Statistics 5350/7110
Forecasting

Lecture 23
Frequency Domain Diagnostics

Professor Stine

Preliminaries

- Questions?
- Assignments
  - Review of Assignment 5
  - Assignment 6
- Longer-than-usual review
  - Periodicity in time series
  - Periodogram
  - Start with a review of properties of the periodogram...

Today's Topics

Text, §6.2, §7.1

- Periodogram
  - Review of sampling properties
- Diagnostics based on the periodogram
  - An isolated frequency (e.g. seasonality)
  - Broader dependence
- Smoothing the periodogram
  - Scatterplot smoothing originated in periodogram smoothing
  - Connection to covariances
  - Spectrum of stationary process

Review of Periodogram

Key concepts from prior lecture

Properties of Periodogram

- Regression coefficients
  - Coefficients at Fourier frequencies

$$a_k = (2/n)\sum X_t \cos(2\pi tk/n) \quad \text{and} \quad b_k = (2/n)\sum X_t \sin(2\pi tk/n)$$

  - Approximately normal, independent

$$a_k, b_k \sim N\left(0, \sigma_w^2 \frac{2}{n}\right) \qquad \text{k = 1, 2, ..., (n/2)-1}$$

$$\text{Cov}(a_k, a_j) = 0 \ (j \neq k) \quad \text{and} \quad \text{Cov}(a_k, b_j) = 0$$

Appendix Property C.3 page 234

- Periodogram
  - Squared amplitude of estimated signusoid

$$X_t = R\cos(2\pi\, t\, k/n + \varphi) + w_t \quad \text{-->} \quad R_k^2 = a_k^2 + b_k^2$$

Caution: this is not the R^2 from regression!

  - Periodogram

$$I(k/n) = \frac{n}{4}\left(a_k^2 + b_k^2\right) \approx \sigma_w^2\,\frac{\chi_2^2}{2}$$

Proportional to an exponential random variable

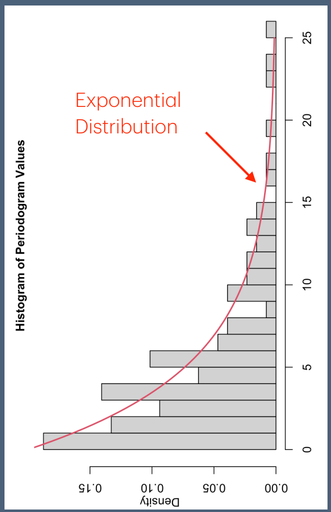
Properties of Periodogram

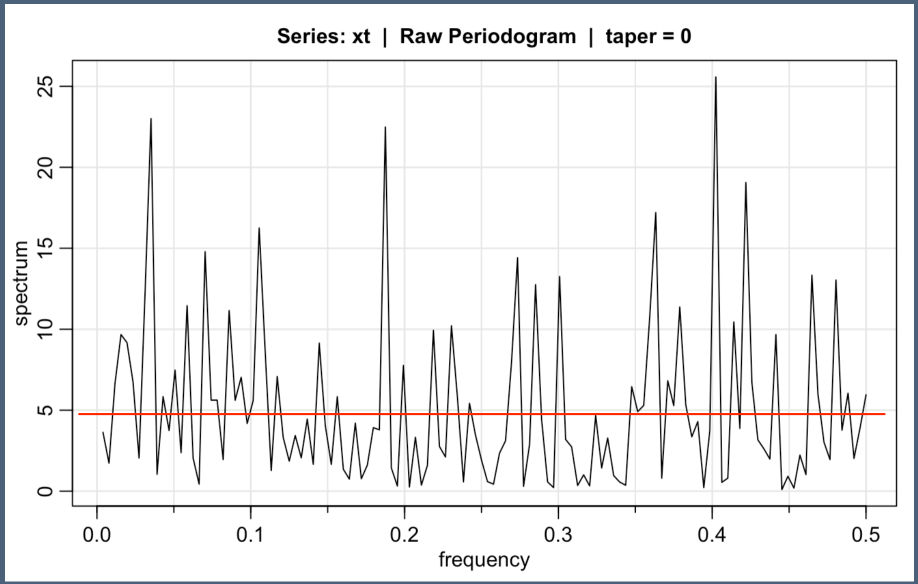
- Periodogram of white noise
  - Proportional to exponential r.v.

$$I_k = I(k/n) = \frac{n}{4}\left(a_k^2 + b_k^2\right) \sim \sigma_w^2 \frac{\chi_2^2}{2}$$

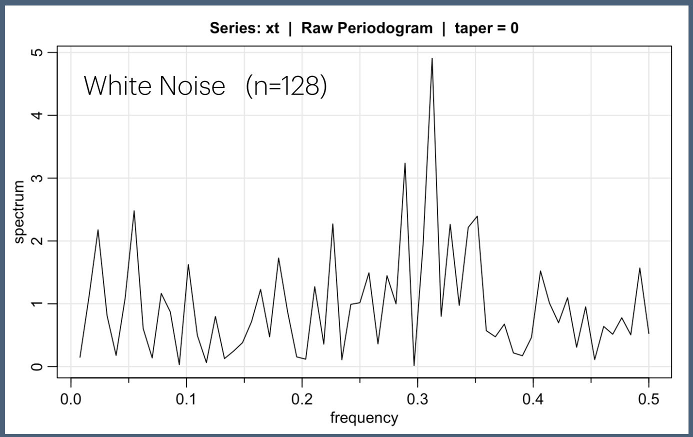
  - Approximately independent values sampled from skewed distribution: Very noisy appearance

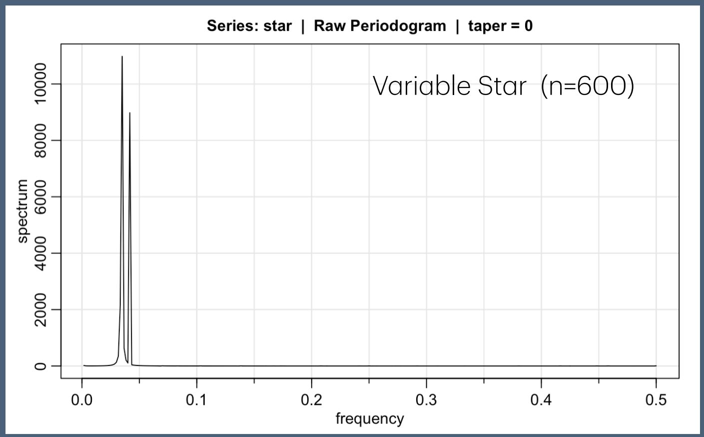
- Example
  - n = 256
  - White noise, $\sigma_w^2 = 5$

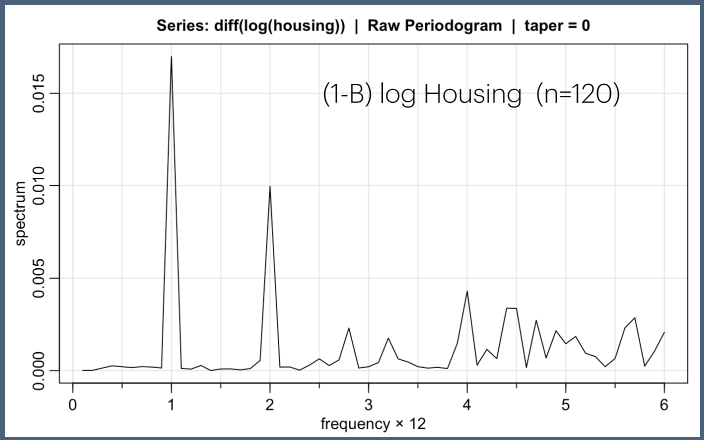


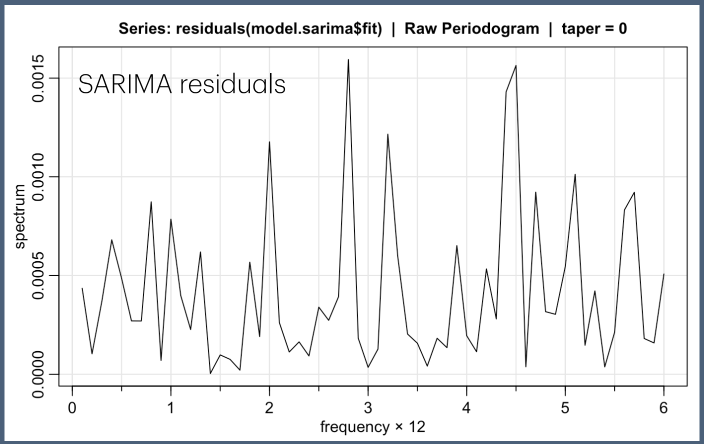


Examples of Periodogram









Max of Periodogram

Distribution of Maximum Periodogram

- Idea
  - Originated 100+ years ago by R. A. Fisher
  - Exponential r.v. imply "nice" behavior of the maximum value
- Application to periodogram
  - Periodogram values I_k are approximately multiple of independent exponential r.v.

