Thurs, 2-28-19

Smoothed Periodogram (cont.)

1. Data window, padding, tapering
2. Null continuum and significance of peaks
3. Confidence intervals with multiple tests
4. Demo06a→AR null continuum
5. Sample run of geosa6

Assignment a6: due Tues after break

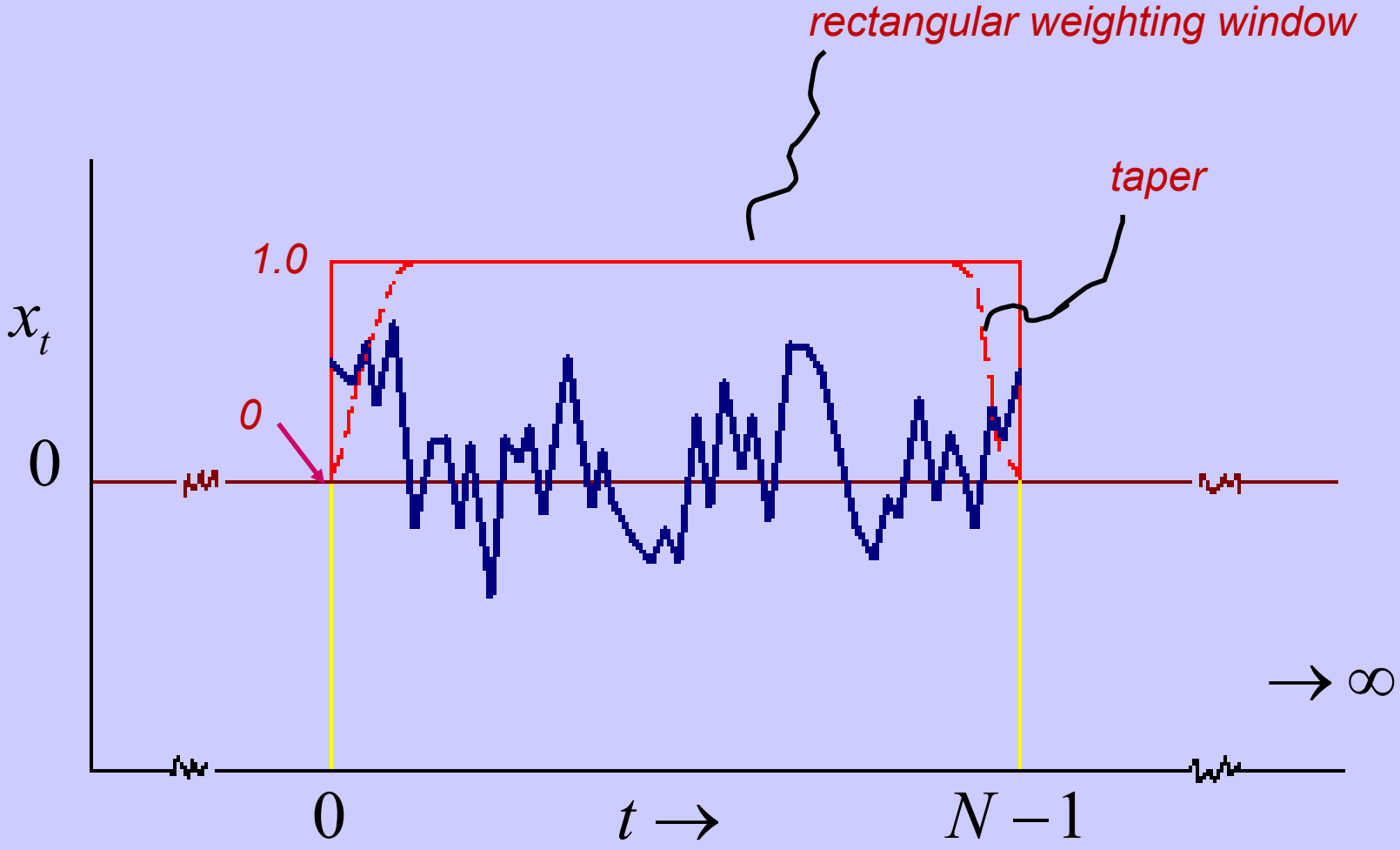
Data Window

1. In using the DFT to estimate the spectrum, the observed time series is viewed as a finite-length snapshot of an infinite-length series
2. You see part of the unknown infinite series through a “data window”
3. Characteristics this window are important to the quality of the spectrum

