Welcome to instats

The Session Will Begin Shortly
(At the top of the hour, Eastern USA time)

1

START

2

Nonlinear Time Series Analysis, Part I: Detecting Nonlinearity

Barney Ricca
Lyda Hill Institute for Human Resilience
University of Colorado Colorado Springs

3

Seminar Overview

- Day 1
  - Session 1: Introduction to Nonlinear Time Series (NTLS)
  - Session 2: Behaviors and State Spaces
- Day 2
  - Session 3: State Spaces (continued)
  - Session 4: Recurrences
- Day 3
  - Session 5: Tests
  - Session 6: Singular Spectrum Analysis and Noise
- Day 4
  - **Session 7: Surrogate Data**
  - Session 8: Convergent Cross Mapping

4

Statistics History

- Fisher (1931): F-test
- Quenouille (1949): Jackknife
- Rosenblatt (1957): Perceptron
- Efron (1979): Bootstrap

5

The Lady Tasting Tea

- Possibly apocryphal
- Fisher
  - Milk before tea or milk after tea?
  - Prepare 10 samples
  - List all possible arrangements (2^{10} = 1024 distinct patterns)
  - Calculate odds
- Wasn't possible in general
  - No computing power!
  - Approximated by the *F*-test

6

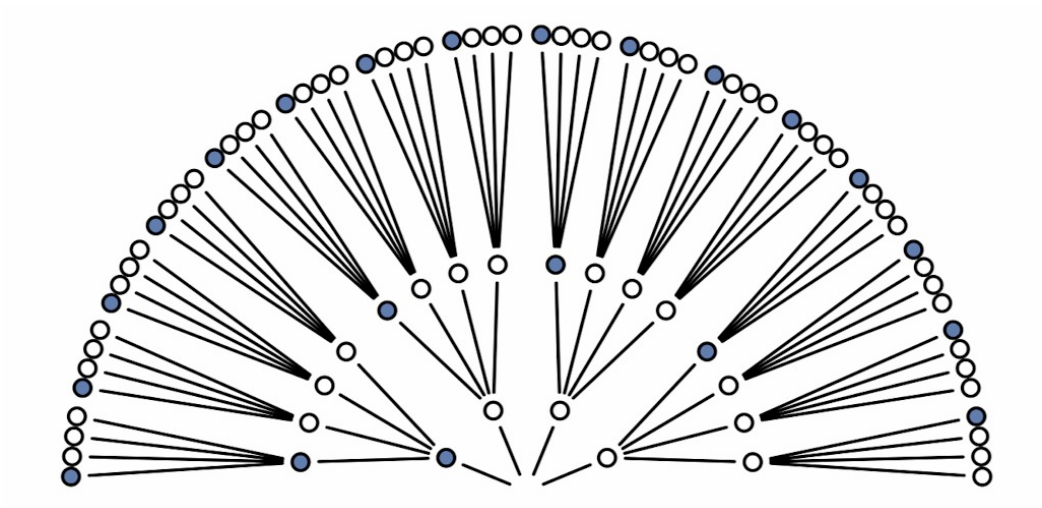

FIGURE 2.2. The 64 possible paths generated by assuming the bag contains one blue and three white marbles.

7

Hardware History

- Zuse (1941): Z3
  - 22 bits x 64 words, 10 Hz
  - *General purpose* computer (software!)
- BCS (1947): Semiconductor transistor
  - Moore's law
- Desktops (1980s-now)
  - 1992: 10 MHz, 16-bit single processor; $40/MB for RAM
  - 2020: 2.3 GHz, 64-bit, 8-processor chip; $5/GB for RAM

8

Jackknife

- Jackknife: Shuffle the data
  - A.k.a., permutation
  - Resample *without* replacement
  - Multiple repetitions to get histogram of possible outcomes
- Jackknives stand-in for the population
  - Empirical, not theoretical, distribution

9

Bootstrap

- Resample *with* replacement
  - More flexible research designs
  - Better (mostly) asymptotic convergence
- Many replications
  - Variance estimates
  - Count number in/out of condition
- Beer, water, and mosquito bites in R

10

Some R Examples

Resampling or Counting	Approximation
fisher.test()	chisq.test()
chisq.test(simulate.p.value = TRUE)	chisq.test()
lmPerm::lmp()	lm()
lmPerm::aovp()	aov()

Beer, water, and mosquito bites in R

11

Dynamics and Resampling

Jackknife: Independent, identically distributed (i.i.d.)

Time Series

Dynamics: State dependent on previous states

12

Surrogate Data

- Frame the null hypothesis about the shadow state space
  - **Watch your assumptions!**
  - E.g., If not i.i.d., need to maintain sequence structure during resampling
- Generate surrogates compatible with the null
  - Maintain some connection between consecutive data points
- Common approaches
  - Block resampling (general time series approach)
  - Fourier transform power spectrum (FTPS)
  - Amplitude-adjusted Fourier transform (AAFT)
  - Pseudo-periodic surrogates (PPS)

13

Block Resampling

- Shuffle blocks of data
  - Each block maintains the dynamics internally
- Lots of details are important
  - Block length
  - How many blocks
  - Do blocks overlap or not
- Surrogate data in R (*package:tsboot*)

14

Phase Resampling: FTPS and AAFT

- Work with the spectra
- Shuffle the phases
  - Maintains each oscillatory component
- Process
  - Fourier transform time-series to spectra
  - Shuffle
  - Transform back to time-domain to create surrogate data
  - Embed and estimate
- Surrogate data in R

15

Pseudo-Periodic Surrogates

- Tests aperiodic oscillations
  - Can’t use Fourier transforms
- Process:
  - Embed data
  - Randomly walk on the shadow attractor
  - First coordinate (e.g., un-delayed coordinate) is surrogate

16

Why Use FTPS and AAFT?

- PPS: Nonlinear
- FTPS & AAFT: Linear
- Q: Why use the linear?
  - A: Suppose PPS and FTPS/AAFT give the same results?
- *Possibly*:
  - Use the comparison to filter out linear signal
  - Maybe

17

General Resampling Notions

- Null hypothesis
- Assumptions
- Discriminating statistics

18

Distinguishing Statistics

- Distinguish between random and deterministic
  - HBR: "…there is general agreement that [these] measures can be used reliably…"
- Correlation dimension
  - Fractional dimension of reconstructed attractor
- Maximum Lyapunov exponent
  - Sensitivity to initial conditions (positive only in nonlinear situations)
- Nonlinear prediction error
  - Nash-Sutcliffe

19

Questions

20

STOP

21

Next session @ UTC 1900

22