$$I_k \sim \sigma_w^2 \left(\frac{\chi_2^2}{2}\right) \quad \Rightarrow \quad P(I_k/\sigma_w^2 \le x) = 1 - e^{-x}$$

  - Let m denote the number of periodogram values considered (m = n/2 - 2 if n is even)
  - Maximum value of the m terms $M = \max I_k$

$$P(M \le x) = (1 - e^{-x})^m$$

  - Because $\lim_{n\to\infty} (1 + x/n)^n = e^x$ it follows that

$$P(M \le x + \log m) = (1 - e^{-x-\log m})^m = (1 - e^{-x}/m)^m \approx \exp(-e^{-x})$$

- Refinement
  - Result has been improved over the years since Fisher (How to estimate σ_w^2)

R function "test_max_periodogram" defined in Lecture_23.Rmd

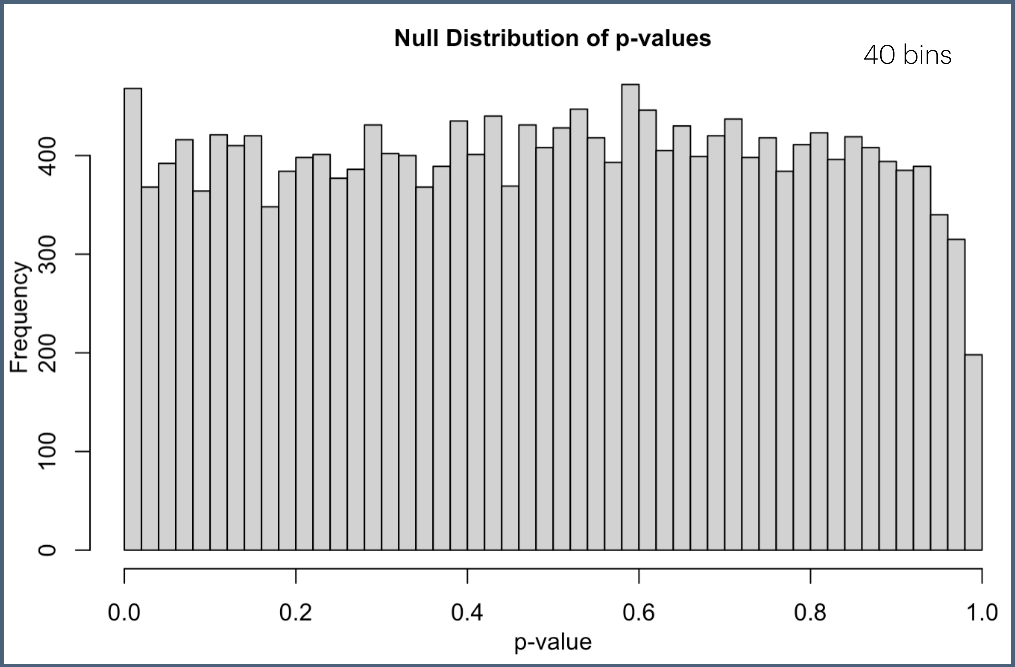
Example of Max Test

- Confirm level of the test
  - Test produces a p-value.
  - Before using, confirm that the test behaves properly
  - If the null hypothesis holds, then the distribution of p-values should be uniform [0, 1]
- Simulate p-values
  - Gaussian white noise with n = 128
  - 20,000 simulated replications
  - Issues:

    Too few close to 1 (who cares)

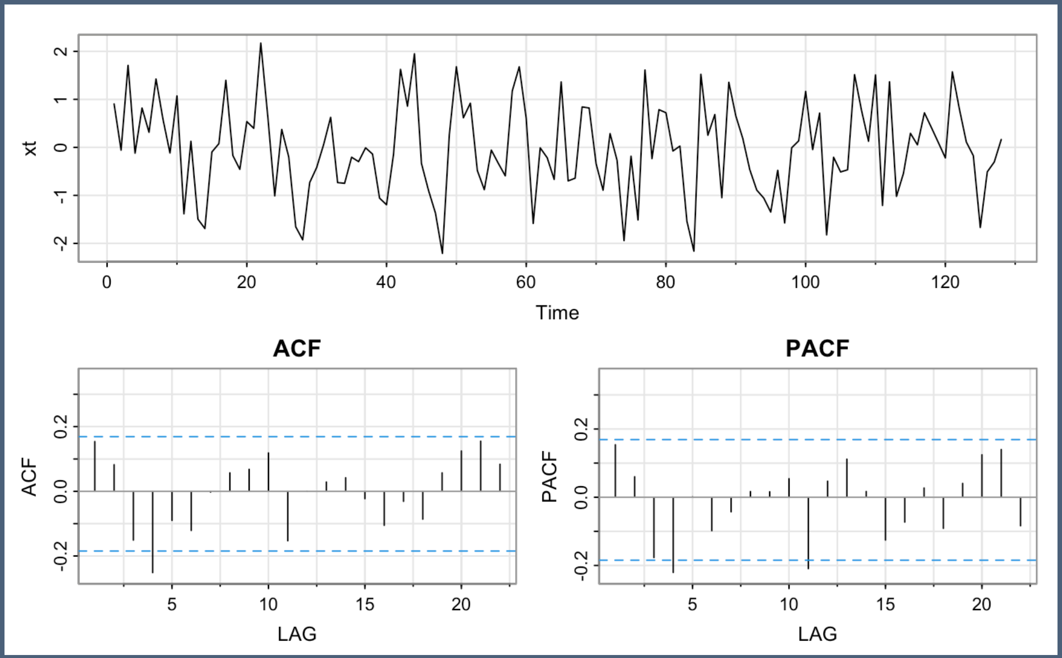
    Too many close to 0 (uh-oh)

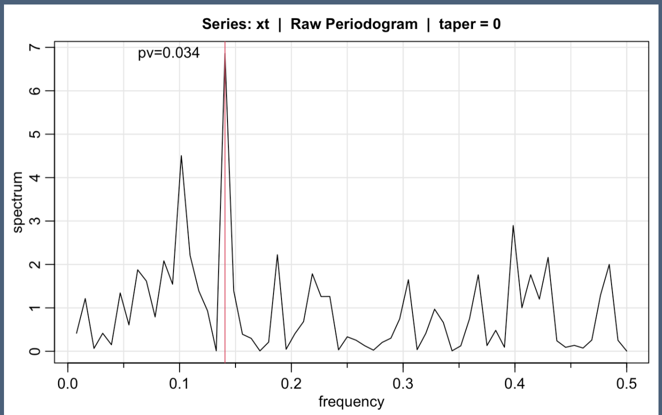
    Not a big problem: 5.1% less than .05
  - Not surprising since n=128 is a "small" length for this application



Example of Max Test

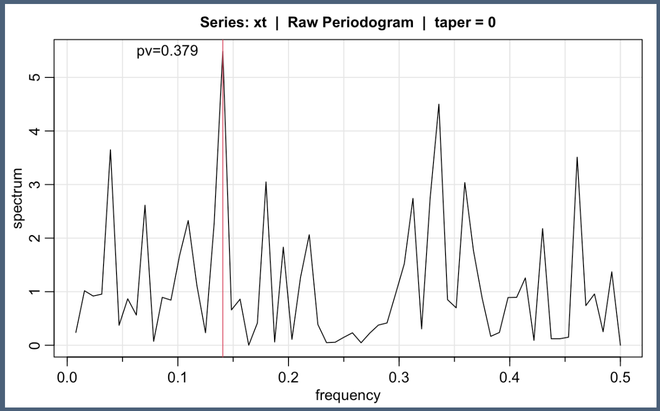
- Simulate application
  - Simulate data that does include a sinusoid.
  - What sort of power does the test have?
- Context
  - White noise with n = 128
  - $X_t = R\cos(2\pi tk/n) + w_t \qquad R = 0.35$
- What do you think?

