Data Mining and Business Intelligence

Lecture 6: Time Series Introduction

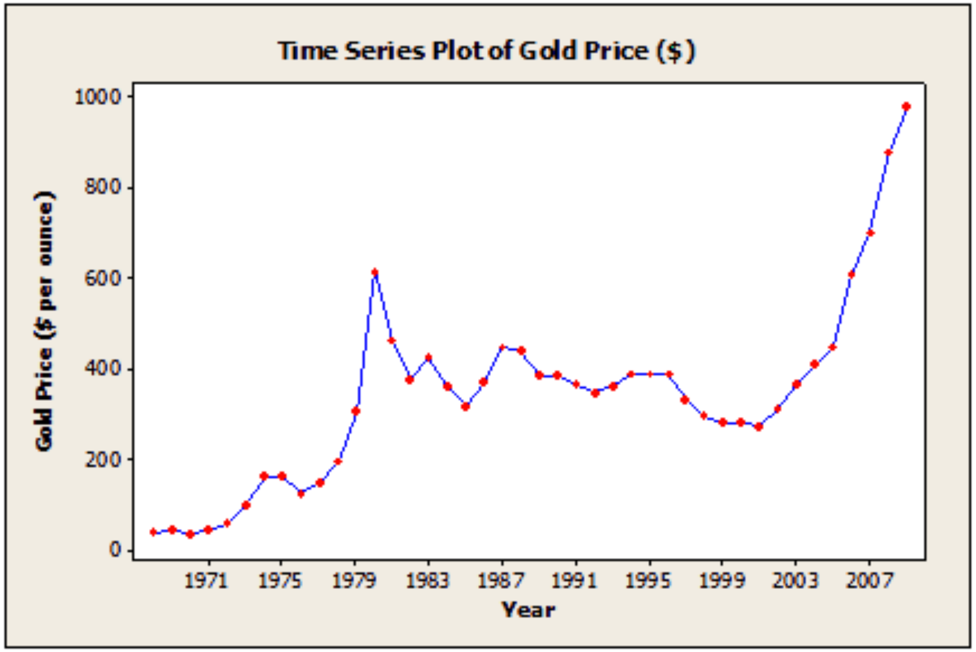
Jing Peng
University of Connecticut

2/27/20

Time Series Basics

What is a time series?

- A **time series** is a sequence of measurements of the same variable collected over time, often at regular time intervals