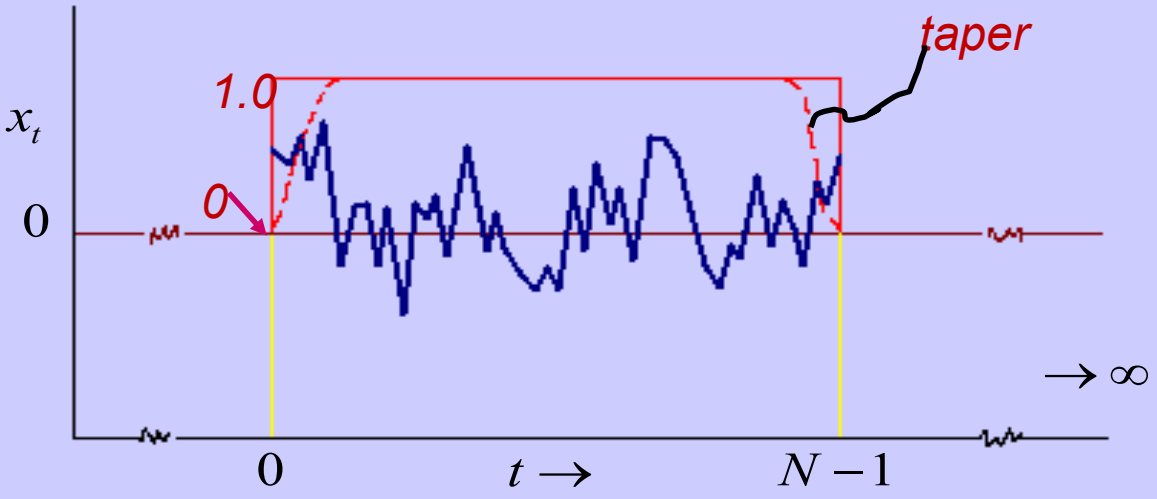
Data Window, cont.