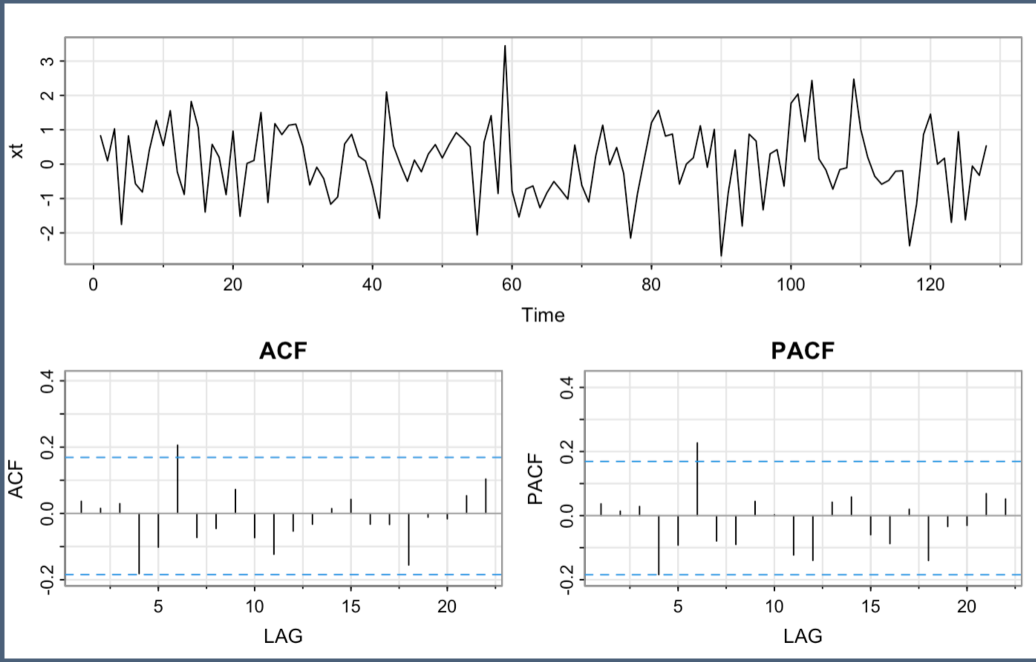




Example of Max Test

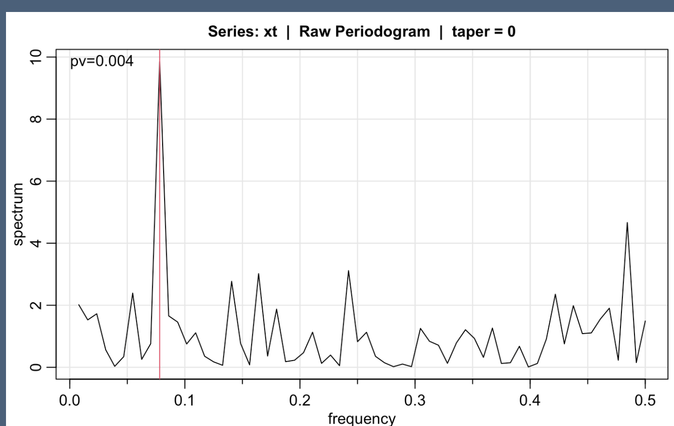
- Simulate application
  - Simulate data that does include a sinusoid.
  - What sort of power does the test have?
- Context
  - White noise with n = 128
  - $X_t = R\cos(2\pi tk/n) + w_t$
- What do you think?

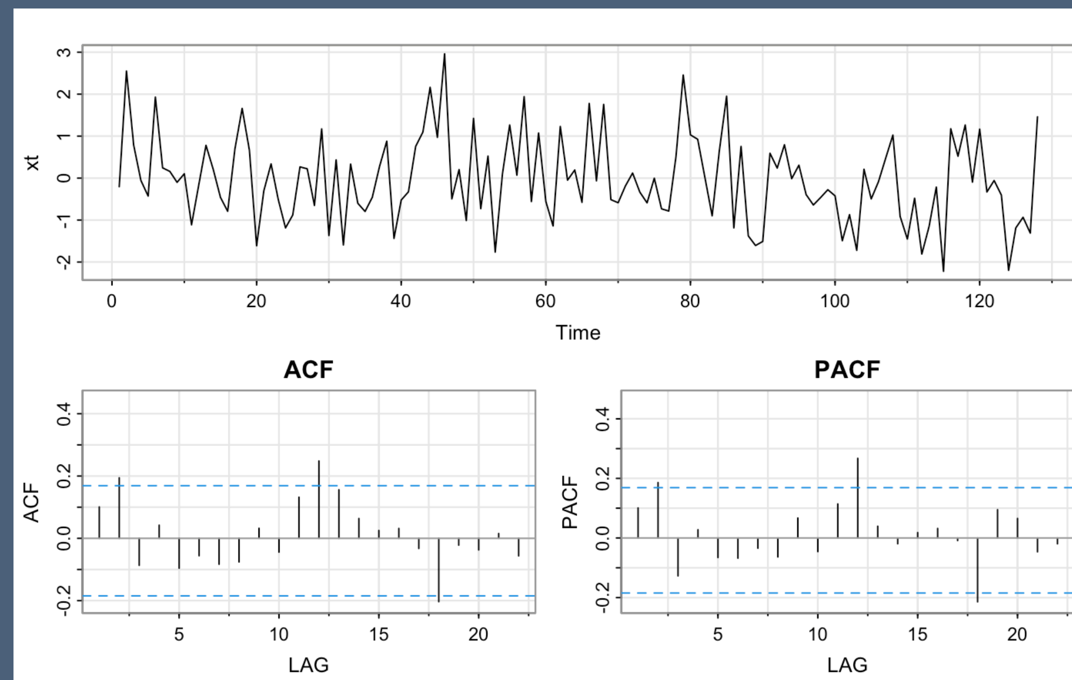




Example of Max Test

- Simulate application
  - Simulate data that does include a sinusoid.
  - What sort of power does the test have?
- Context
  - White noise with n = 128
  - $X_t = R\cos(2\pi tk/n) + w_t$
- What do you think?

