analysis of power spectra

- **Joint Power Statistic (cont.)**
  - The following function of X is distributed exponentially:

$$\boldsymbol{J_2 = -\ln\left(2X\,K_1(2X)\right)}$$

where $K1$ is the Bessel function of the second kind.

  - A good approximation to $\boldsymbol{J_2}$ is found to be:

$$J_{2A} = \frac{1.943\,X^2}{0.650 + X}.$$

Combined analysis of power spectra

- **Joint Power Statistic (cont.)**

  - Let's now consider joint power statistics of higher orders, and consider the following geometric mean of *n* powers:

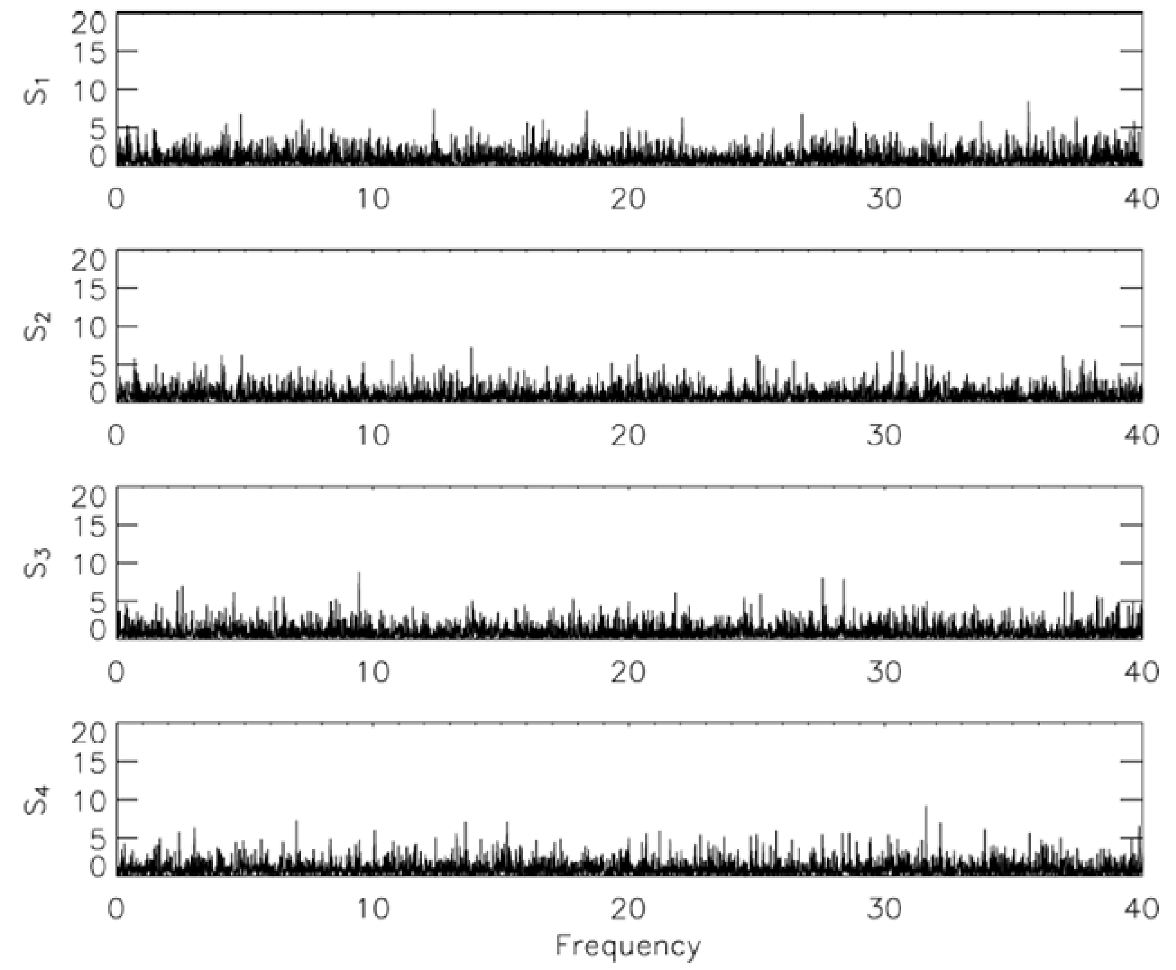
$$X = (S_1 \ldots S_n)^{1/n}$$

  - There is no useful analytical functions of *X* that are distributed exponentially, but there are very good approximations:

$$J_{3A} = \frac{2.916X^2}{1.022 + X}, \qquad J_{4A} = \frac{3.881X^2}{1.269 + X}.$$

Combined analysis of power spectra

Four synthetic spectra, each with a signal of power 5 at $\nu = 20$.



Combined analysis of power spectra

The combined power statistic, minimum power statistic, and joint power statistic, formed from the four synthetic spectra.