Periodogram & Power Spectrum

- The periodogram is an estimate of the spectral density of a signal. The term was coined by Arthur Schuster in 1898 (*the Schuster Periodogram*).
- A Power Density Spectrum is computed as the squared Fourier amplitudes with **some normalization**:

$$a_j = \sum_{k=0}^{N-1} x_k \, e^{2\pi i jk/N} \qquad j = -\frac{N}{2} + 1, \dots, \frac{N}{2}$$

$$P_j = (Normalization)|a_j|^2$$

Power Spectrum – Leahy Normalization

- We will adopt the Leahy et al. (1983) normalization:

$$P_j = \frac{2}{N_{tot}} |a_j|^2 \quad j = 0, \dots, \frac{N}{2}; \; where \; N_{tot} = N_{ph} = \sum_k x_k = a_0$$

N_{tot} – dispersion of the total number of counts in the time series. For the Poisson process, the variance (square of the standard deviation) is equal to the total number of counts.

Properties of Leahy normalized PDS

- Variance in the real time series x_k:

$$Var(x_k) \equiv \sum_k (x_k - \overline{x})^2 = \sum_k x_k{}^2 - \frac{1}{N}\left(\sum_k x_k\right)^2 =$$

$$= \frac{1}{N}\sum_j |a_j|^2 - \frac{1}{N} a_0{}^2 = \frac{1}{N}\sum_{j\neq 0} |a_j|^2$$

Parseval's theorem

$$Var(x_k) = \frac{N_{tot}}{N}\left(\sum_{j=1}^{N/2-1} P_j + \frac{1}{2} P_{N/2}\right)$$

variance is sum of powers!

The dimension of P_j is the same as x_k and a_j: $[P_j] = [a_j] = [x_k]$

Properties of Leahy normalized PDS

- Fractional rms (root-mean-square) amplitude of a signal in a time series x_k:

$$r \equiv \frac{\sqrt{\frac{1}{N} Var(x_k)}}{\bar{x}} = \frac{N}{N_{tot}} \sqrt{\frac{N_{tot}}{N^2} \left(\sum_{j=1}^{N/2-1} P_j + \frac{1}{2} P_{\frac{N}{2}} \right)}$$

$$= \sqrt{\frac{1}{N_{tot}} \left(\sum_{j=1}^{N/2-1} P_j + \frac{1}{2} P_{N/2} \right)}$$

r is dimensionless and often expressed in % (percentage rms variation).

Properties of Leahy normalized PDS

- "rms normalized" power density: $q(\nu_j) \equiv TP_j/N_{ph}$
  physical unit of $q(\nu_j)$ is (rms/mean)²/Hz

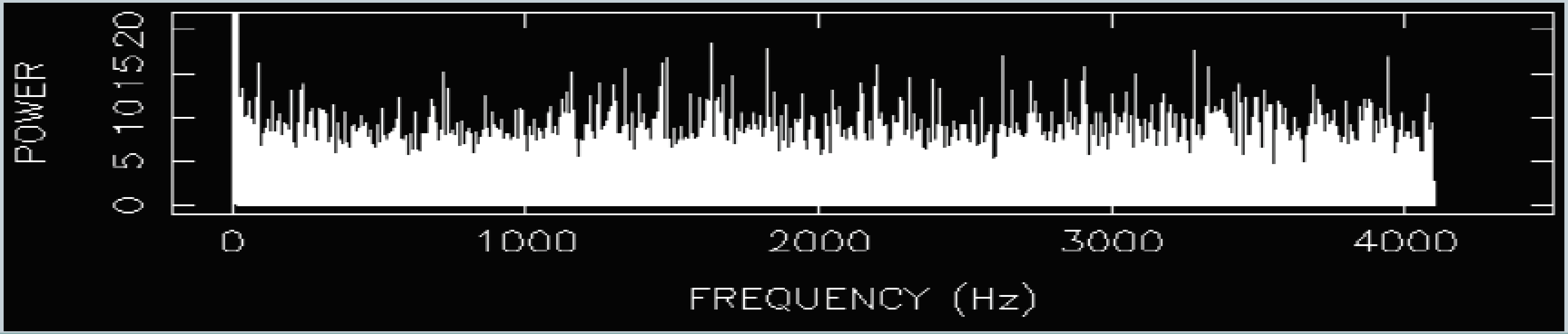
- "Source" fractional rms amplitude: If the x_k are the sum of source and background: $x_k = b_k + s_k$, then the rms amplitude as a fraction of just the s_k:
  $r_s = r\frac{B+S}{S}$,
  where B and S are sums of the b_k and s_k, so $B+S = \sum_k x_k = N_{ph}$

- "Source rms normalized" power density ("Miyamoto" normalization):
  $q_S \equiv q\left(\frac{B+S}{S}\right)^2 = TP_j\frac{B+S}{S^2}$
  the same unit as q: (rms/mean)²/Hz

  Requires a model or a measurement of the background count rate.

Properties of Leahy normalized PDS

- The Leahy normalization is chosen such that if the x_k are Poisson distributed, then the P_j exactly follow the chi-squared distribution with 2 dof, χ^2.
- Properties of this distribution:
  - The mean power is 2;
  - the standard deviation is 2!
- So, the power spectrum is very noisy. This does not improve with:
  - longer observation — you just get more powers
  - broader time bins — you just get a lower ν_{Ny}