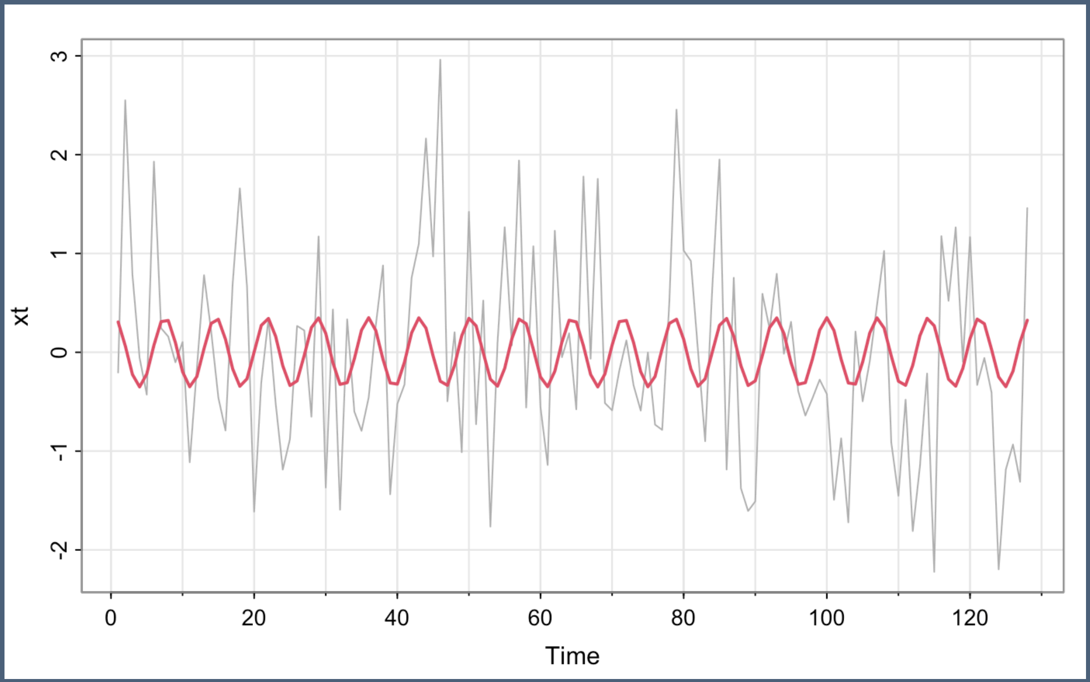




Signal Revealed

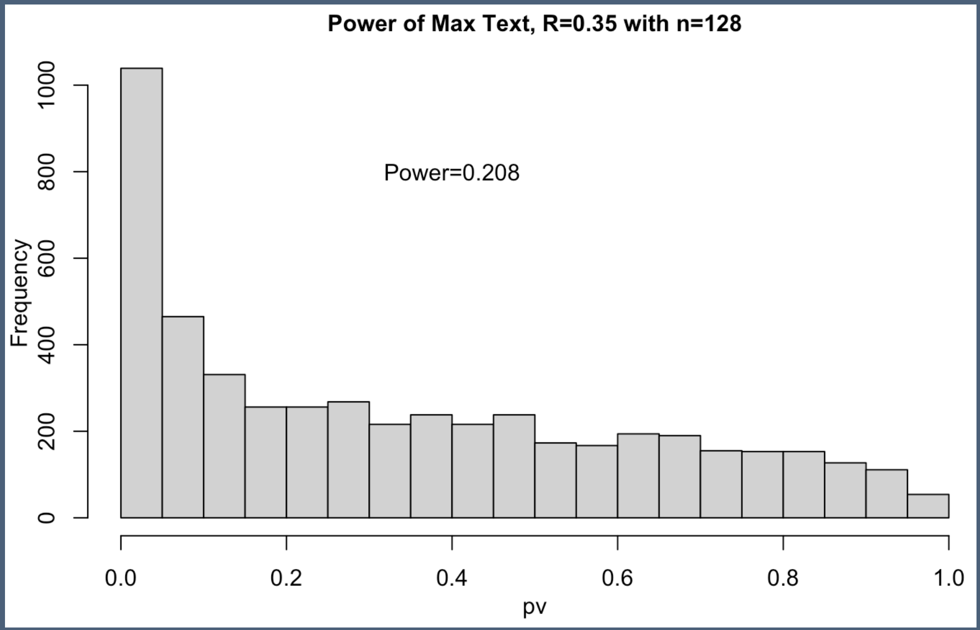
- Lower power?
  - Hard to see signal in noise
  - Hints in ACF

- Periodogram
  - Very evident in periodogram
  - It's the continuing regularity that makes this frequency so evident in the periodogram
  - Wave detectors at ocean



Power of Max Test

- Setting
  - n = 128
  - R = 0.35
- Simulate for Gaussian white noise

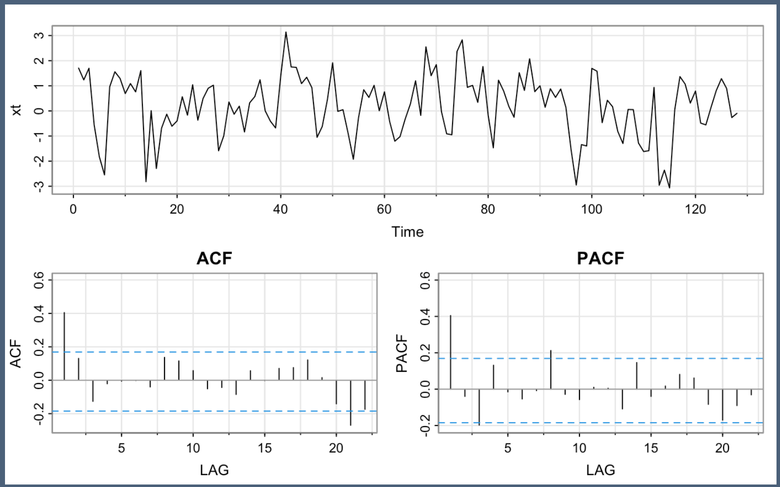


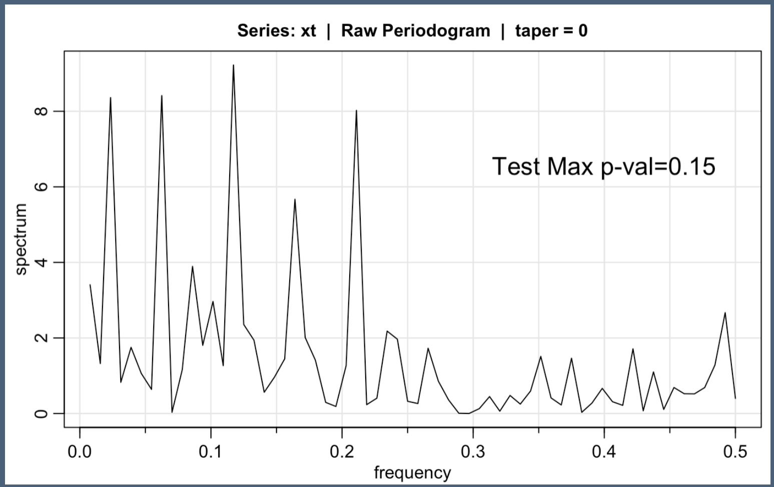
20 bins, so power of test that rejects if
$p < 0.05$
is fraction of p-values in the left-most bin.

Cumulative Periodogram

Alternative Test from Periodogram

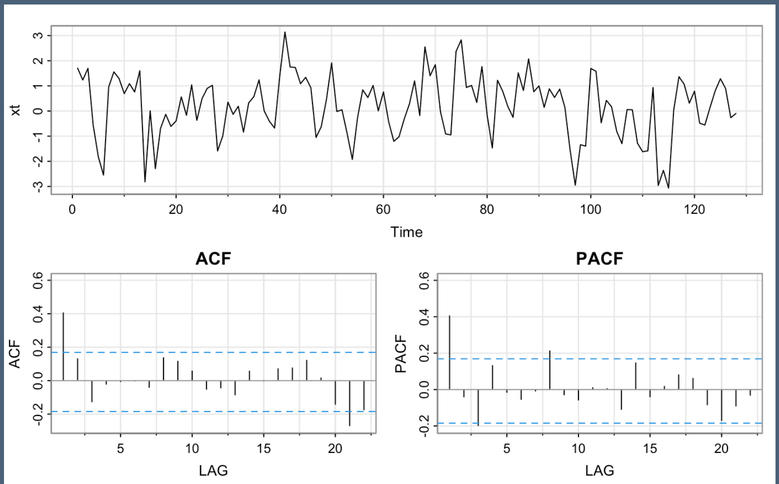
- Max test
  - Great for finding single sinusoid hidden in a time series, what about other questions
  - What if you want to test for other components, such as the second peak?
  - What if the deviation from white noise isn't confined to a single frequency?
- Example
  - Suppose data is an AR(1) with $\phi = 0.4$
  - Max test isn't very helpful here

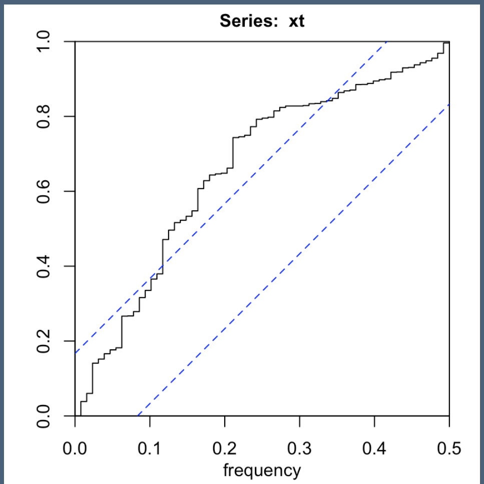




Alternative Test from Periodogram

- Max test
  - Great for finding single sinusoid hidden in a time series, what about other questions
  - What if you want to test for other components, such as the second peak?
  - What if the deviation from white noise isn't confined to a single frequency?
- Example
  - Suppose data is an AR(1) with $\phi = 0.4$
  - Max test isn't very helpful here, but inspecting the cumulative periodogram is