$x_t, t = 0, N-1$ observed time series

x_t times weights is windowed, tapered data

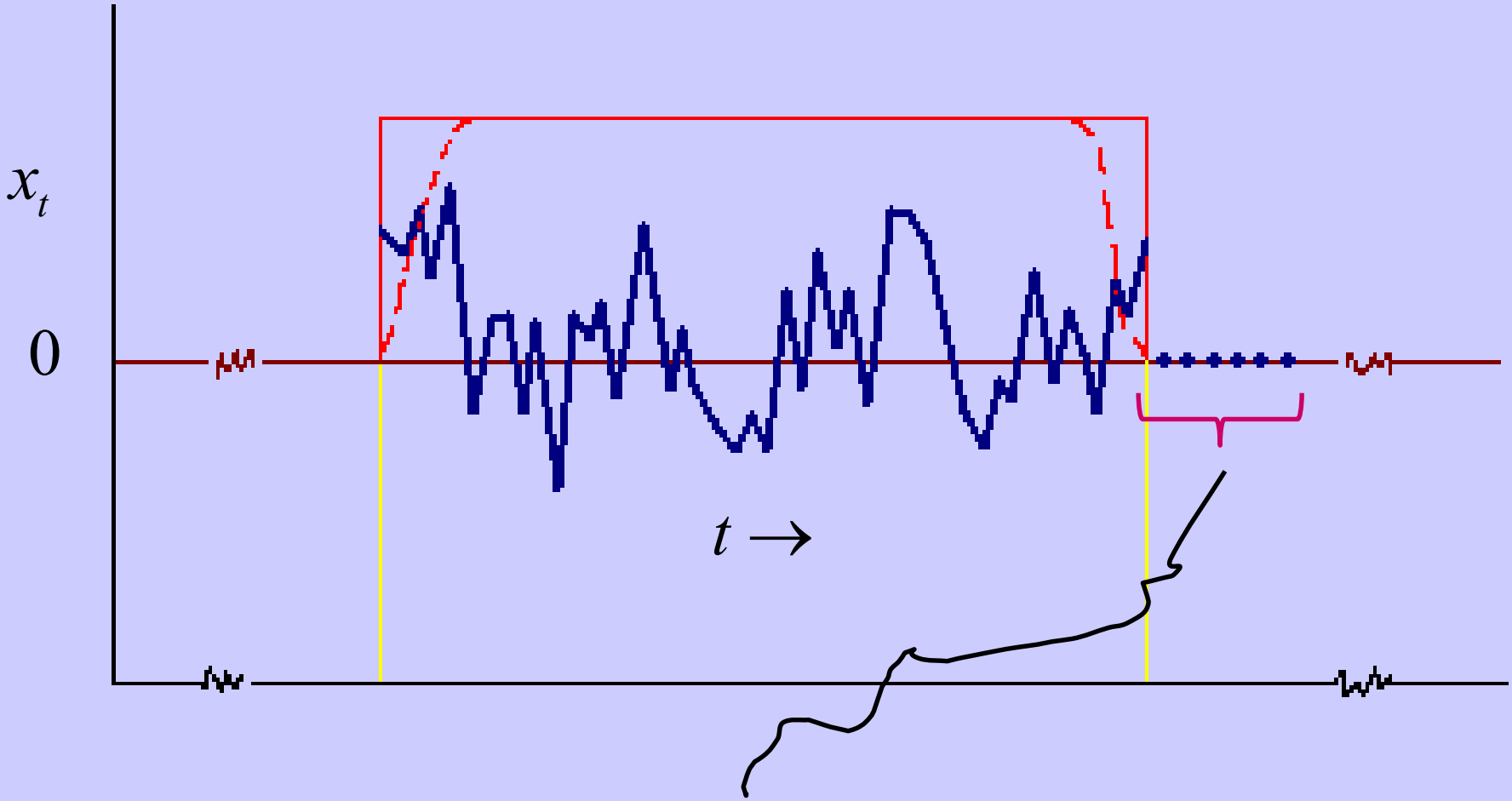


Tapering



- Typically applied to small percentage (e.g., 5%) on each end
- Purpose is to reduce distortion in the spectrum by leakage of variance from non-Fourier frequencies to Fourier frequencies when computing DFT

Padding



After tapering, series is padded to a length equal to the next power of 2 higher than the original sample length by tacking on zeros. Purpose is to facilitate FFT, or fast Fourier transform

Padding -- consequences

- Can use FFT for the DFT (computational convenience)
- Fourier, or standard, frequencies have finer spacing than the original (1/N' <1/N)

N' is padded length

For example, if original time series length is N=370 years, padded length is 512 years, and standard frequencies and wavelengths are now