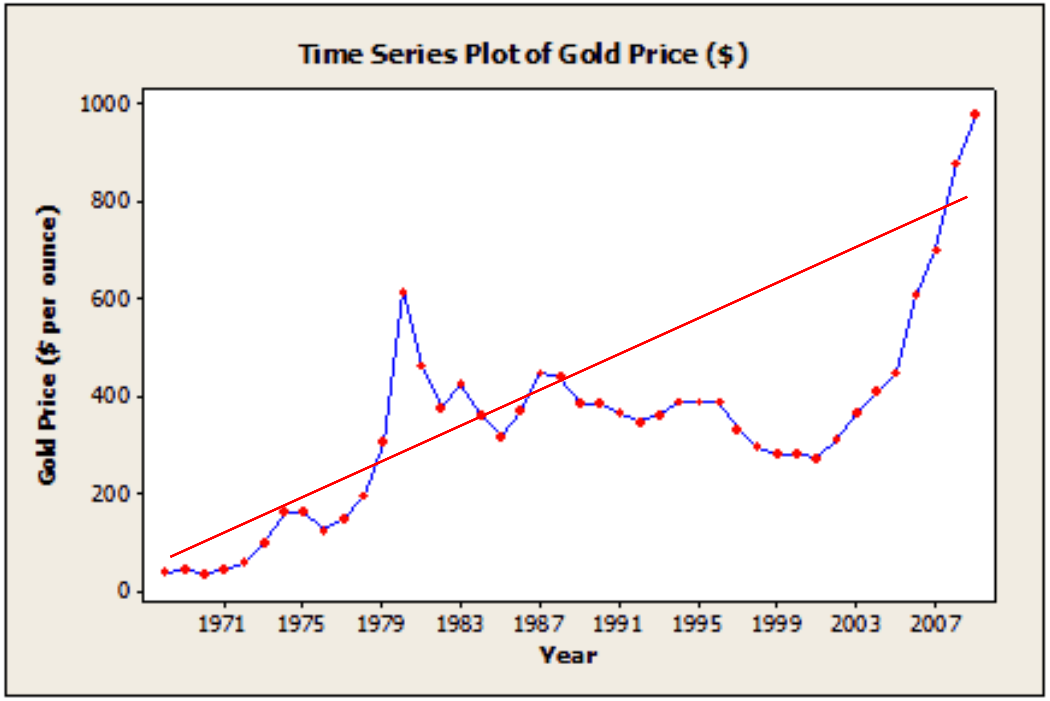


Common Patterns of Time Series

- Trend
- Seasonality
- Events

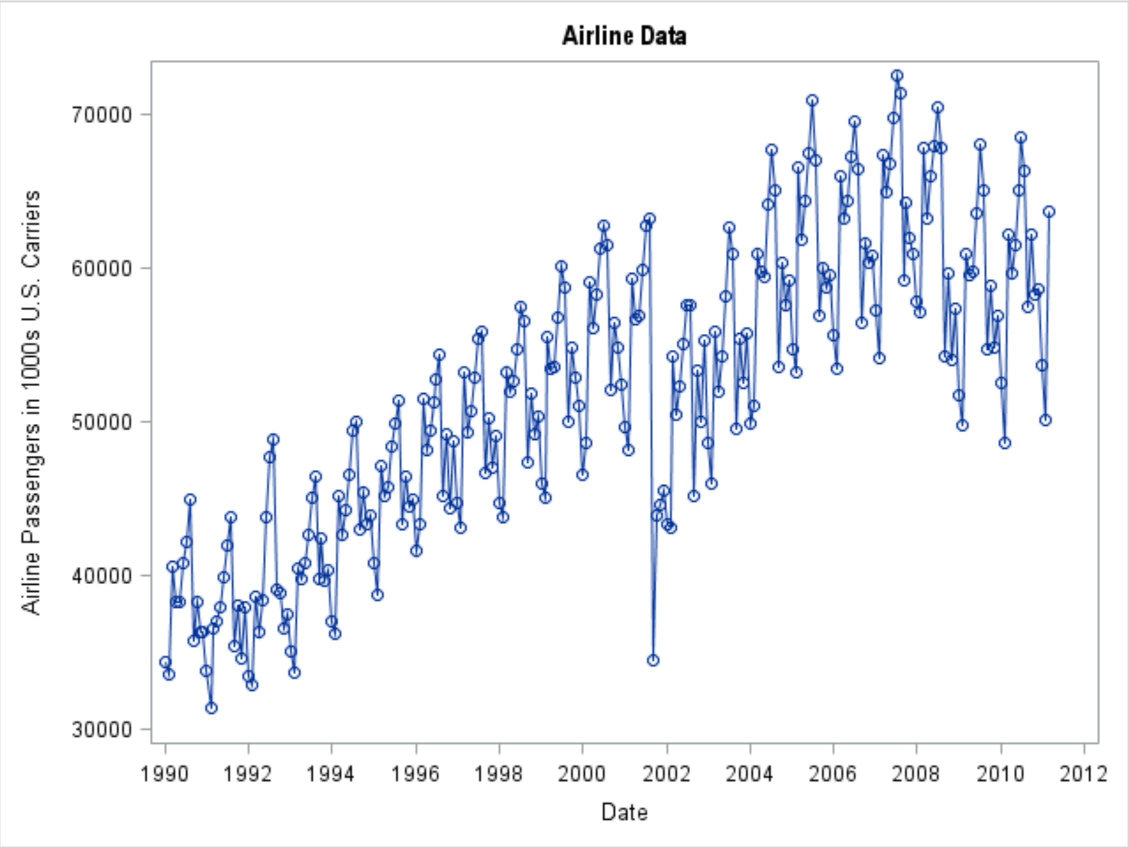
Pattern: Trend

- Trend: long term increase or decrease in the measurement



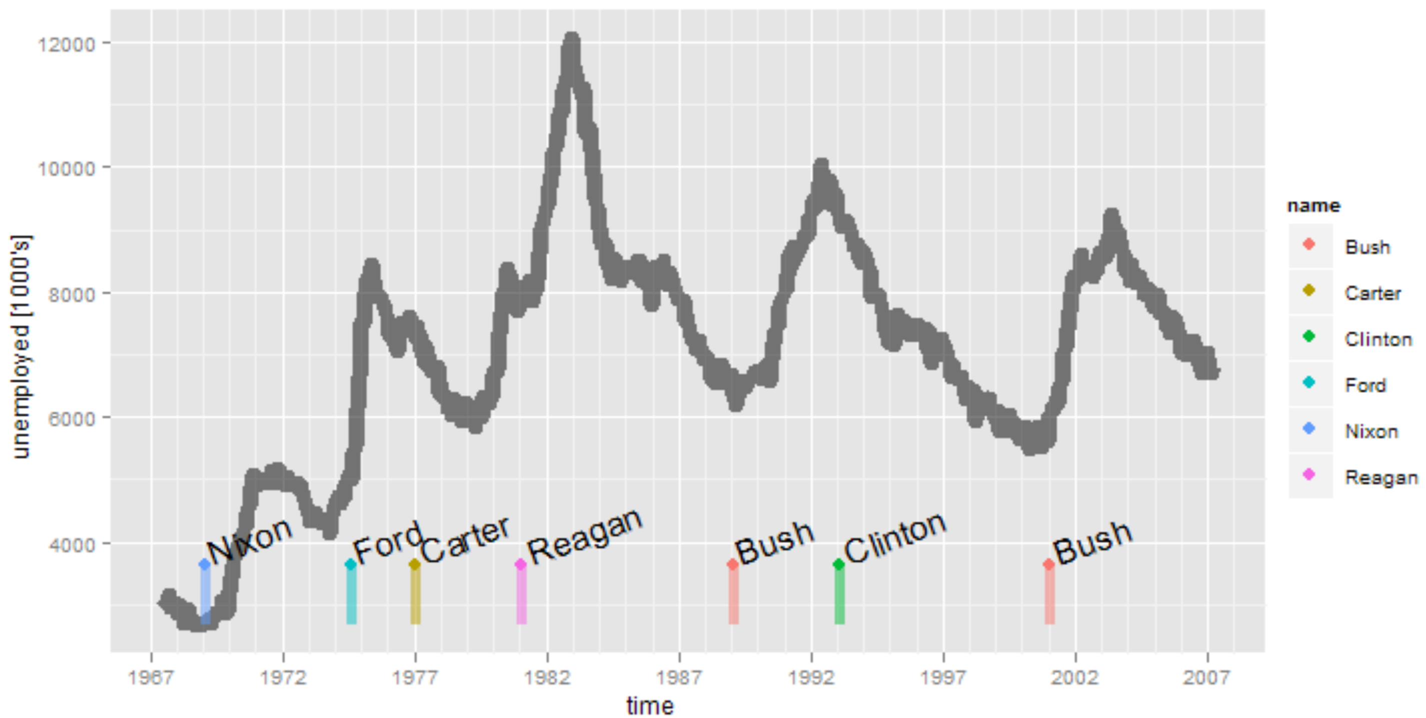
Pattern: Seasonality

- Seasonality: patterns reoccurring at a regular interval (e.g., day, week, year)



Pattern: Events

- Event: a factor that results in abrupt changes in outcome



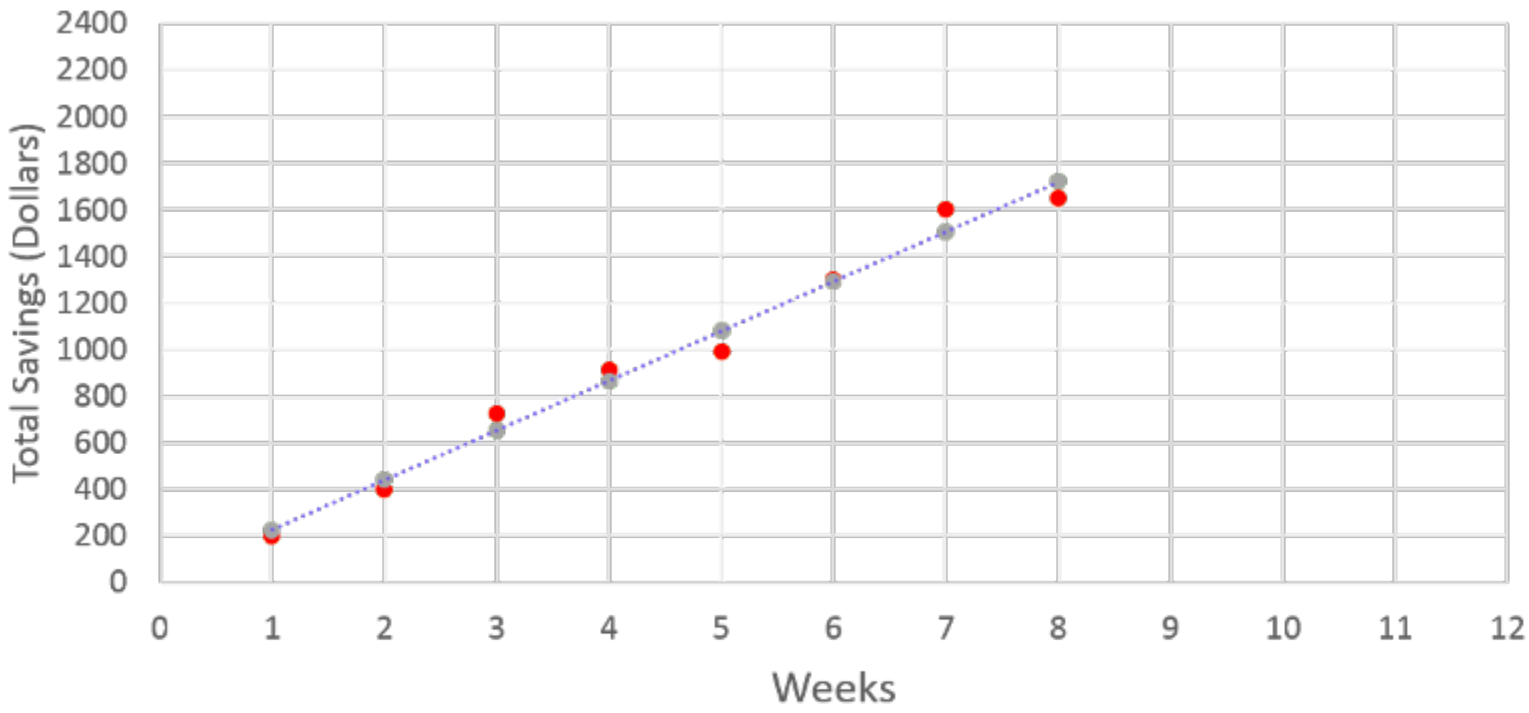
Source: https://stackoverflow.com/questions/8317584/r-ggplot-time-series-with-events

Time Series Models

Models for Time Series

- Linear Regression
- Exponential Smoothing Models
- Autoregressive Integrated Moving Average Models (ARIMA)
  - Autoregressive (AR)
  - Moving Average (MA)
- Recurrent Neural Networks: Long Short-Term Memory

Linear Regression



$$Savings_t = \beta_0 + \beta_1 Week_t + \varepsilon_t \ (t = 1, \dots, 8)$$

Beyond Linear Regression: Example

- Bob opened a bank account with \$500 initial deposit
- Bob deposits \$1000 at the end of each week
- How much money does Bob have in his account after 10 weeks?
  - What if Bob spends 30% of his money at the beginning of each week?
  - What if there is some random variation in the amount Bob deposits each week (e.g., \$995 in week 1, \$1001 in week 2, …)

Beyond Linear Regression: Example (cont.)

- A recursive approach

$$Y_0 = 500$$

$$Y_t = 1000 + 0.7Y_{t-1}$$

- With random variation

$$Y_t = 1000 + 0.7Y_{t-1} + \varepsilon_t$$

$$\Downarrow$$

Autoregressive Model

Autoregressive (AR) Model

- AR (1): autoregressive model of order 1

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

- AR(p): autoregressive model of order p

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

- Examples: global temperature, stock price

Moving Average (MA) Model

- MA (1): moving average model of order 1

$$Y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

- MA(q): moving average model of order q

$$Y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

- Examples
  - Unobserved shocks with sustaining effect (e.g., coupons and disasters)
  - Data manipulation (e.g., smoothing and interpolation)

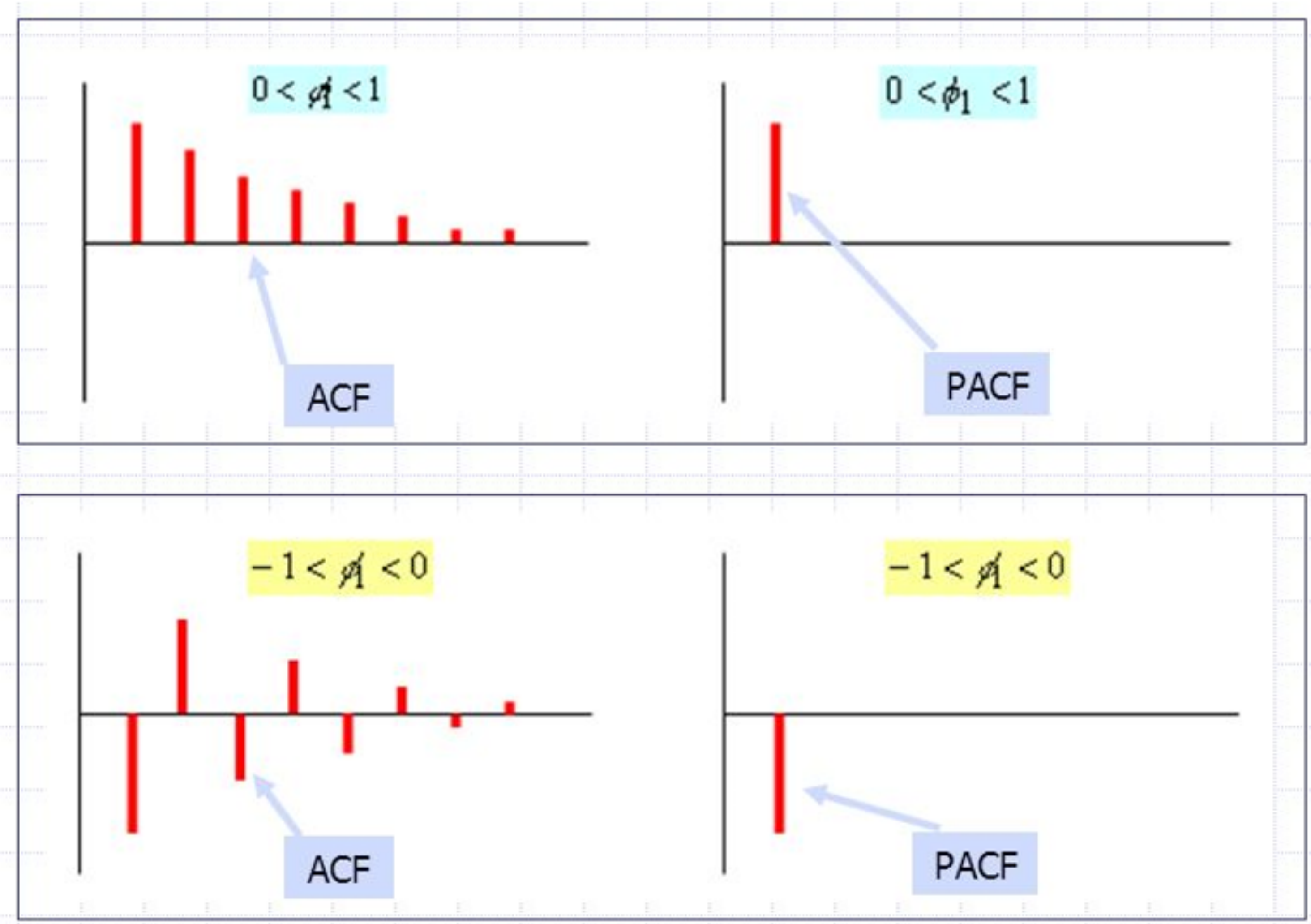
https://stats.stackexchange.com/questions/45026/real-life-examples-of-moving-average-processes