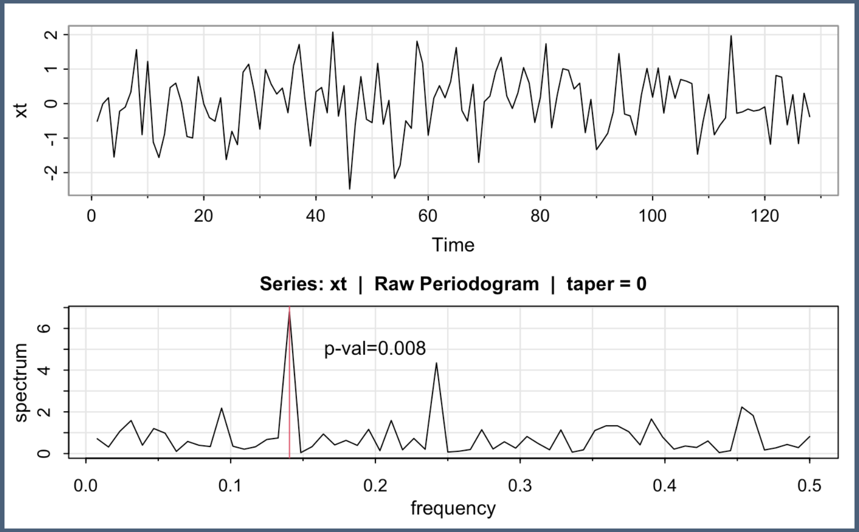


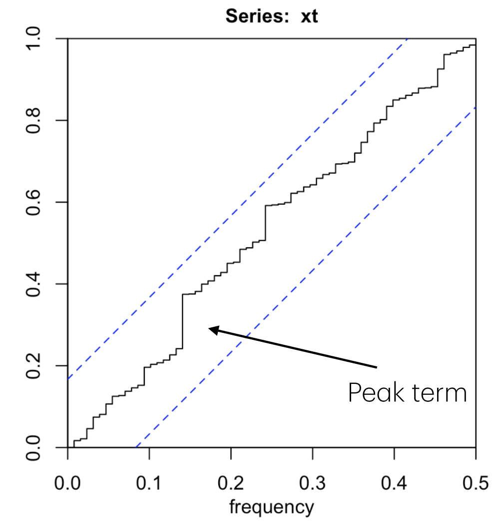
Test uses boundaries like those that define significant deviations from the diagonal in a QQ-plot

Comparison

- Each test has its domain
  - Power of tests based on the maximum or cumulative periodogram depend on deviation from H_0
  - Null hypothesis H_0: white noise (ideally, Gaussian white noise)

- Deviation: Hidden periodic component
  - Max test has more power
- Deviation: Broad shift
  - Cumulative periodogram has more power





The Spectrum

What's the nature of that "broad deviation" from white noise?

Spectral Density

- Alternative to the autocovariance function
  - The spectral density is the discrete Fourier transform of the autocovariances

$$f(\nu) = \gamma(0) + 2\sum_{h=1}^{\infty} \gamma(h)\ \cos(2\pi\nu h), \quad -1/2 \le \nu \le 1/2$$

  - The inverse relationship also holds,

$$\gamma(h) = \int_{-1/2}^{1/2} f(\nu)\ \cos(2\pi\nu h)\, d\nu, \qquad h = 0, \pm 1, \pm 2, \ldots$$

Property 6.6
without the
complex exponential

- Spectral density is ...
  - Equivalent to autocovariances
  - Symmetric around 0, f(u) = f(-u)
  - Offers a frequency-domain decomposition of the variance of the process

$$\mathrm{Var}(X_t) = \gamma(0) = \int f(\nu) d(\nu)$$

Could anticipate from
the periodogram

Spectral Density of ARMA

- White noise
  - Origin of the name "white noise"
  - If {wt} is white noise with variance σ_w^2, then

$$f_w(\nu) = \gamma(0) + 2 \sum \gamma(h) \cos(2\pi\nu h) = \gamma(0) = \sigma_w^2$$

Example 6.7

  - Implies equal variance associated with every frequency
- MA process
  - Reconstruct from the definition and autocovariances
  - If $X_t = w_t + \theta\, w_{t-1}$, then $\gamma(0) = \sigma_w^2(1+\theta^2)$, $\gamma(1) = \sigma_w^2\,\theta$, $\gamma(h) = 0$, $h=\pm2,\pm3,\ldots$

Example 6.9

$$f_x(\nu) = \sigma_w^2 \left(1 + \theta^2 + 2\theta \cos(2\pi\nu)\right)$$

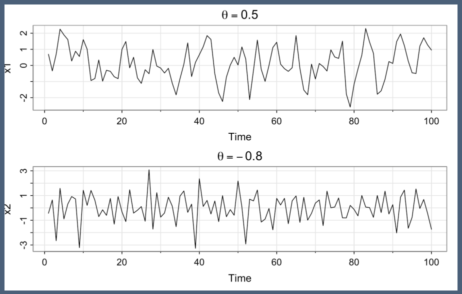
- AR process
  - Practical derivation requires knowing a bit of complex variables
  - For AR(1) with coefficient ϕ,

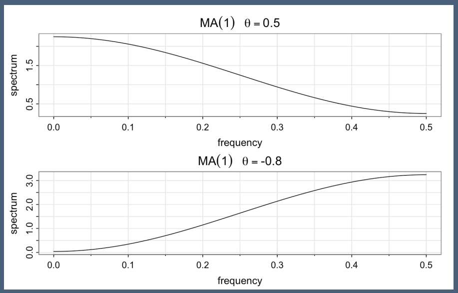
Example 6.10 shows AR(2)

$$f_x(\nu) = \frac{\sigma_w^2}{1 + \phi^2 - 2\phi \cos(2\pi\nu)}$$

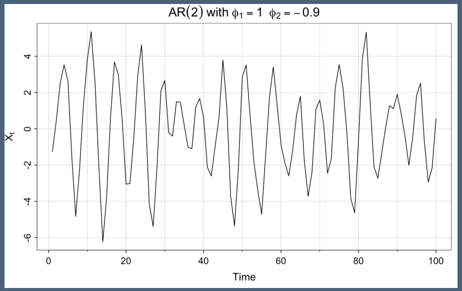
Examples of ARMA Spectra

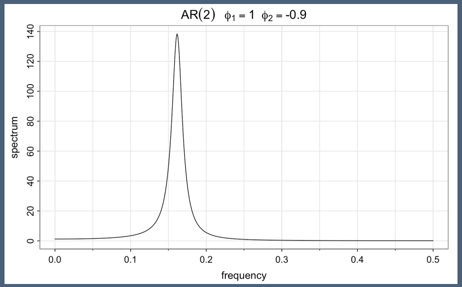
- Moving averages
  - Positive coefficient leads to more low frequency
  - Negative coefficient leads to more high frequency





- AR(2) process can show predominant cycle
  - $\phi_1 = 1$, $\phi_2 = -0.9$





Connection to Periodogram

- Periodogram of white noise
  - Multiple of exponential

$$I_k = \frac{n}{4}\left(a_k^2 + b_k^2\right) \approx \sigma_w^2\,\frac{\chi_2^2}{2}$$

  - Re-express using what we know about spectral density of white noise

$$I_k \approx f(k/n)\,\frac{\chi_2^2}{2}$$

- Periodogram of stationary process
  - Analogous property holds: periodogram is an exponential multiple of the spectral density
- Smoothing periodogram
  - Periodogram is not a consistent estimator of the spectral density: larger n just means more noisy values
  - Need to smooth: combine adjacent values to reduce variance (origin of scatterplot smoothing)