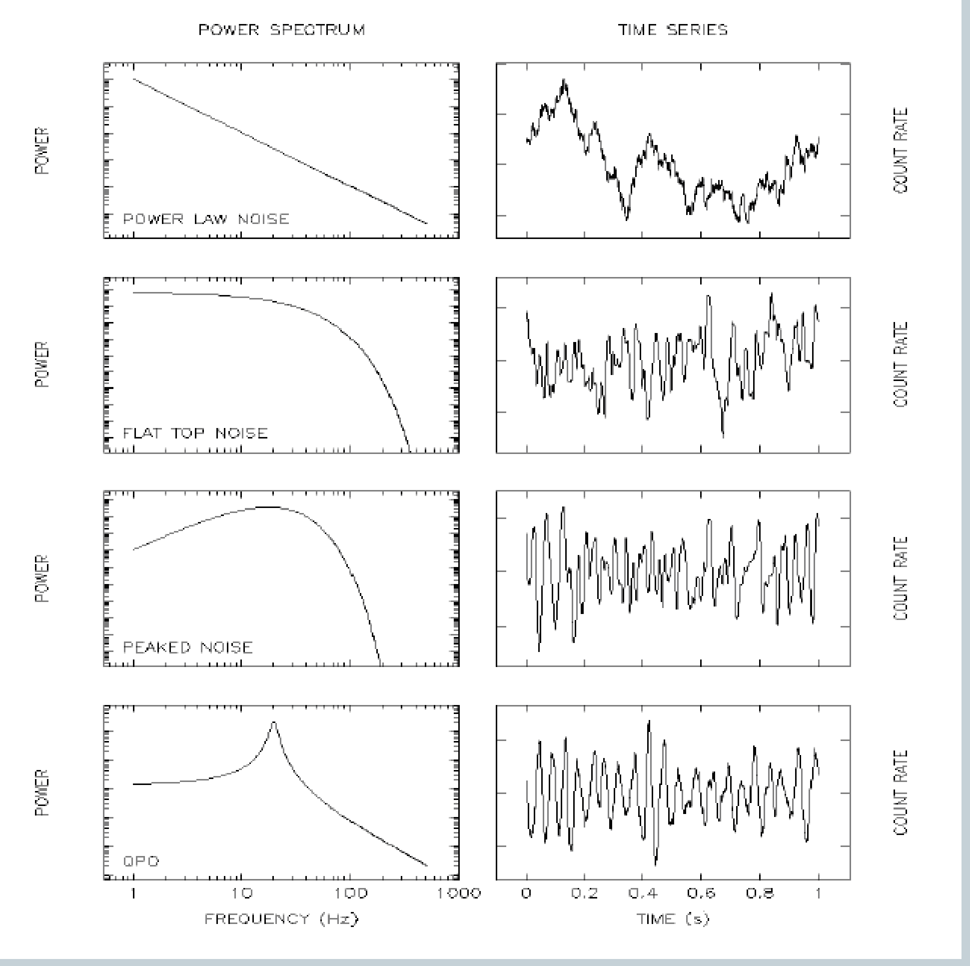
Statistics of Power Spectra

- Flux measurements are always accompanied by noise.
- The light curve can be divided into its independent components: the deterministic signal S and the noise N. For an individual time bin, the total number of counts is composed of the sum of the signal and the noise, i.e., $x_k = s_k + n_k$.
- Examples of deterministic signals:
  - a non-periodic deterministic variation, such as a nova light curve;
  - A periodic variation, such as an eclipsing binary or a RR Lyr light curve;
  - a multiply periodic variation, such as a spectroscopic triple system;
  - a modulated periodic variation where either the amplitude, frequency, or phase may vary with time - for example a pulsating system in a binary orbit.

Statistics of Power Spectra

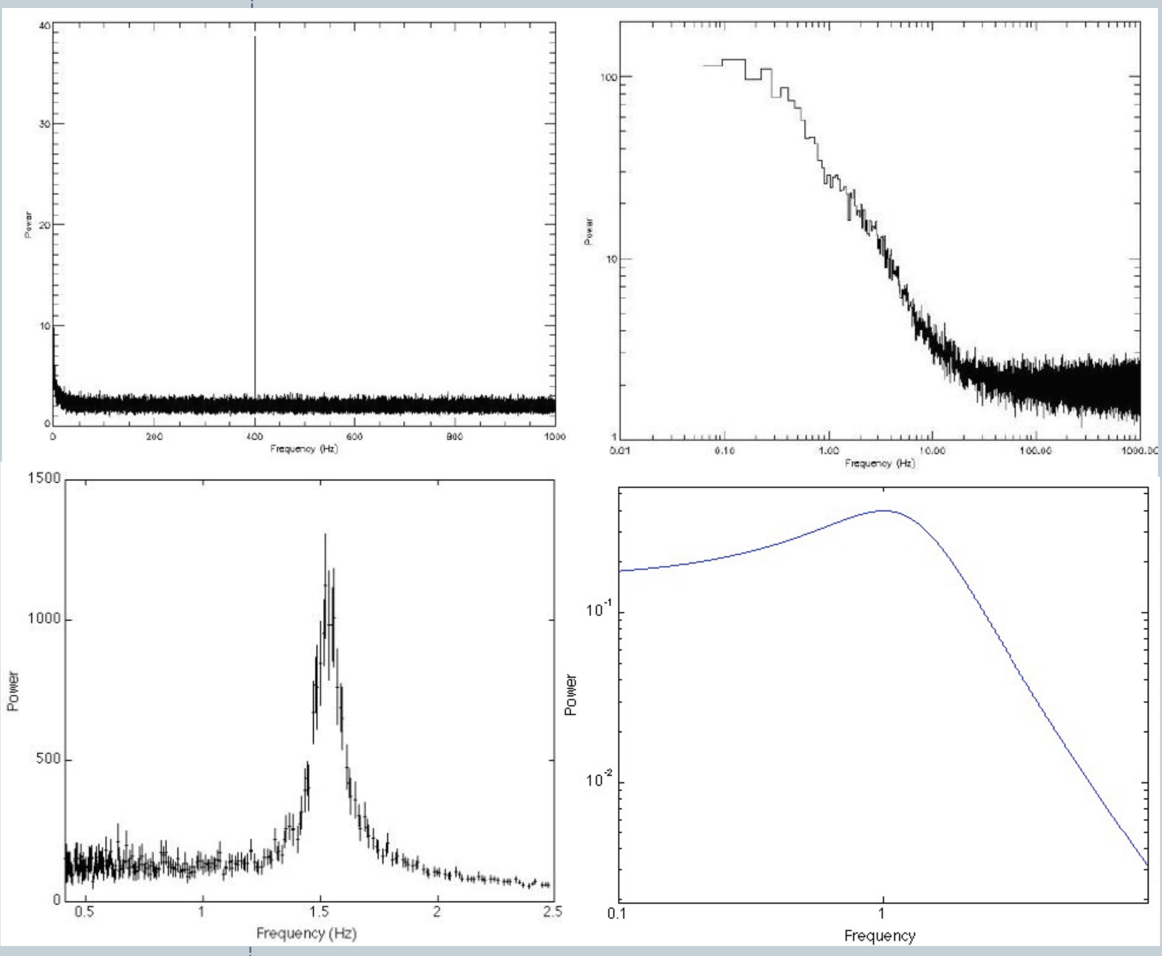
- 'Noise' (= random aka stochastic processes) in the light curve produces peaks and broad components in the power spectrum.
- Examples of noise:
  - Counting statistics noise (Poisson noise) -> white noise;
  - Poisson noise modified by instrumental effects (e.g. deadtime) and other instrumental noise;
  - Noise that is (stochastic) intrinsic source variability: QPO, band limited noise, red noise, etc.
- All these can occur at the same time, possibly together with deterministic signals.
- They can be the **background** against which you are trying to detect something else
- Or they can be the **signal** you are trying to detect.