Autocorrelations

Correlation Functions

- Autocorrelation, also known as serial correlation, is the correlation of a signal with a delayed copy of itself as a function of delay.
- **Autocorrelation Function** (ACF): Correlation of observations k time units apart
- **Partial Autocorrelation Function** (PACF): Correlation of observations k time units apart, **conditional on** observing observations in-between

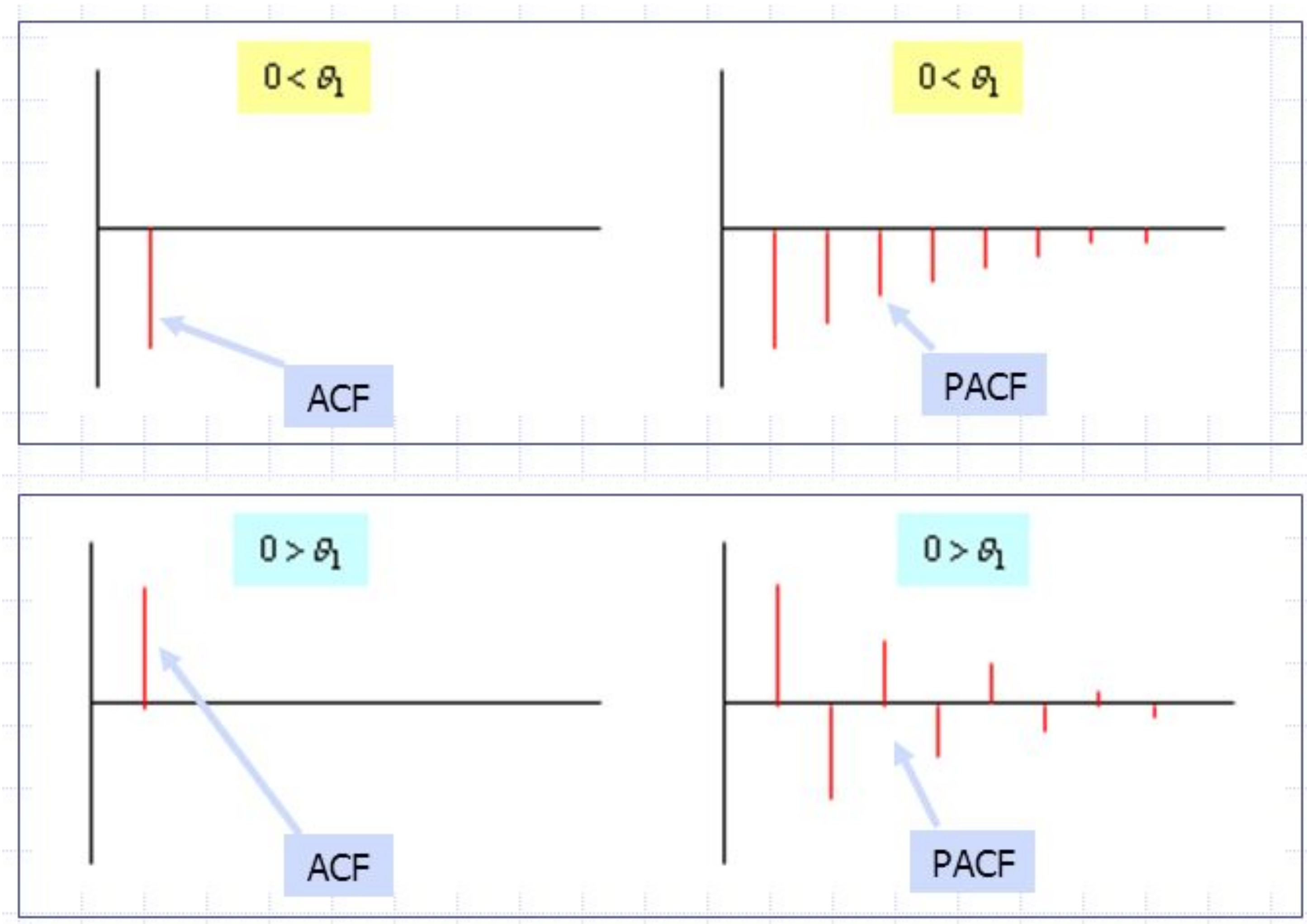
Theoretical ACF and PACF for AR(1)



$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

Source: http://slideplayer.com/slide/6399540/

Theoretical ACF and PACF for MA(1)



$$Y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Source: http://slideplayer.com/slide/6399540/

Theoretical Patterns of ACF and PACF

Type of Model	Typical Pattern of ACF	Typical Pattern of PACF
AR (p)	Decays exponentially or with damped sine wave pattern or both	**Cut-off after lags *p***
MA (q)	**Cut-off after lags *q***	Declines exponentially
ARMA (p,q)	Exponential decay	Exponential decay

Source: http://slideplayer.com/slide/1507028/

Conditional Dependence vs. Independence

- Conditional independence: for AR(1) model, Y(t) and Y(t-2) are unrelated conditional on Y(t-1)

- Conditional dependence: for MA(1) model, Y(t) and Y(t-2) become correlated conditional on Y(t-1)

- Conditional dependence: events A and B are unrelated, but they become correlated conditional on event C
  - A: money spent on food
  - B: money spent on clothing
  - C: total money spent

Inverse Autocorrelation Function

- IACF is the ACF of the inverse model

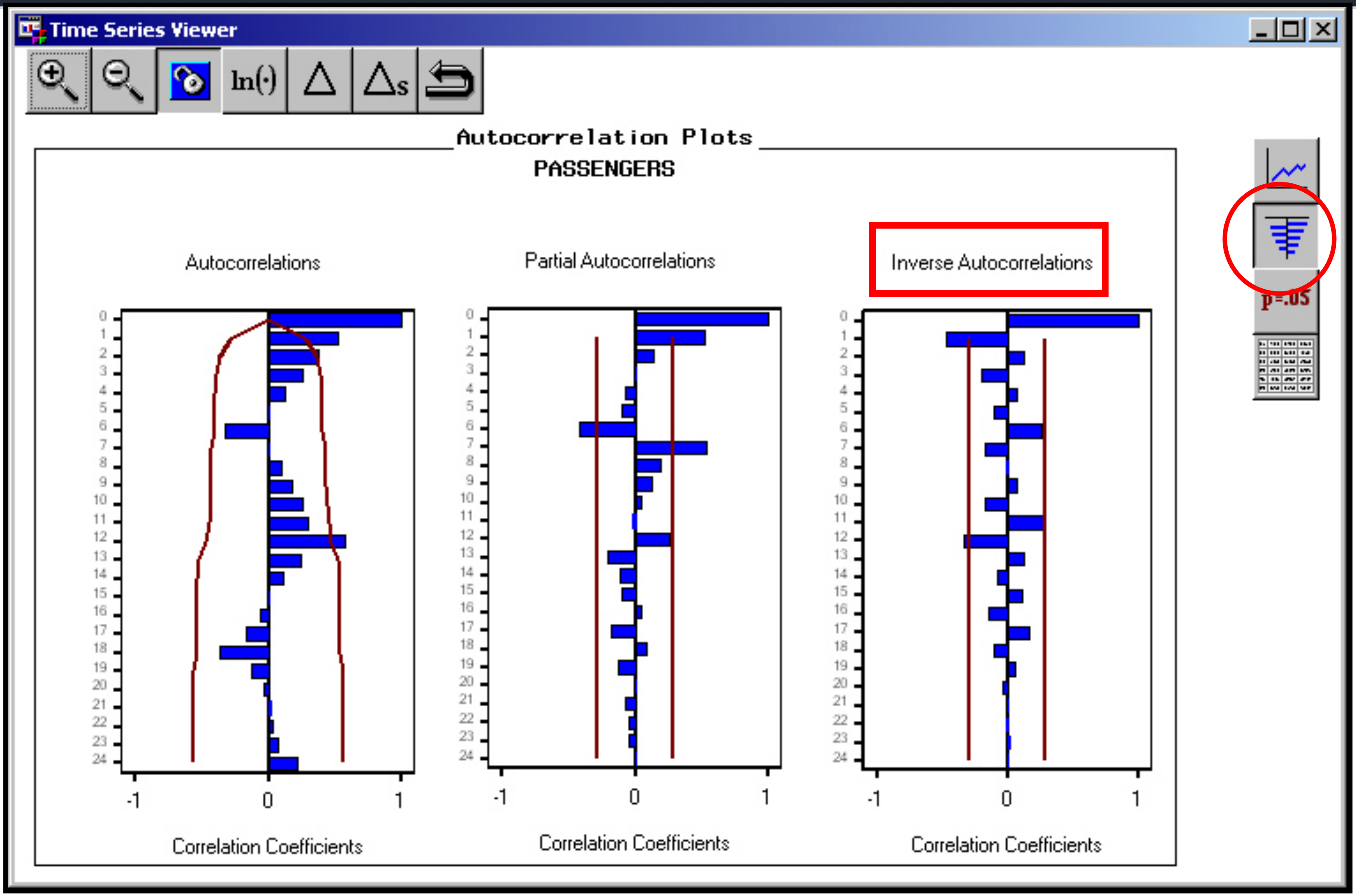
$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

$$\Updownarrow \quad \text{inverse model}$$

$$Y_t = \theta_0 + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \varepsilon$$