Estimating Spectrum

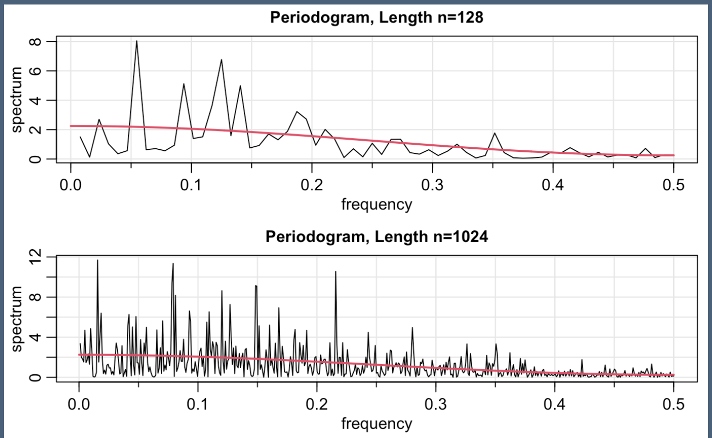
- Periodogram is not consistent estimator
  - Doesn't improve as the sample size increases; just get more noisy estimates

$$E\,I_k = f(k/n), \quad \text{Var}(I_k) = f(k/n)^2$$

  - Periodogram is an random exponential multiple of the spectrum

$$I(k/n) = \left(\frac{\chi_2^2}{2}\right) f(k/n)$$

  - Hence, periodogram is an inconsistent estimator of the spectral density.
  - On the good side, values I(k/n) are approximately uncorrelated… unlike estimated autocovariances
- Example
  - Assume process is MA(1) as in prior example
  - Observe either n=128 (top) or n=1024

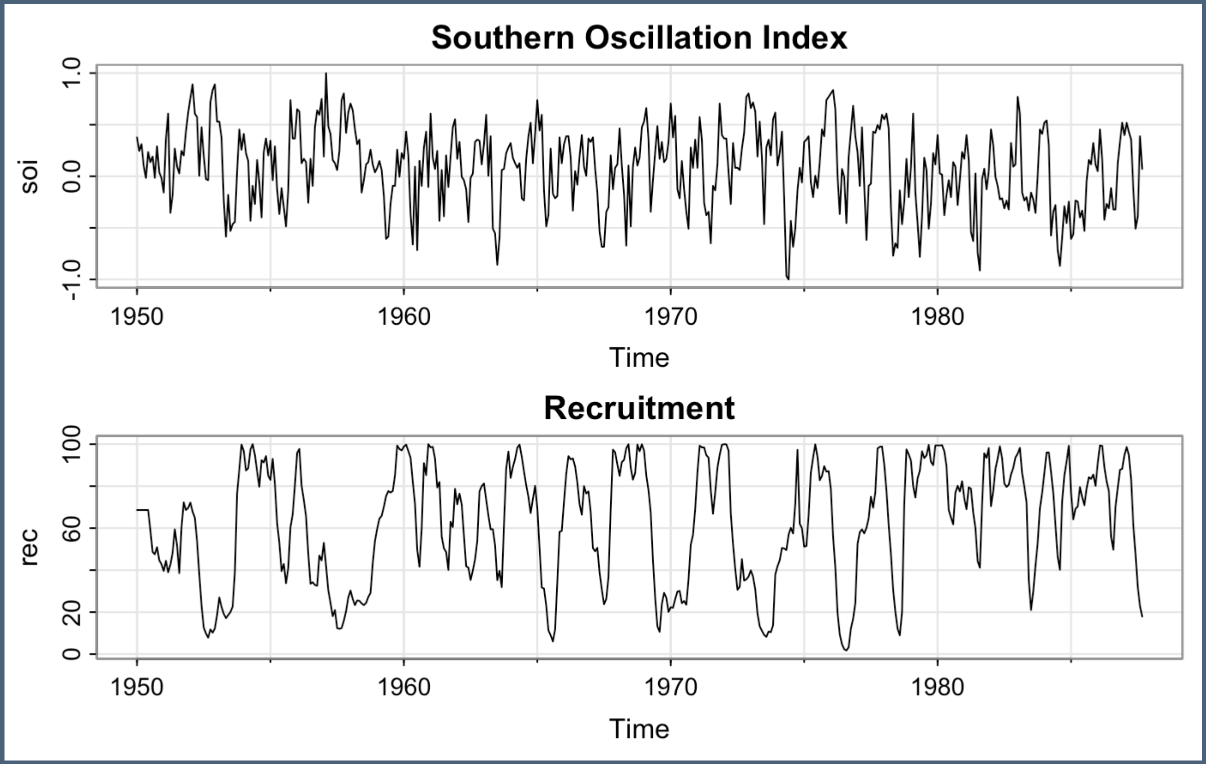


Climate Example

Southern Oscillation Index

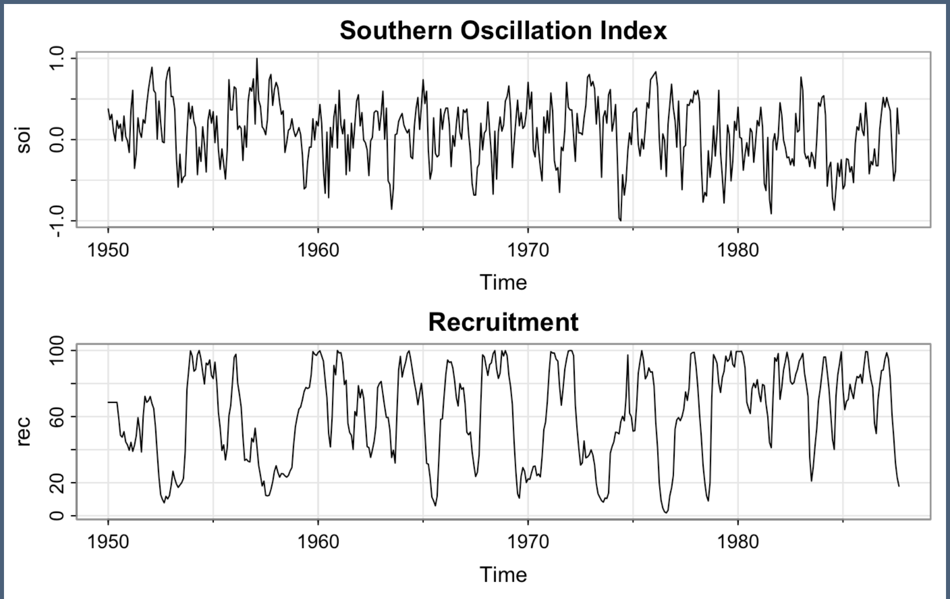
SOI Example

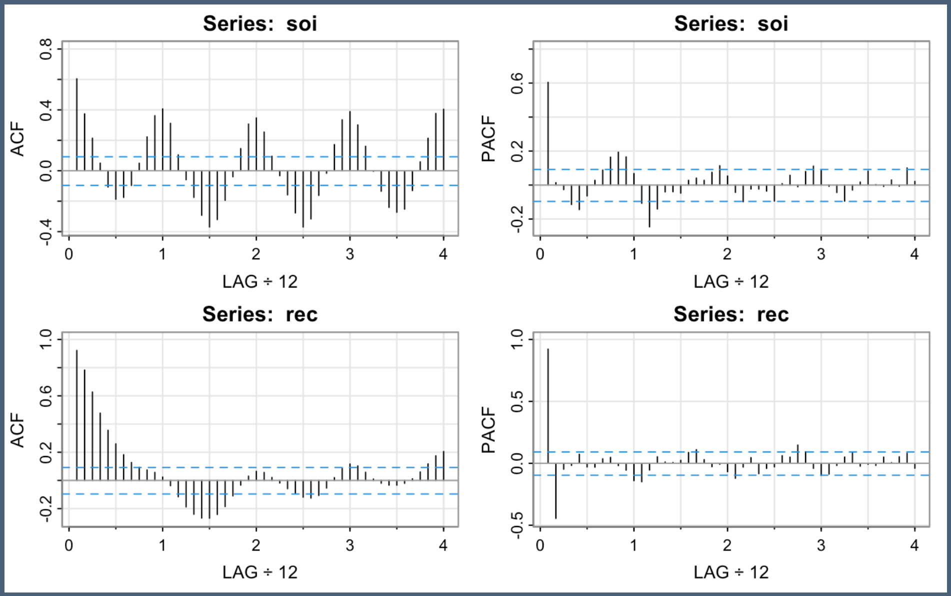
- Data
  - Southern oscillation index (soi, air pressure differences in Pacific) and recruitment (rec, measure of new fish)
  - Monthly, n = 453