Statistics of Power Spectra



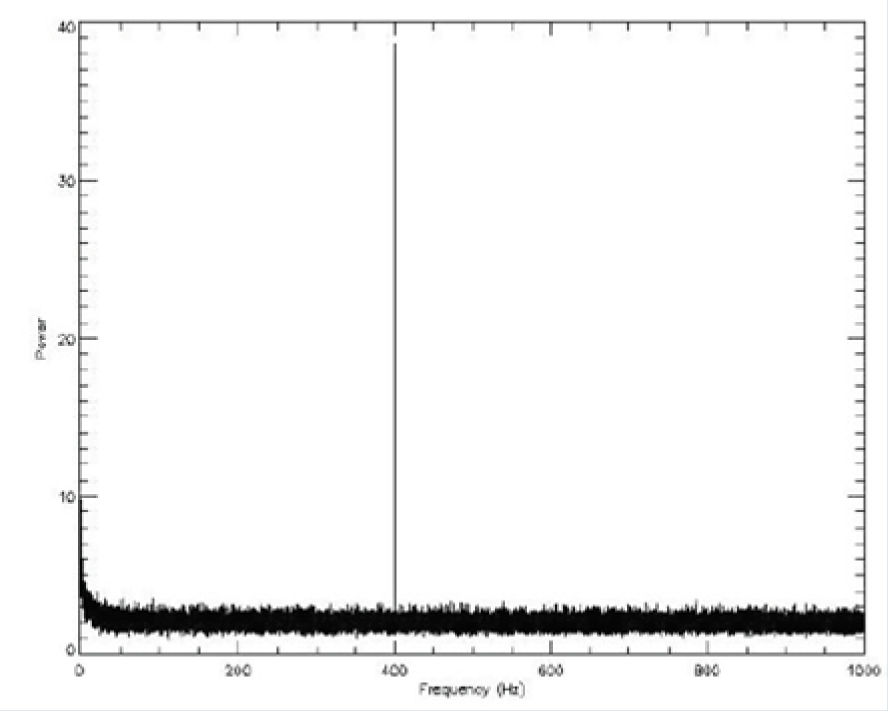
Main types of signals

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- Peaked-noise



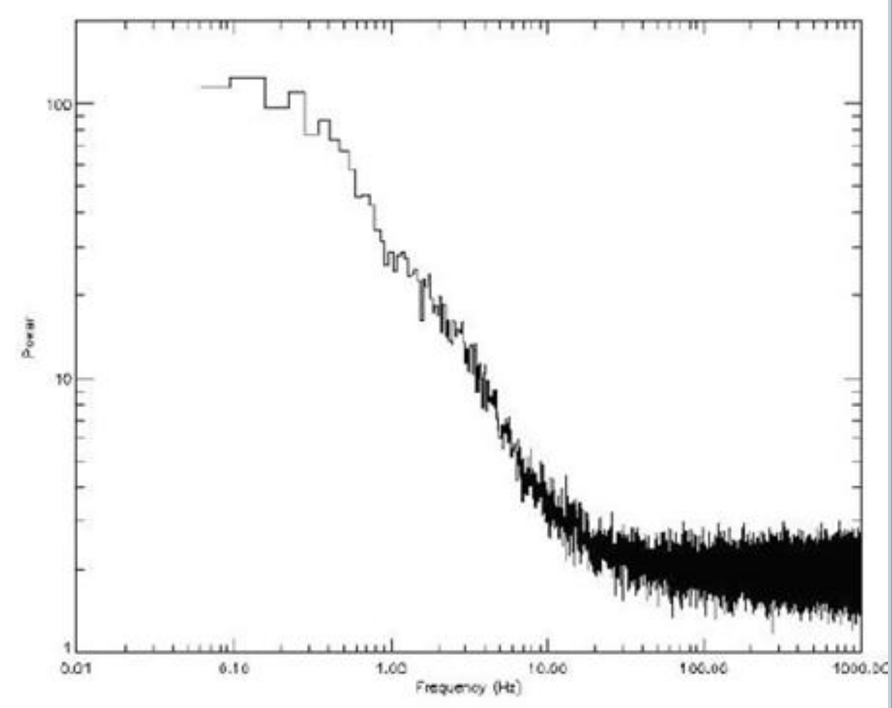
Main types of signals

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- Peaked-noise



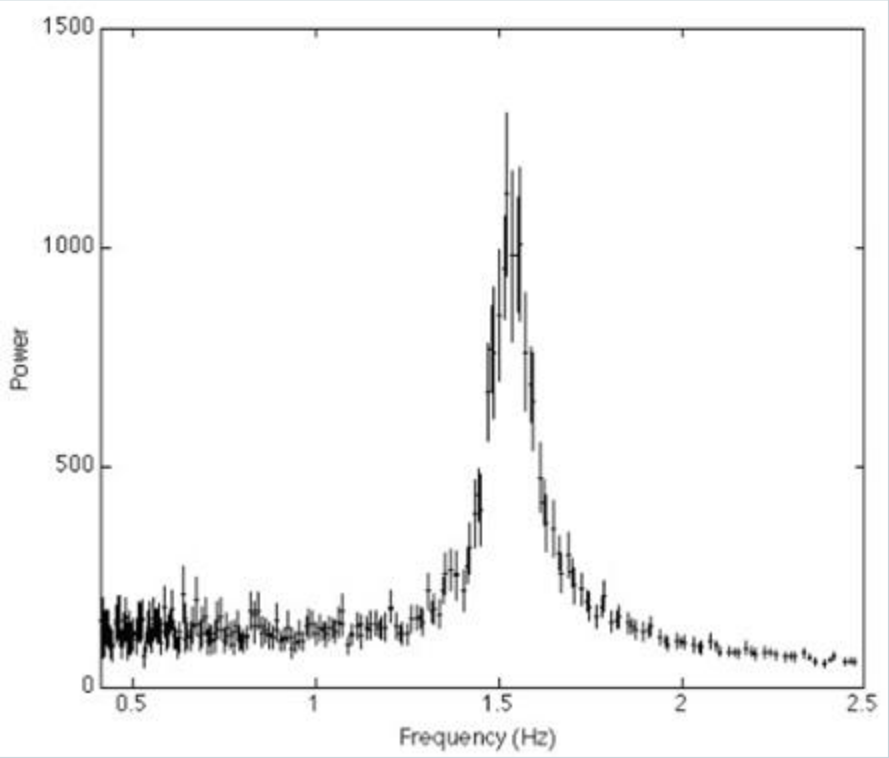
Main types of signals

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- Peaked-noise



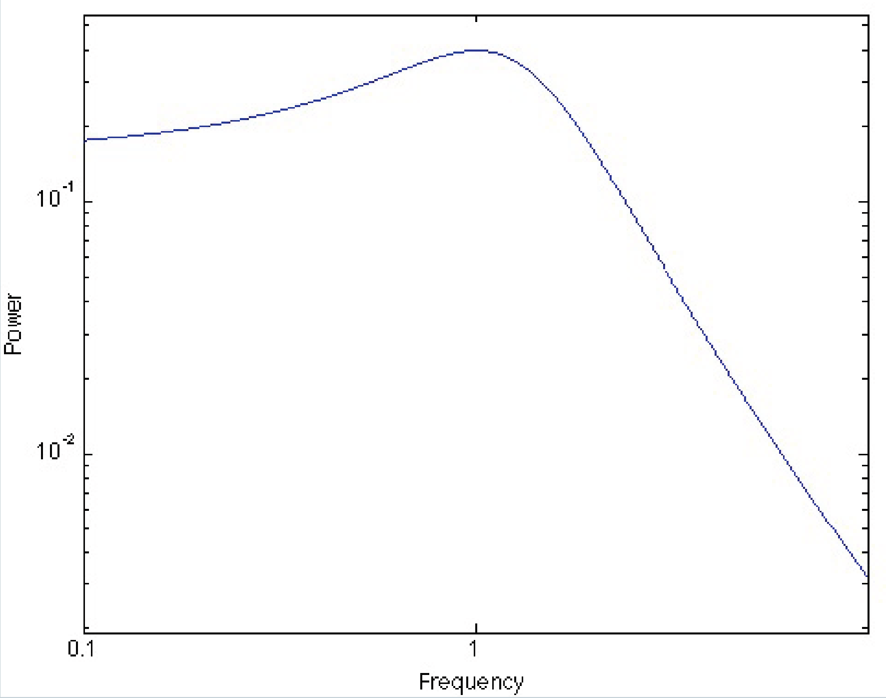
Main types of signals