- IACF is similar to PACF, but more sensitive than PACF for some features of autocorrelation

Autocorrelation Plots in SAS



Autocorrelation Detection

Detection of Autocorrelation

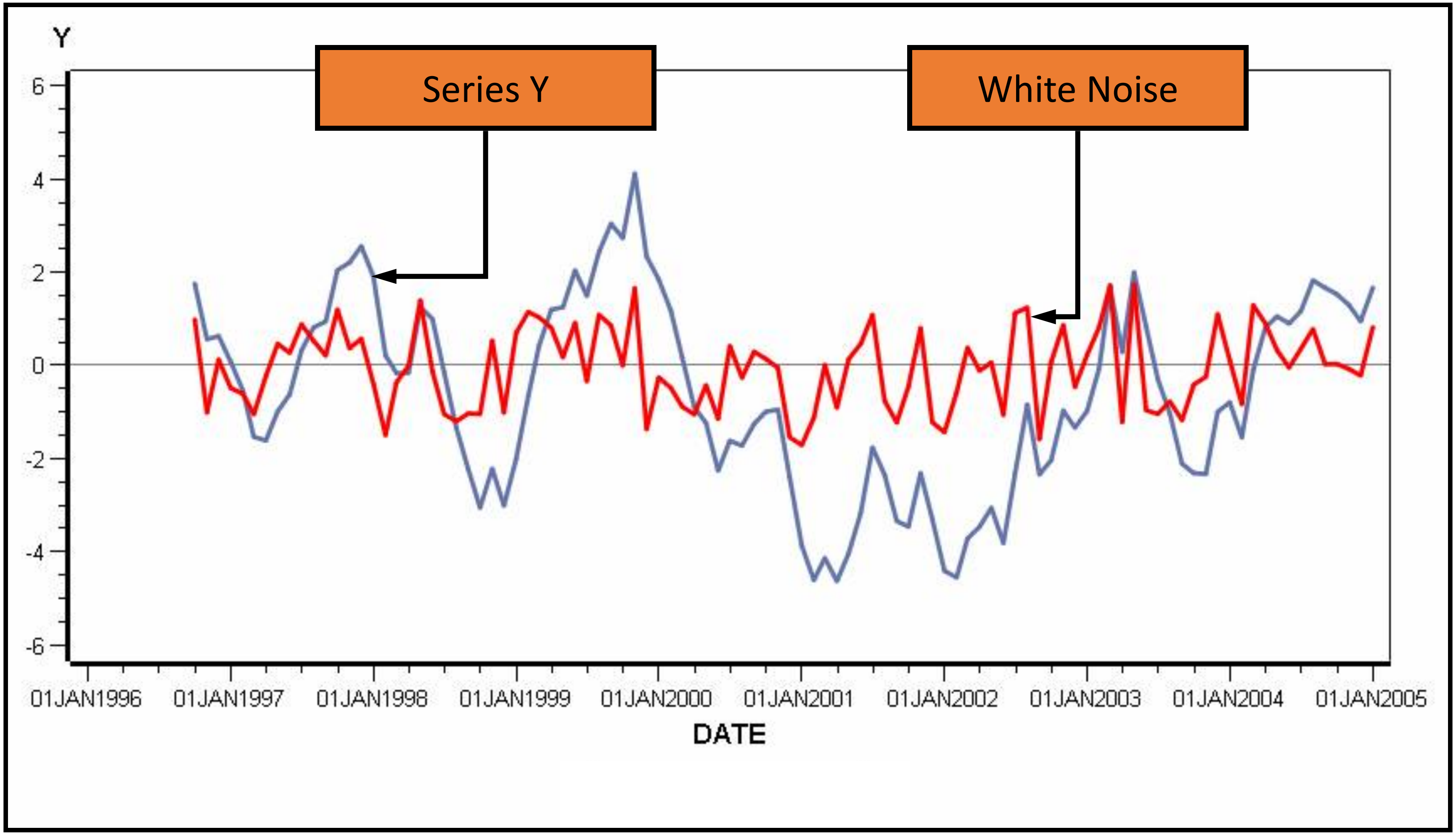
- Approach 1: Manually inspection
- Approach 2: Autocorrelation plots

Comparison of Two Time Series

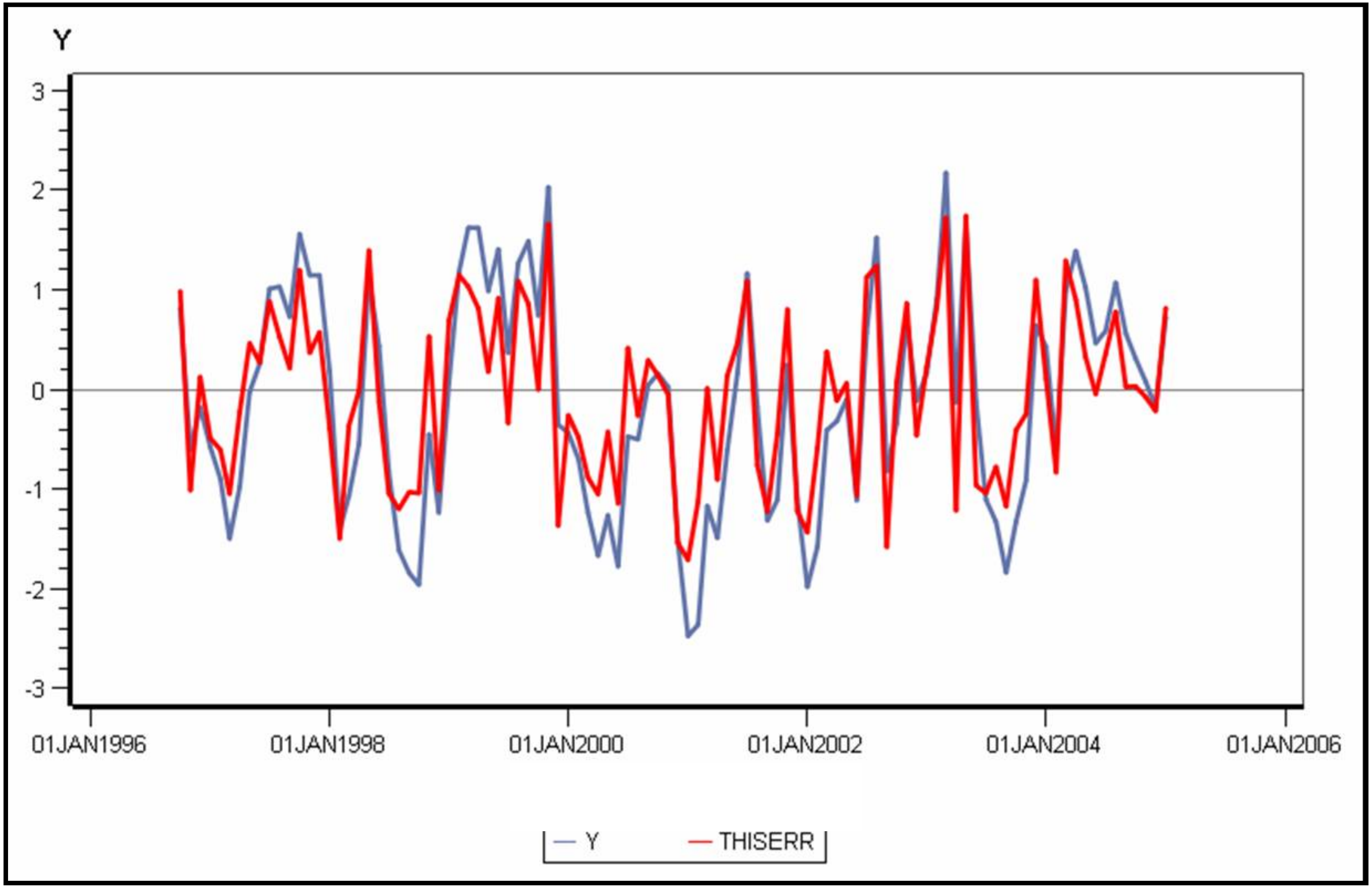
- AR(1): $Y_t = \phi Y_{t-1} + \varepsilon_t$