SOI Example

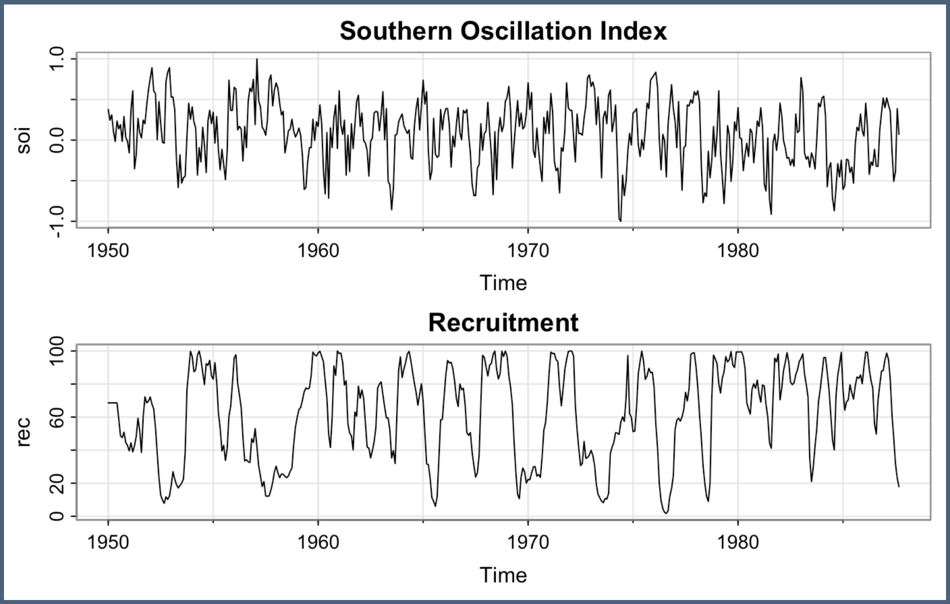
- ACF/PACF
  - SOI has long-term annual dependence, whereas REC resembles stationary AR(2)

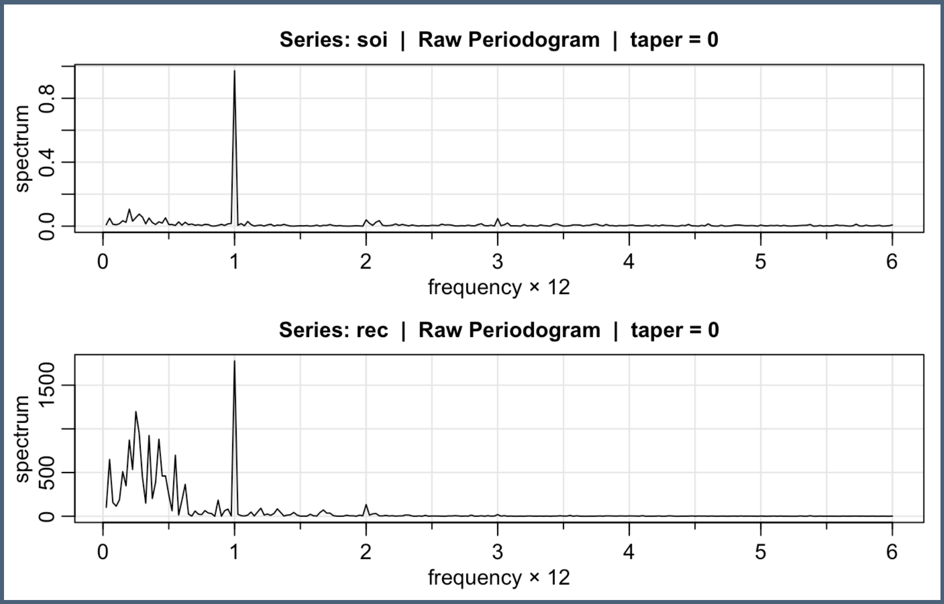




SOI Example

- Periodograms
  - SOI has very strong annual cycle, whereas REC has more low-frequency variation
  - Frequency scale runs from 0 to 12 (1/2) = 6 months

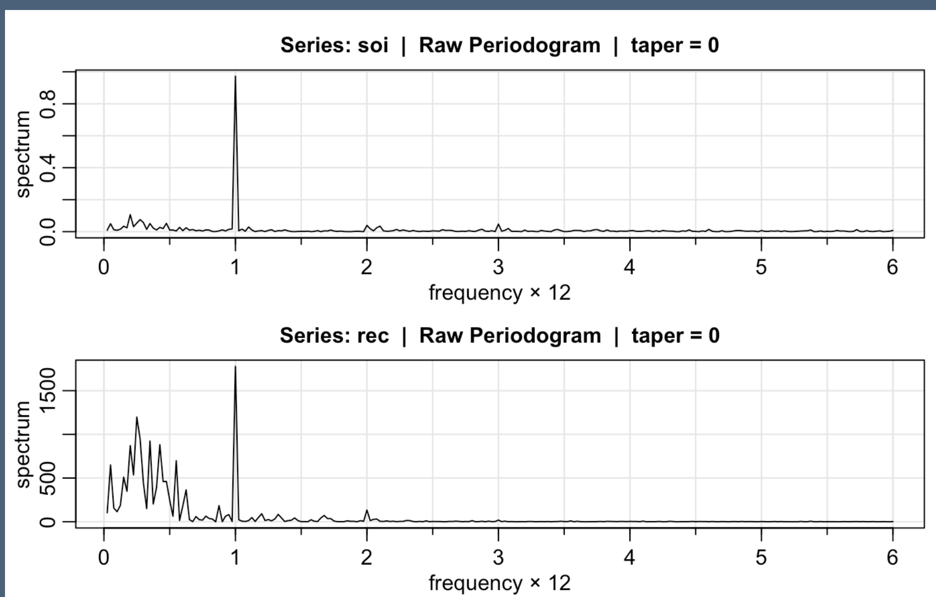


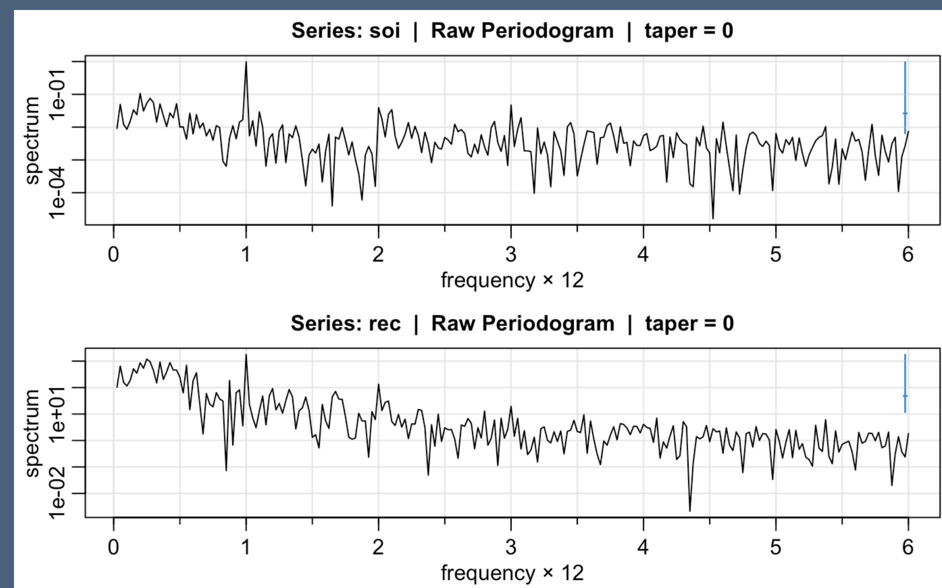


SOI Example

- Log periodograms
  - Variance of I(k/n) proportional to square of f(k/n)
  - Software indicates confidence interval in upper right corner
- Results
  - Peak doesn't look so impressive
  - Need for smoothing more evident