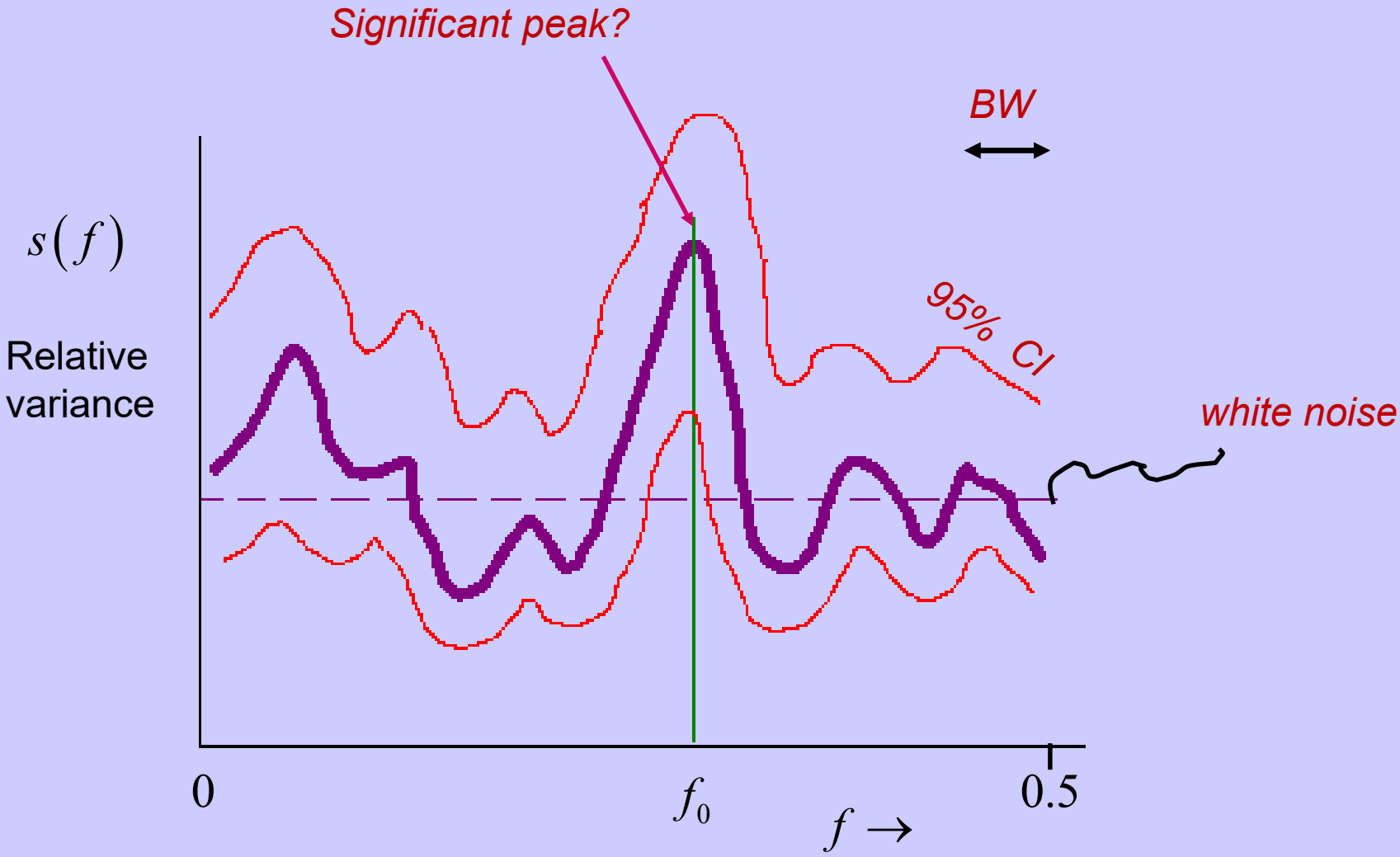
$$f_1 = 1/512, \quad \lambda_1 = 512\,yr$$
$$f_2 = 2/512, \quad \lambda_1 = 256\,yr$$
$$\vdots$$
$$f_{N/2} = 1/2, \quad \lambda_1 = 2\,yr$$

One consequence is a finer spacing of frequencies in the raw periodogram (1/512 instead of 1/370)

Null continuum

- Baseline for judging significance of peaks
- Question is does the spectrum have more variance at some particular frequency than some assumed comparison spectrum
- This assumed comparison spectrum is the null continuum
- The null continuum is usually white noise or red noise