- White noise: ε_t

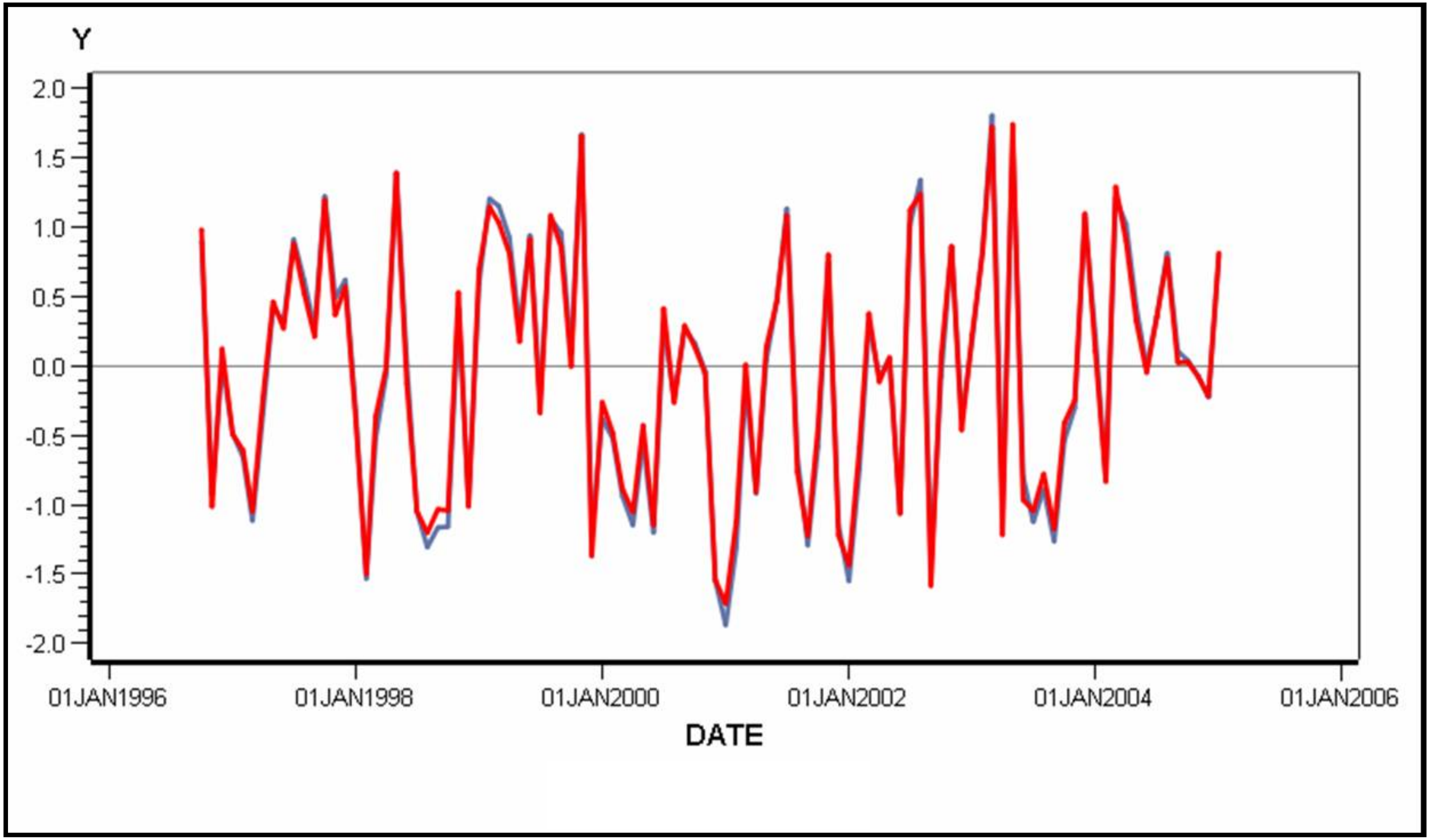
Plot of Series with ACF(1)=0.90 Overlaid with the Error Term (White Noise)



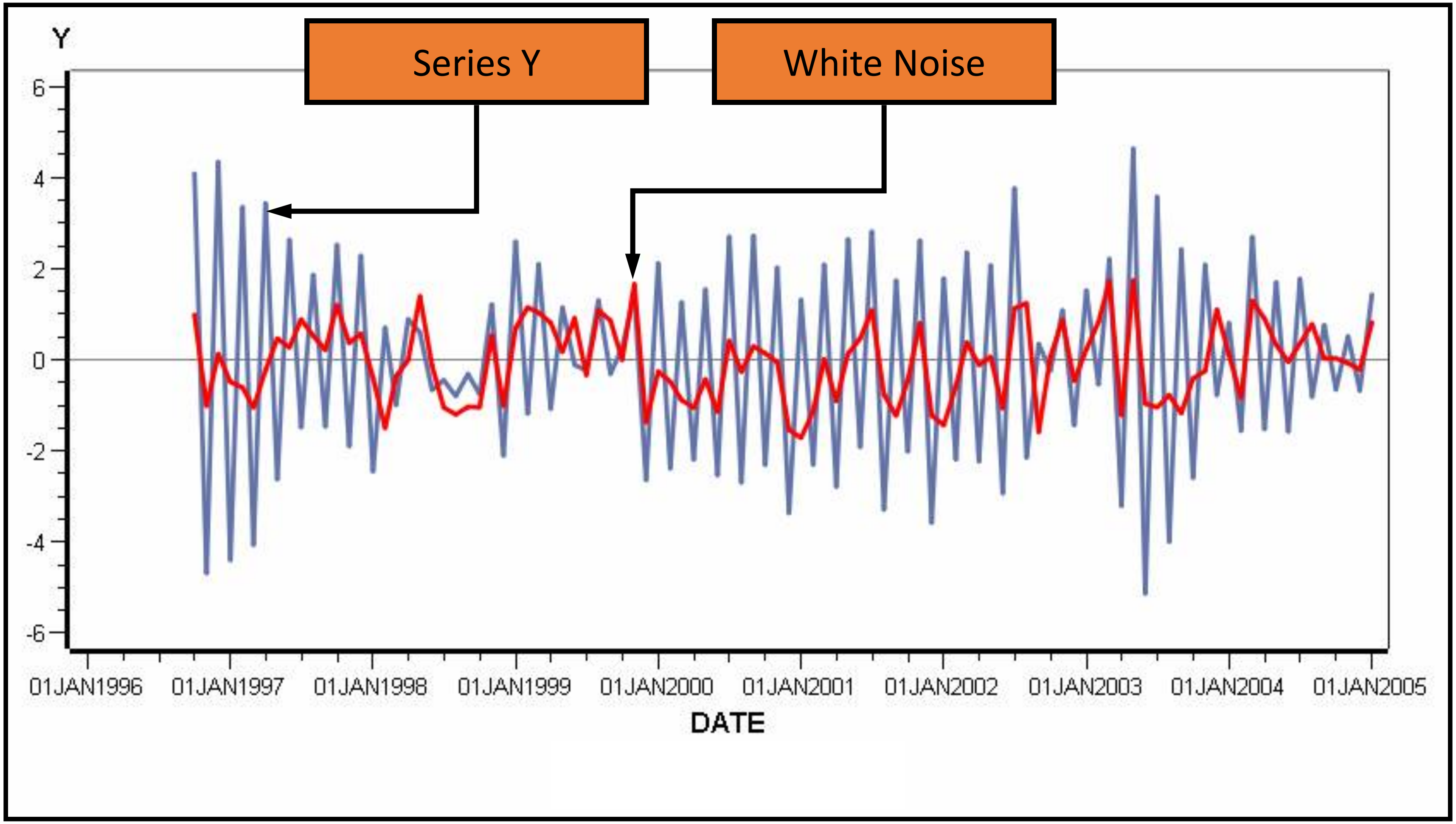
Plot of Series with ACF(1)=0.50 Overlaid with the Error Term (White Noise)



Plot of Series with ACF(1)=0.10 Overlaid with the Error Term (White Noise)



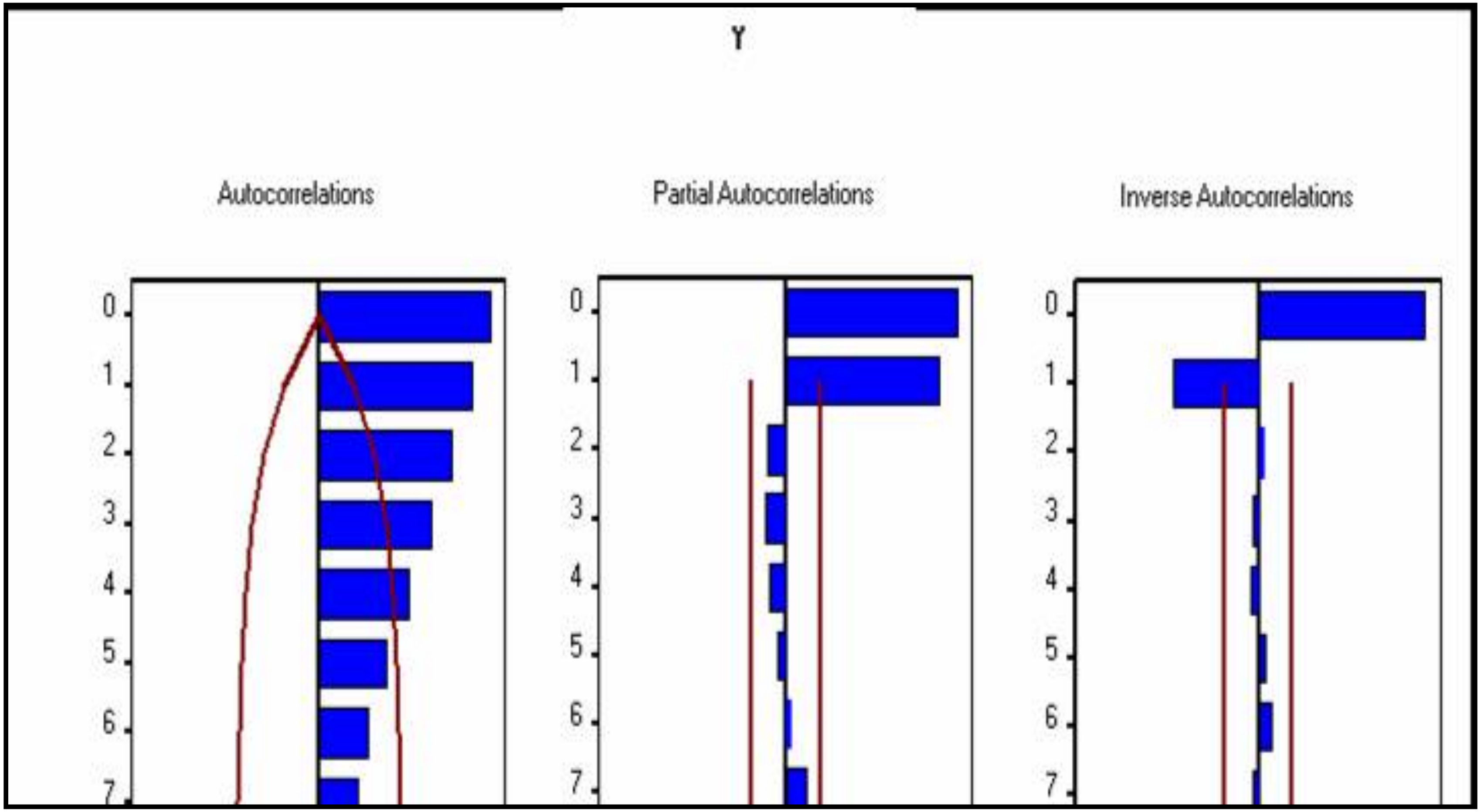
Plot of Series with ACF(1)= -0.90 Overlaid with the Error Term (White Noise)