$$\log I(k/n) = \log f(k/n) + \log(\chi_2^2/2)$$

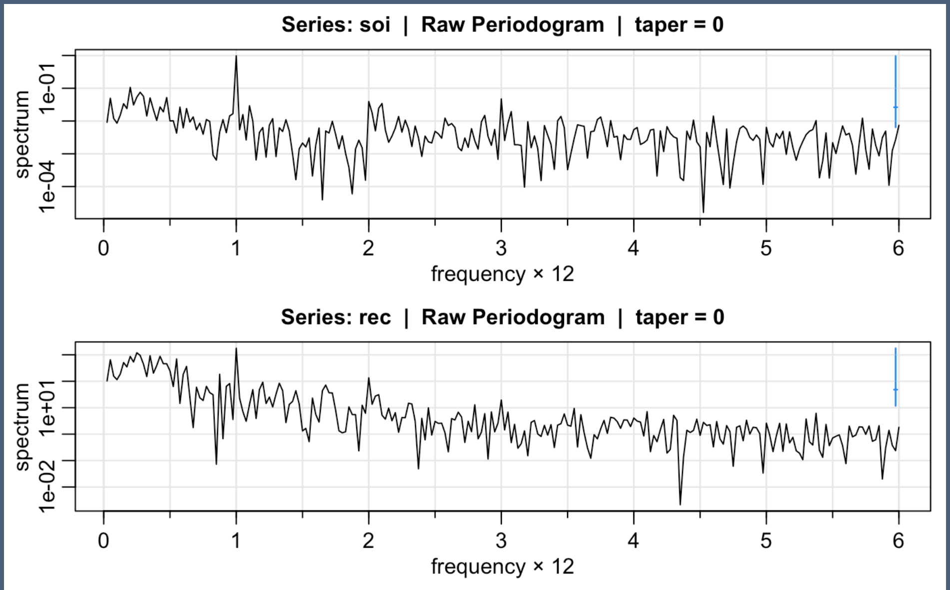


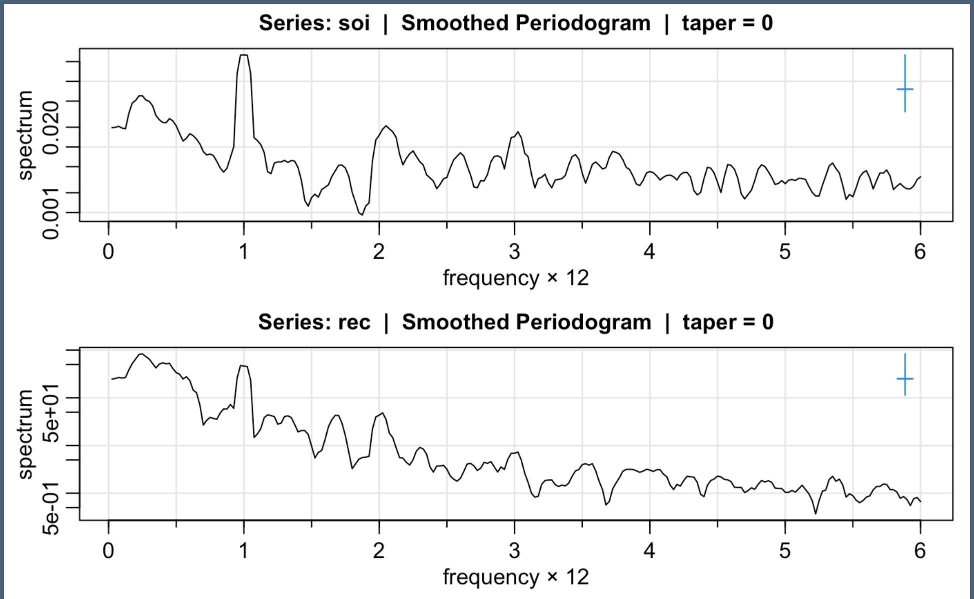


SOI Example

Figure 7.4, 7.5

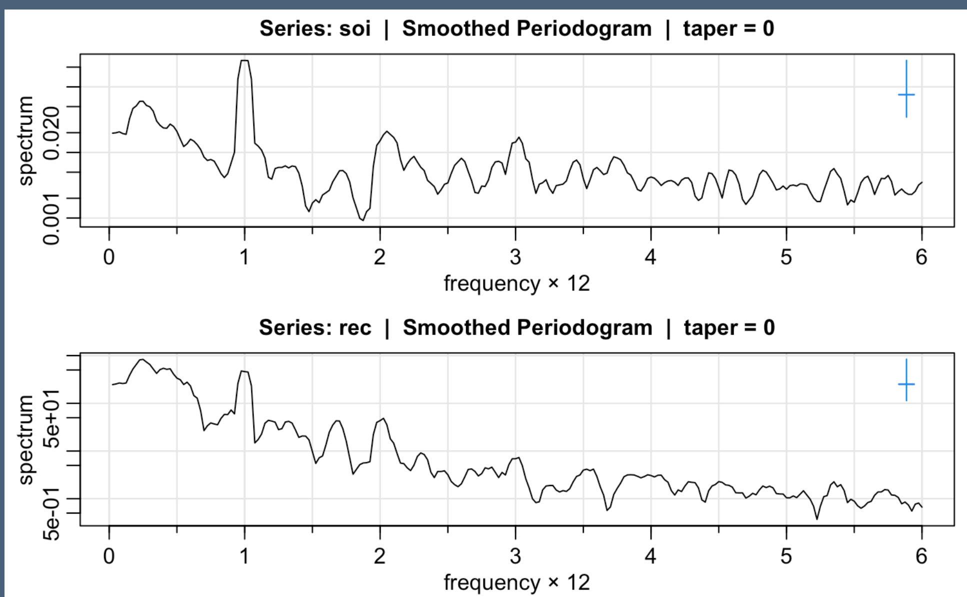
- Log periodograms with smoothing
  - Simple moving average smoother, width 5
  - Smoothing introduces bias in order to reduce variance: Peak in SOI becomes more spread out
  - Confidence interval is narrower

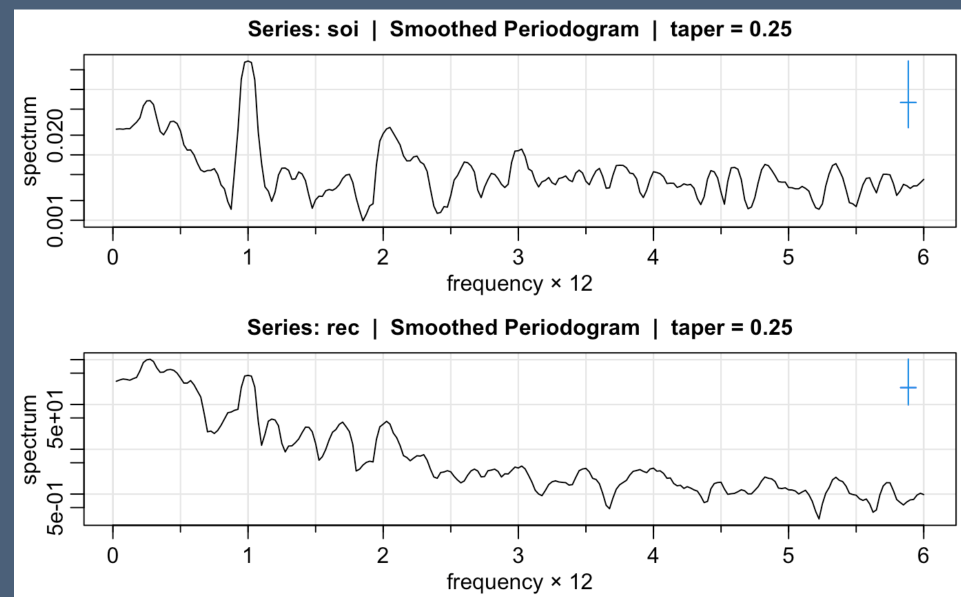




SOI Example

- Log periodograms with smoothing and tapering
  - Simple moving average smoother, width 5
  - Tapering sharpens peaks
  - Less leakage (bias) for SOI





What's next?

- Forecast revisions

- Context
  - Your company is ordering supplies for a future event
  - You've provided a forecast for how much is needed at the event
  - As the date of the event gets closer, your forecasts change
  - Are such changes appropriate?