White noise is simplest null continuum



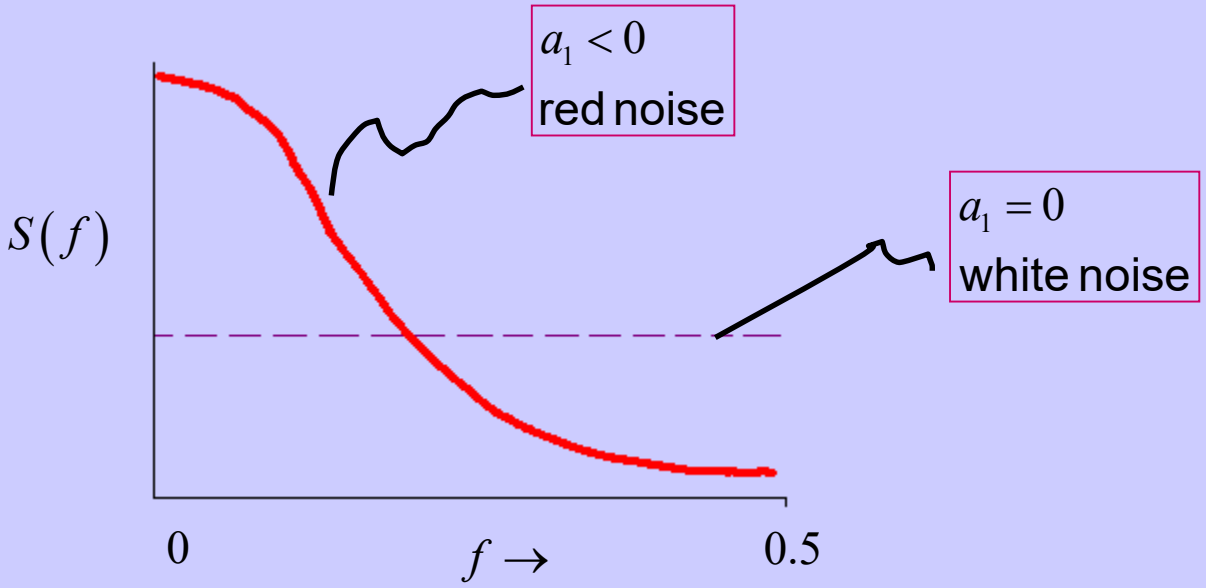
Red noise null continuum

- White noise is not appropriate for testing significance of spectral peaks for a dynamc time series with short-term memory
- Simplest alternative is significance relative to the spectrum of an AR(1) process with same variance as the time series – red noise
- Know that 1) spectrum is Fourier transform of acvf, and 2) AR(1) process has a theoretical acvf (e.g., geometric decay from lag 1)
- Theoretical spectrum of red noise is Fourier transform of the acvf of AR(1) process …