Takeaway for Manual Inspection

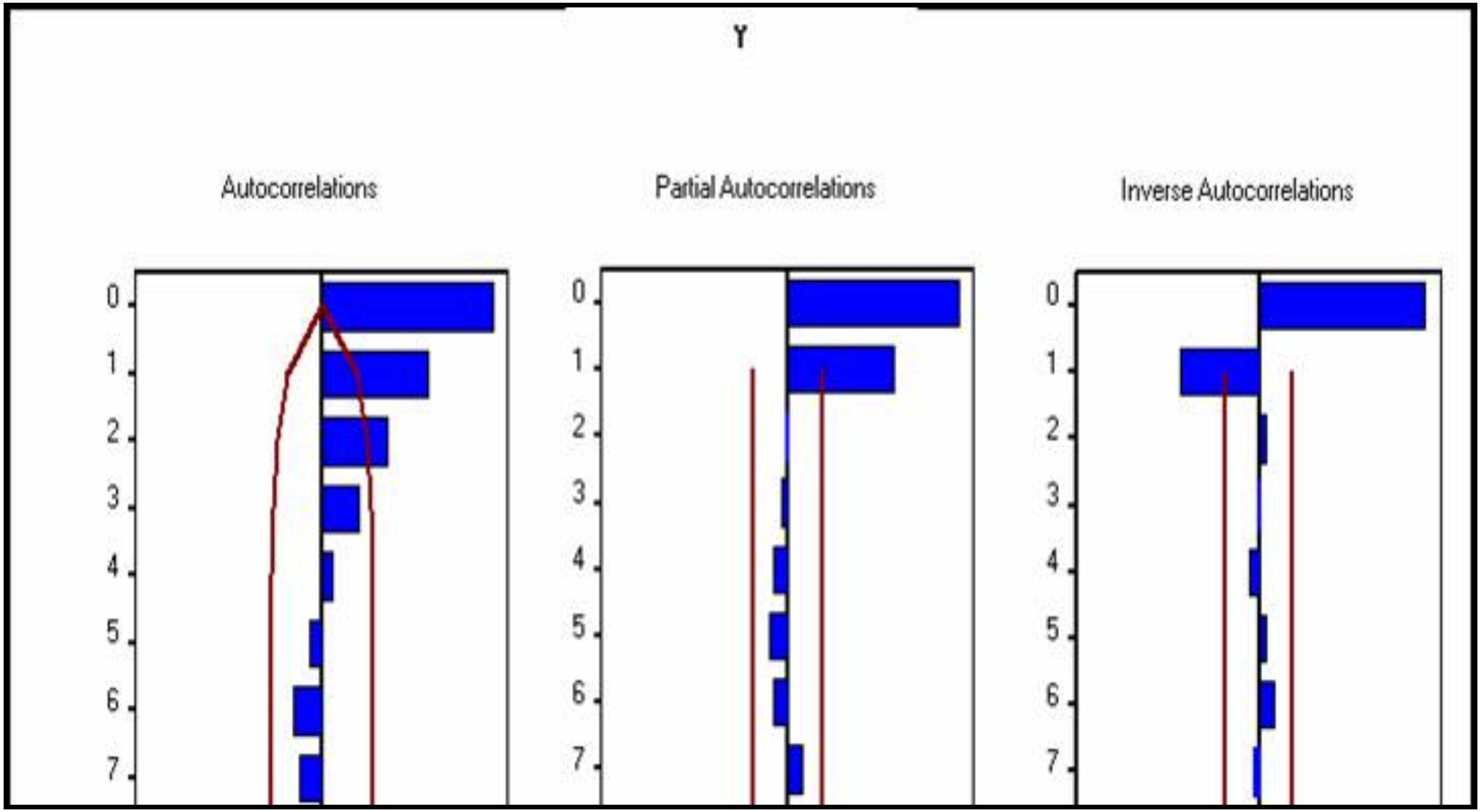
- **Positive autocorrelation**: remain above or below the mean for extended periods of time
- **Negative autocorrelation**: rapidly switch between above and below the mean
- Autocorrelation is NOT always detectable from plots of the data

Autocorrelation Plots for Series with ACF(1)=0.90

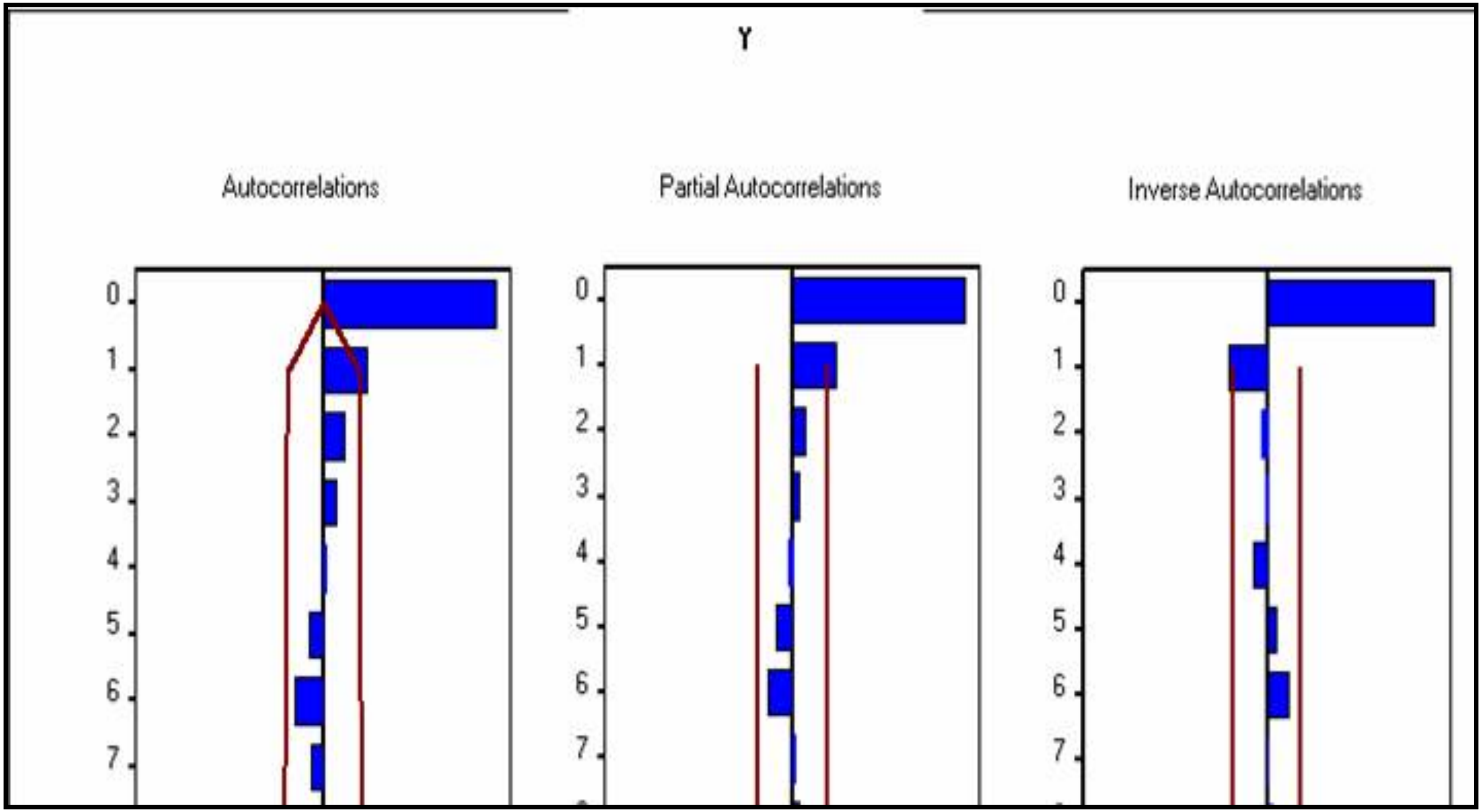


continued...