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- Peaked-noise



Main types of signals

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- Peaked-noise

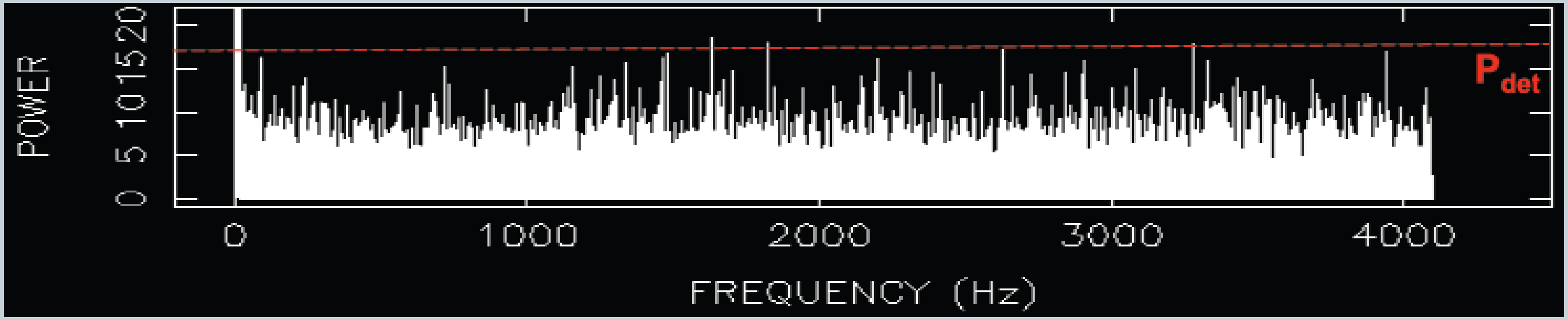


Coherent Signals

- Much analysis involves “coherent” signals, i.e. periodic signals whose phase is constant over the relevant duration
  - $Q = \nu/\Delta\nu >> 1000$
- Examples:
  - Pulses from rotating pulsars;
  - Orbital modulation or eclipses;
  - Precession periods.