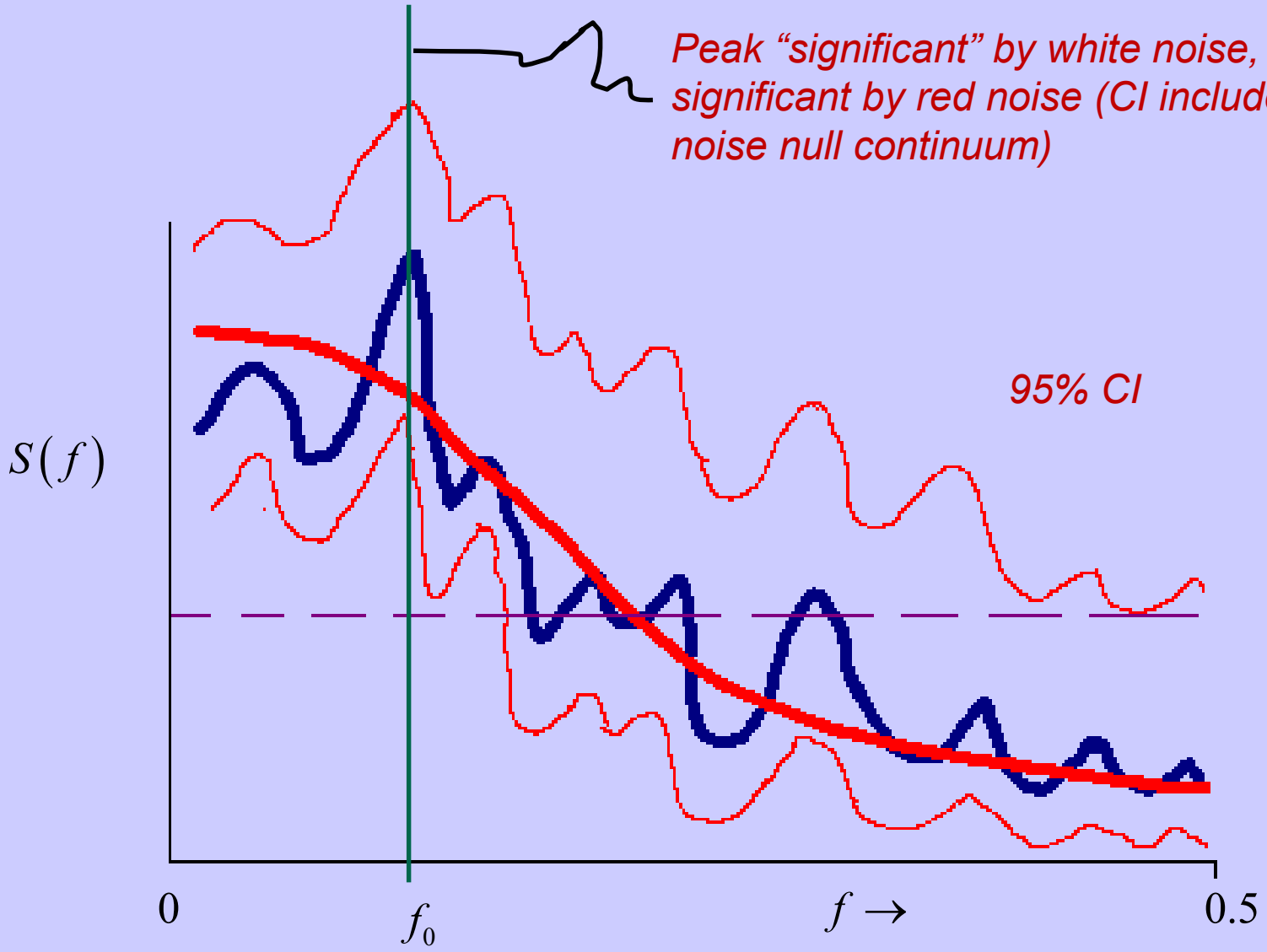
Red noise null continuum

AR(1): $y_t + a_1 y_{t-1} = e_t$

$$S(f) = \frac{4\sigma_e^2 / N}{1 + a_1^2 + 2a_1 \cos(2\pi f)}, \quad 0 \le f \le 0.5$$



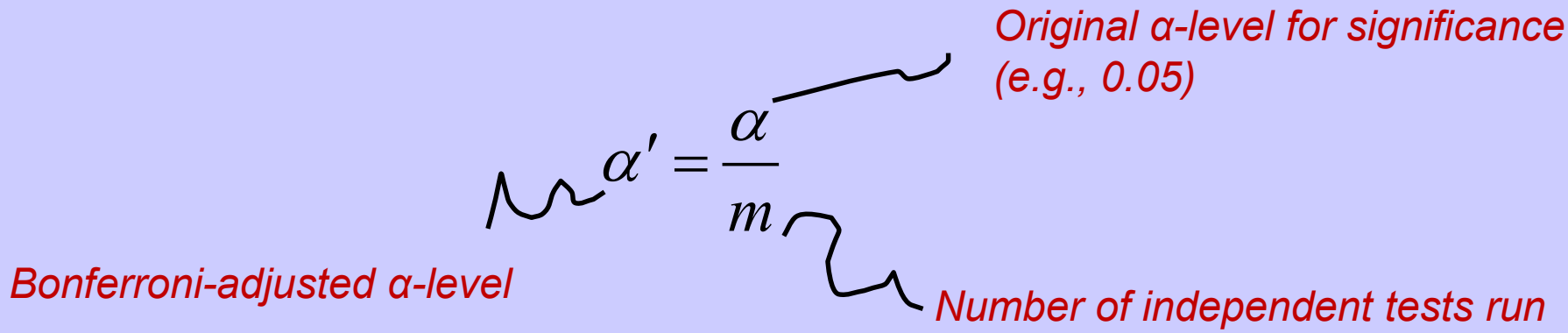
Choice of continuum can determine “significance”