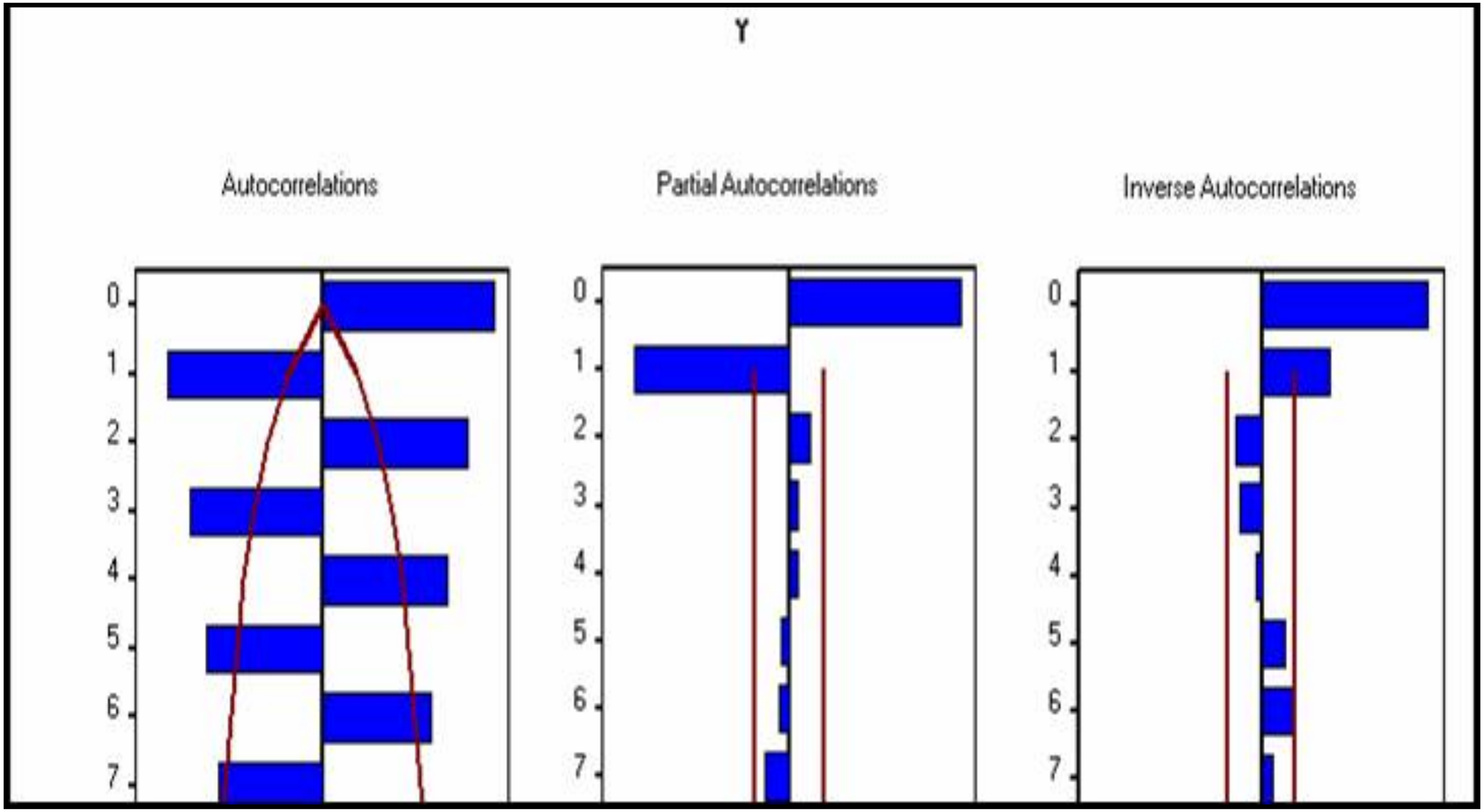
Autocorrelation Plots for Series with ACF(1)=0.50



Autocorrelation Plot for Series with ACF(1)=0.10



Autocorrelation Plots for Series with ACF(1)=-0.90



Takeaway for Autocorrelation Plots

- Can effectively detect autocorrelation even if the autocorrelation is small
- There are 5% probability that a significant spike is observed in white noise series (definition of p-value)
- IACF plot for the same dataset may vary across software, and even different versions of the same software
- IACF and PACF may have different number of significant spikes

Illustration in R

- The data underlying the above plots are not provided, now let’s generate similar data on our own

- Overlaid.R

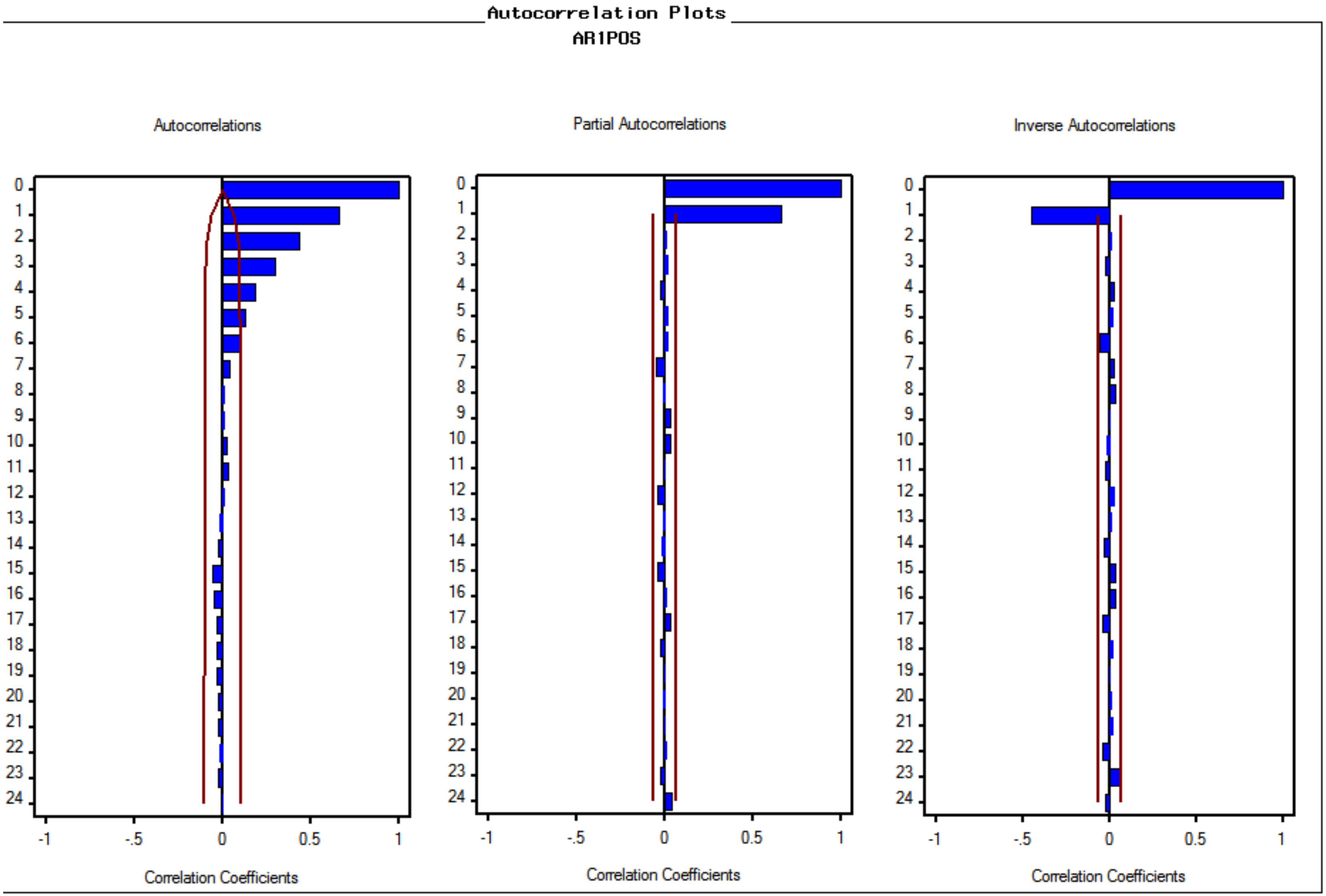
Simulations

Demo on Simulated Datasets

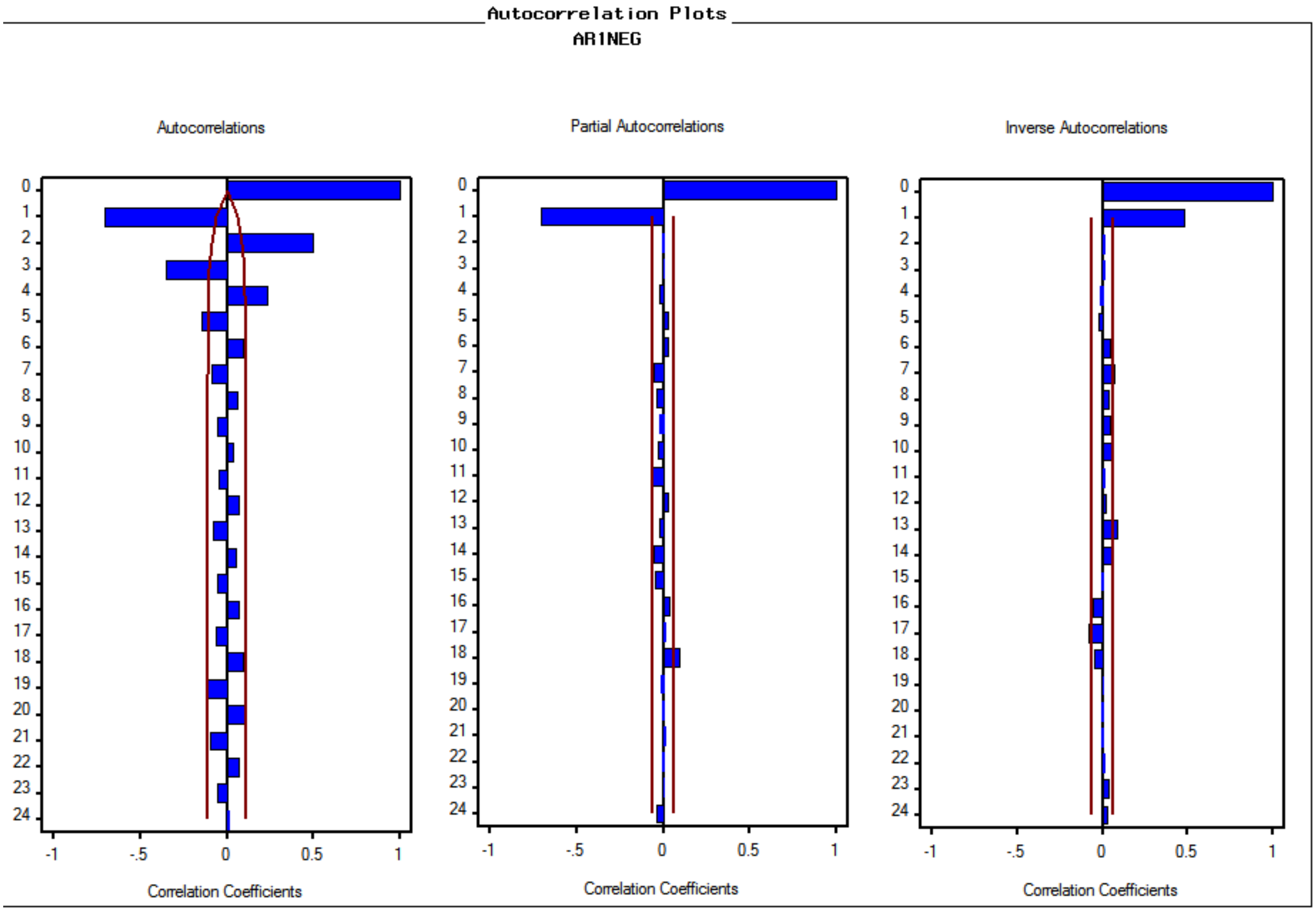
- Four time series
  - AR(1) with $\phi = 0.7$
  - AR(1) with $\phi = -0.7$
  - MA(1) with $\theta = 0.7$
  - MA(1) with$\theta = -0.7$