Statistics of Power Spectra

- How to determine the significance of peaks found in power spectra? How big must a power be to constitute a significant excess over the noise?
- Let's define ε as the probability that a noise fluctuation exceeds P_{det}. The (1- ε) confidence detection level P_{det} is a level that has a false alarm probability of ε. If there is just noise, $\mathrm{Prob}(P_j > P_{det}) = \varepsilon$. We want ε to be small, e.g., ε =1% for 99% confidence.
- If $P_j > P_{det}$ then with 99% confidence there is something else than just noise, a source signal.



Statistics of Power Spectra

- To determine P_{det}, we need to know **the noise power distribution**.

- **Warning:** Because in High-Energy Astrophysics we are counting individual photons, the relevant statistics are *Poisson*, not *Gaussian*.
- The Leahy normalization is chosen such that if the x_k are Poisson distributed, then the P_j exactly follow the chi-squared distribution with 2 degrees of freedom, χ^2. This is actually an exponential distribution:

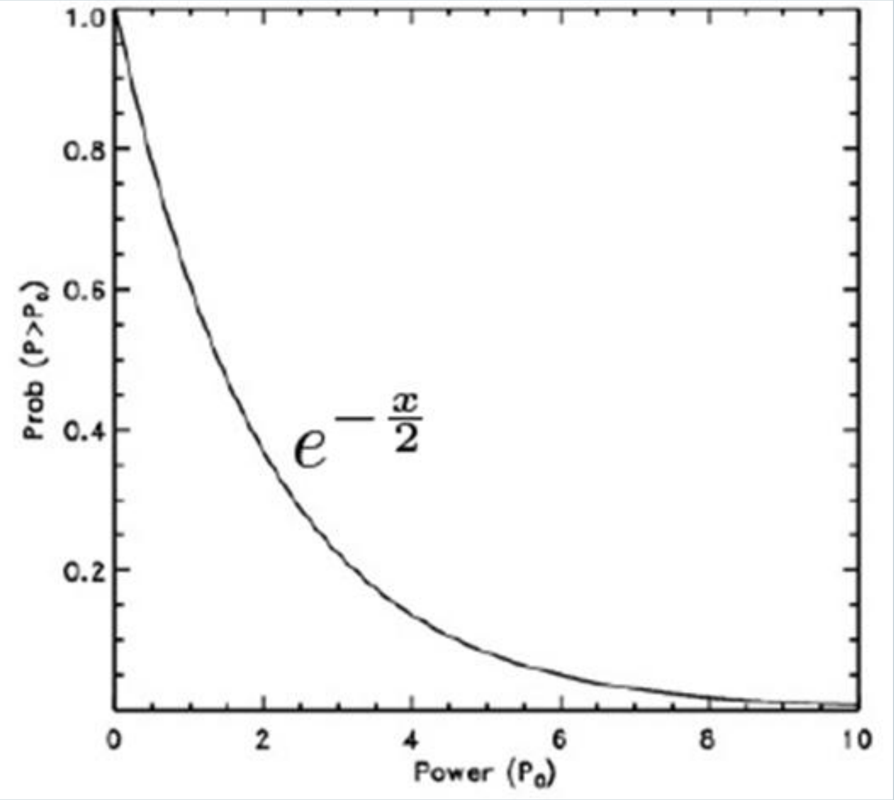
$$\varepsilon = Prob_{single}(P_j > P_{det}) = e^{-P_{det}/2} \longrightarrow P_{det} = -2\ln\varepsilon$$

- Properties of this distribution: $\langle P_{noise} \rangle = 2$; $Var(P_{noise}) = 4$

Statistics of Power Spectra

- Examples:
  - ε=1% corresponds to P_{det}=9.2;
  - a power of 50 has a probability of $e^{-40/2}=2\times10^{-9}$ of being noise.
- Since a large number of independent frequencies N_{trial} are examined, the detection threshold has to be defined as that power that has an ε (small) probability to be exceeded in one frequency bin out of the N_{trial} examined.
  - One should divide ε by the number of trials.

$$\varepsilon = N_{triale}{}^{-P_{det}/2}$$



Statistics of Power Spectra

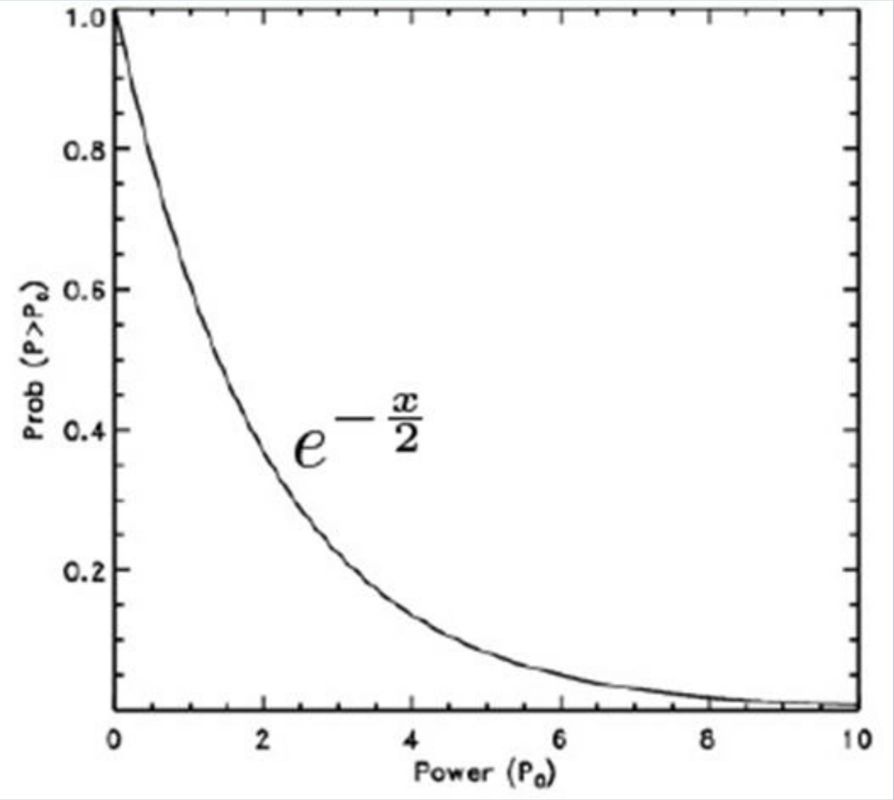
Important! The number of trial powers N_{trial} over which the search has been carried out:

N_{trial} = to the powers in the PSD if all the Fourier frequencies are considered;

N_{trial} < than the powers in the PSD if a smaller range of frequencies has been considered.

- Examples (cont.): N_{trial}=10 000
  - ε=1% corresponds to P_{det}=27.6;
  - a power of 40 has a probability of $e^{-40/2}=2\times10^{-5}$ of being noise.

Still significant!!



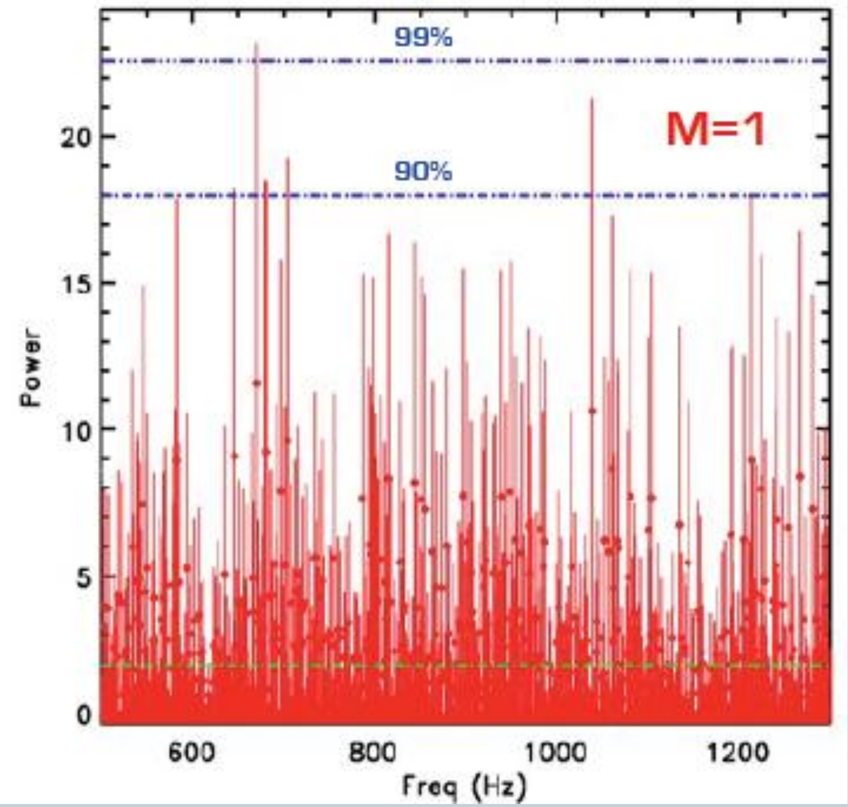
Rebinning and Averaging

- The power spectrum is very noisy. Smoothing methods:
  - Average several power spectra of subsegments of the time series;
  - Average adjacent bins in a power spectrum: rebinning;
  - Windowing is also possible.
- Averaged power distribution:
  - Individual P_j follow the chi-squared distribution with 2 dof.
  - Additive property of χ^2 distribution: sum of M powers is distributed as χ^2_{2M}
- W – the number of the time series, N – Frequency rebinning factor:

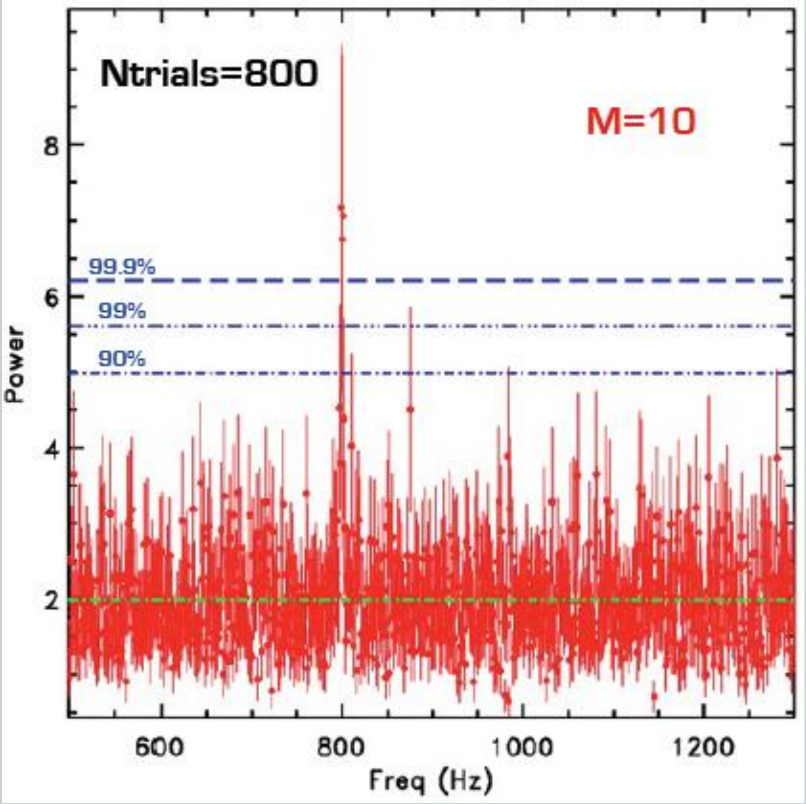
  $\langle P_{noise} \rangle = 2$; $Var(P_{noise}) = 4/MW$ (the number of trials decreases)
- **Central limit theorem: for large MW the distribution of $\overline{P_{WM}}$ tends to normal (Gaussian), with mean 2 and standard deviation $2/\sqrt{MW}$**