Caveat:
confidence intervals for fishing expeditions

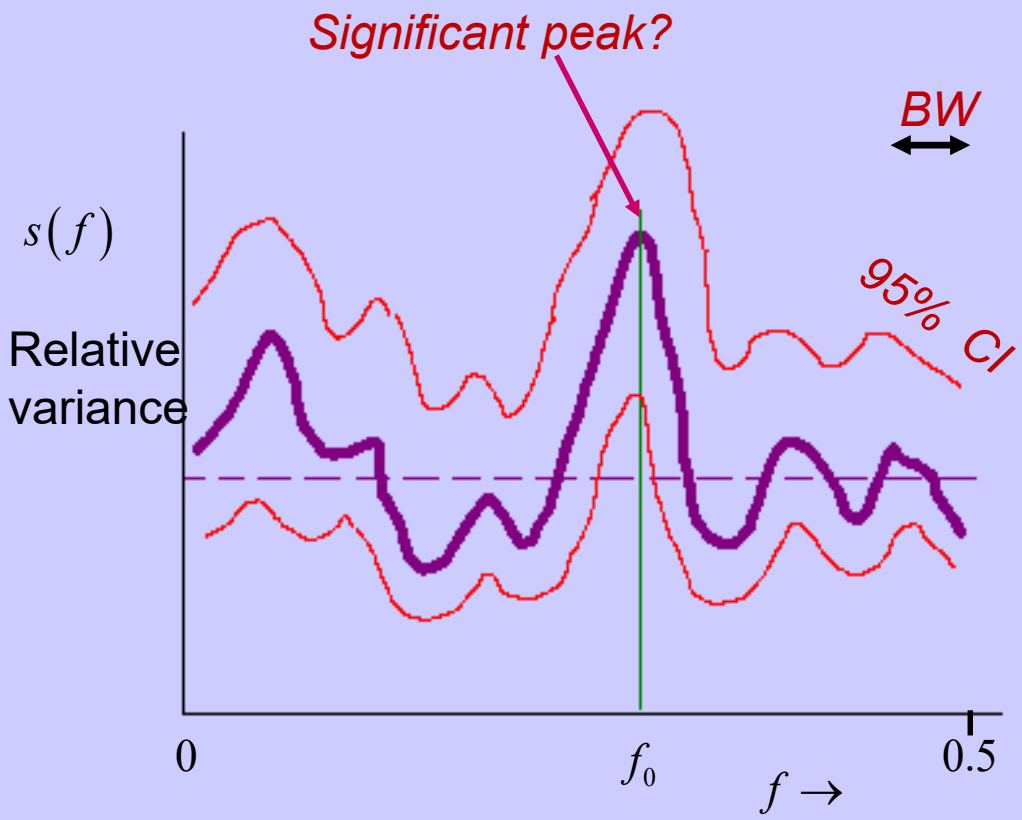
- The chi-squared-based confidence intervals for the spectrum described so far are not “simultaneous”: they are not designed for multiple simultaneous testing of peaks
- Legitimate use is to test the significance of a peak at a frequency of a-prior interest
- A fishing expedition consists of 1) spectral analysis with no particular frequency of prior interest, followed by 2) testing whatever peaks happen to show up for significance

Bonferroni-adjusted confidence interval



- *For the estimated spectrum, the number of independent tests run can be considered the number of spectral estimates independent from one another.*
- *This number is approximate equal to the number of times the 0 to 0.5 frequency axis is divisible by the bandwidth of spectral estimate*
- *Can illustrate with previous sketch…*

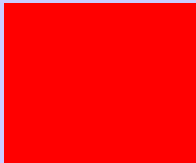
Peak still significant?



- *Bandwidth is only about 1/10 the length of the frequency axis (m=10)*
- *Should use α'=0.005 instead of α=0.05 to compute the 95% confidence interval*
- *This change would widen the CI, possibly removing significance of peaks*

Demo06a—theoretical AR spectrum

1. Generate and plot theoretical AR(1) and AR(2) spectra
2. Show sensitivity of shape of these possible “null continua” to AR order and parameter setting



Trial runs of geosa6…