simulate.R

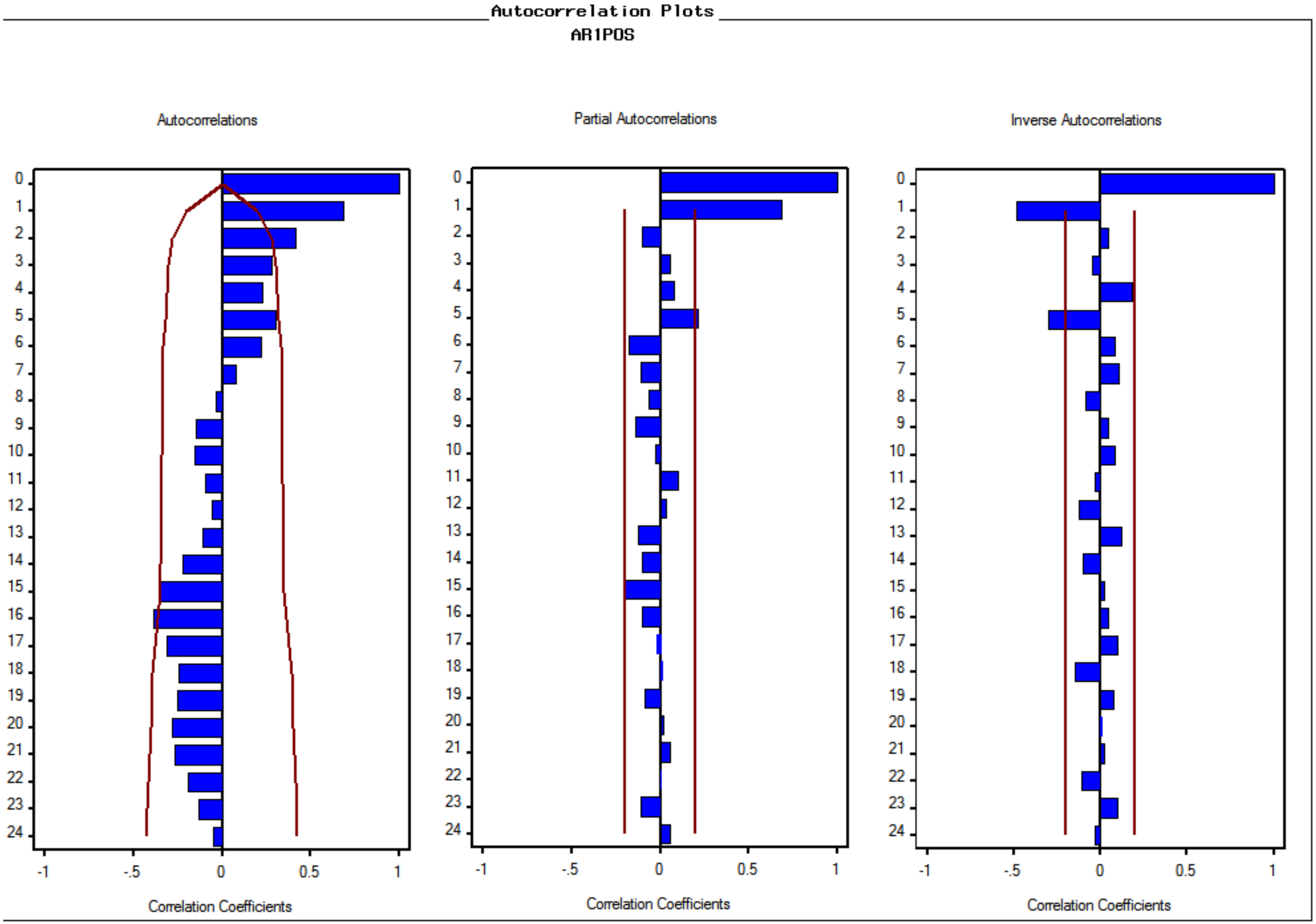
AR(1): $\phi = 0.7$ (1000 obs)



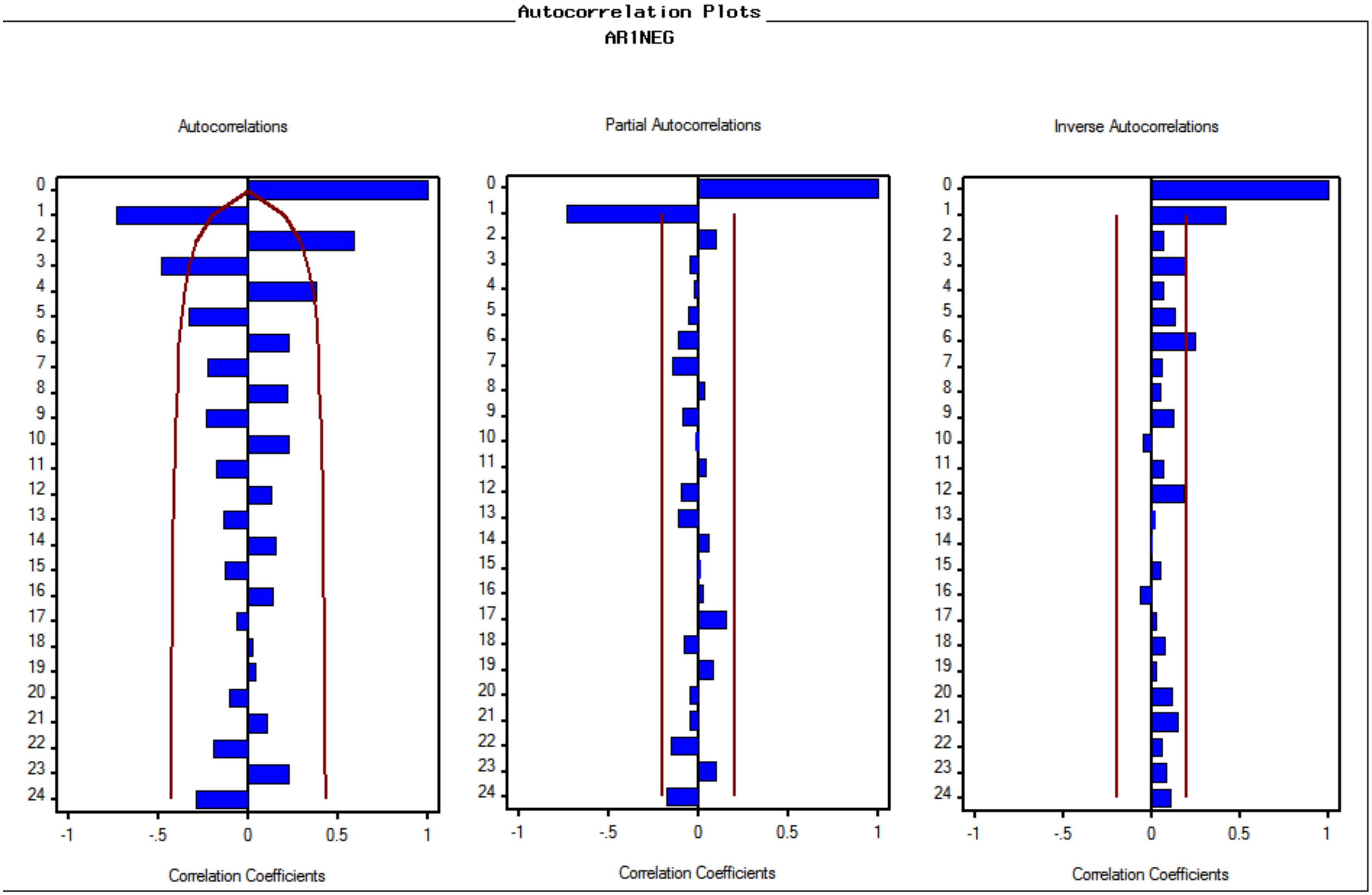
AR(1): $\phi = -0.7$ (1000 obs)



AR(1): $\phi = 0.7$ (100 obs)



AR(1): $\phi = -0.7$ (100 obs)



Exercise

- Examine the autocorrelation plots for different MA models in both simulation100 and simulation1000 datasets

- Examine the autocorrelation plots for white noise in simulation100 dataset

- Copy and paste those figures in the slides

Further Readings

- Forecasting Chapter 1
- https://onlinecourses.science.psu.edu/stat510/