Signal Detection

M=1,
Noisy PDS



M=10,
A signal is clearly detected



A note about rebinning

- **Coherent peak:** narrow power distribution – the longer the observation span, the better. The signal power to decrease by 1/MW.
  Is it worth to average or rebin? No.
  - The signal power decreases faster than the threshold power when averaging/rebinning;
  - If the frequency varies (orbital motion) is even worse as you average signal with noise.

- **Broad peak:** broad power distribution - length of observation not crucial - rebinning helps.

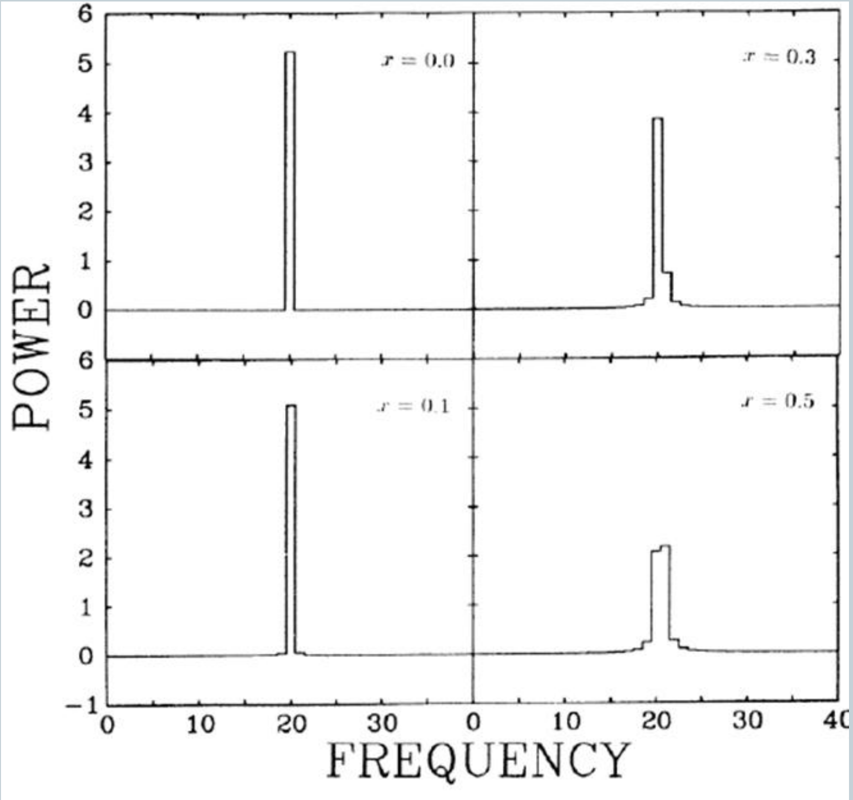
Signal detection optimization

- The power spectrum of a sinusoidal signal $x_k = A\cos(2\pi\nu_{sine}t_k + \varphi)$:

$$|a_j|^2 \approx \frac{1}{4}A^2N^2\left(\frac{\sin \pi x}{\pi x}\right)^2$$

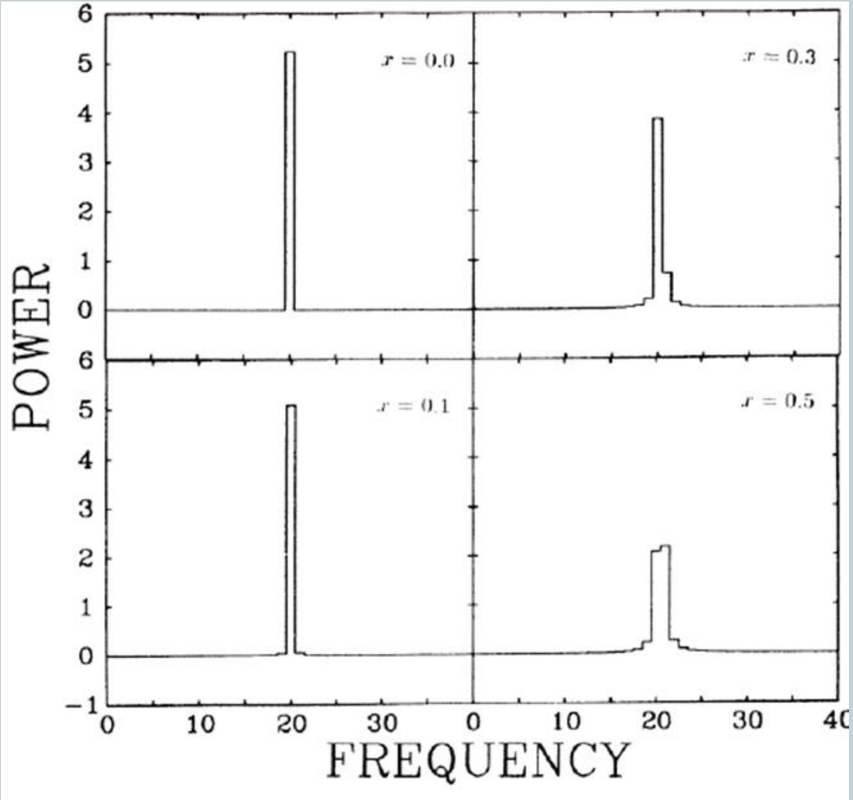
where $x = (\nu_{sine} - \nu_j)T$

- The highest power in the signal power spectrum will be obtained at the Fourier frequency ν_j closest to ν_{sine}. Normalized to a power of 1 for $\nu_{sine} = \nu_j$ (x = 0), this power varies between 0.405 and 1, with an average value of 0.773



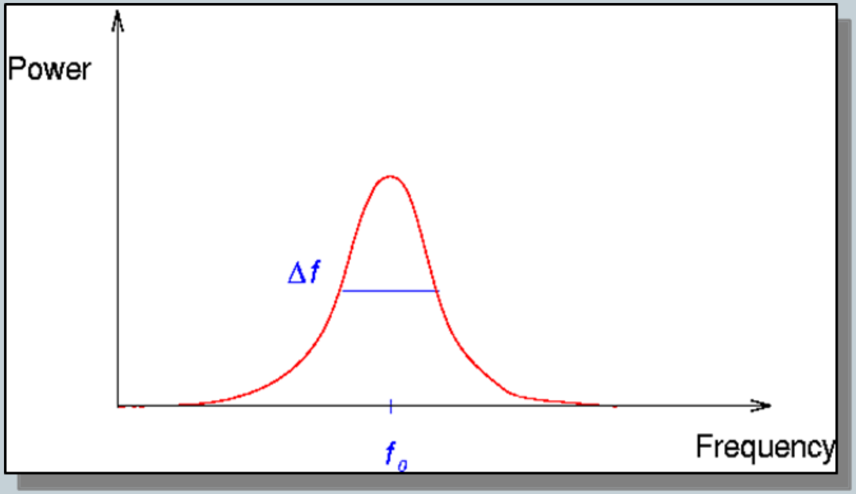
Signal detection optimization

- **Implications:** When searching for strictly coherent signals it is important to rely upon the original/maximum Fourier resolution (1/T).



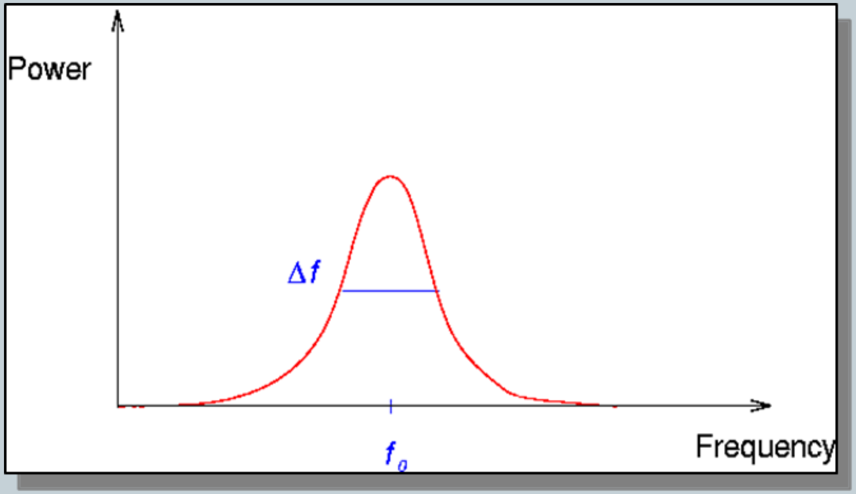
Signal detection optimization

- Similar reasoning shows that the signal power for a feature with finite width $\Delta\nu$ drops proportionally to $1/MW$ when degrading the Fourier resolution. However, as long as feature width exceeds the frequency resolution, $\Delta\nu > MW/T$, the signal power in each Fourier frequency within the feature remains approx. constant. When $\Delta\nu < MW/T$ the signal power begins to drop.



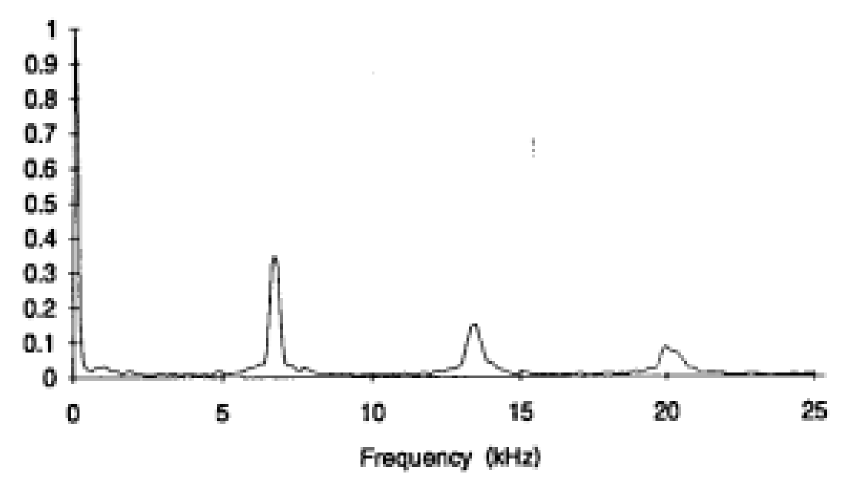
Signal detection optimization

- **Implications:** The search for QPOs is a three step interactive process.
- Firstly, estimate (roughly) the feature width.
- Secondly, run again a PSD by setting the optimal value of MW equal to ~$\Delta\nu$ T. Two or three iterations are likely needed.
- Finally, use χ^2 hypothesis testing to derive significance of the feature, its centroid and r.m.s.



Periodic Non-sinusoidal Signals

- Power for Periodic Nonsinusoidal Signals is spread over harmonics of the modulation frequency: Confidence lower.



Summary: Detecting something in a power spectrum

The process of detecting something in a power spectrum against the background of noise has several steps:

- knowledge of the probability distribution of the noise powers;
- knowledge of the interaction between the noise and the signal powers (determination of the signal upper limit);
- The detection level: Number of trials (frequencies and/or sample);
- Specific issues related to the intrinsic source variability (non Poissonian noise);
- Specific issues related to a given instrument/satellite (spurious signals – spacecraft orbit, wobble motion, large data gaps, etc.).