# Time Series Analysis


*By*

*Tanmay Kr. Maity*

*Assistant Professor, Dept. of Statistics,*

*Haldia Govt. College, Vidyasagar University*

# Time Series

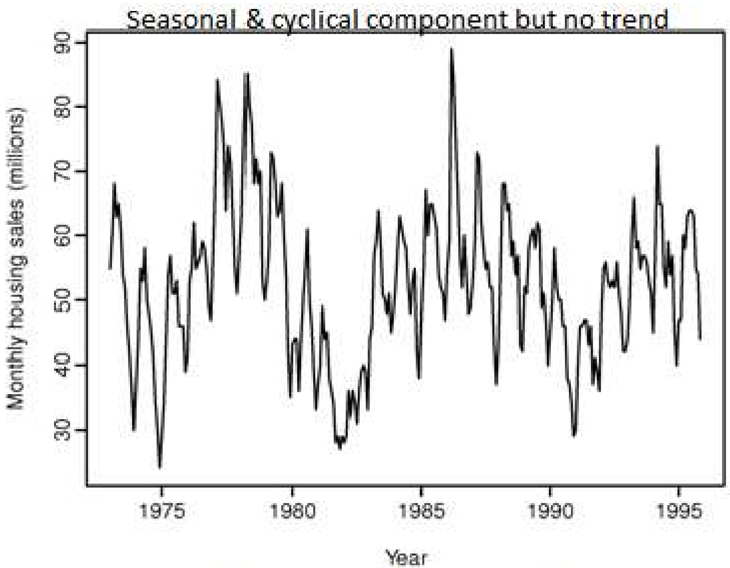
- Time series → A series of observations recorded over time
- Examples:
  – Daily closing price of a stock over several period
  – Turn-over of a firm over a no. of months
  – Sales of a business establishment over no. of weeks
  – Exchange rate observed over certain interval of time
  – Quarterly profit of a company over several quarters

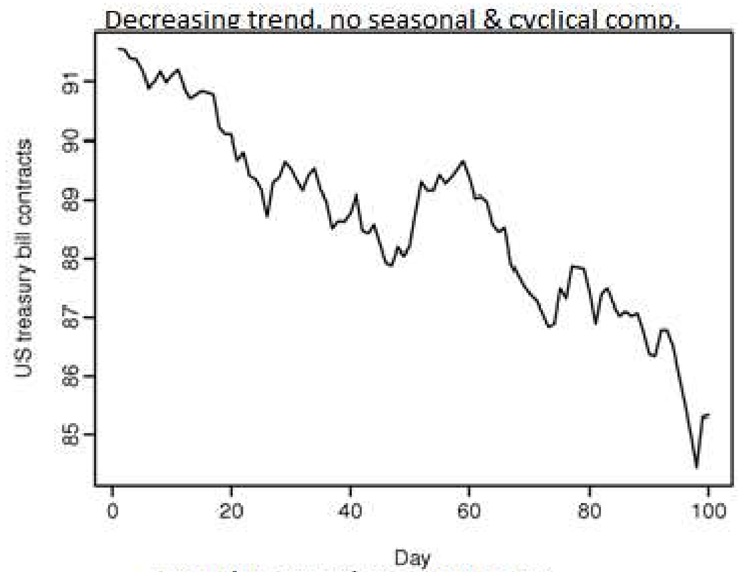
# Component of Time series

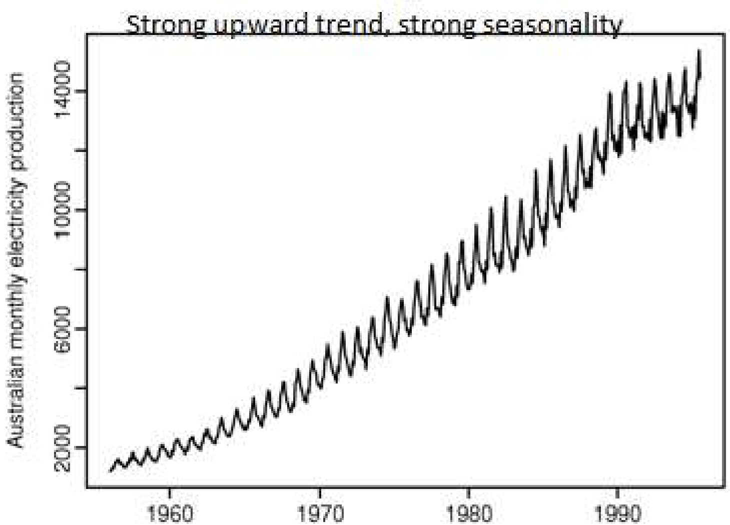
- Systematic component
  - Trend: Smooth, regular, long-term movement of time series
  - Seasonal variation: Periodic movement where the period is not longer than one year
  - Cyclical variation: Oscillatory movement where the period of oscillation being more than a year
- Random or Irregular component

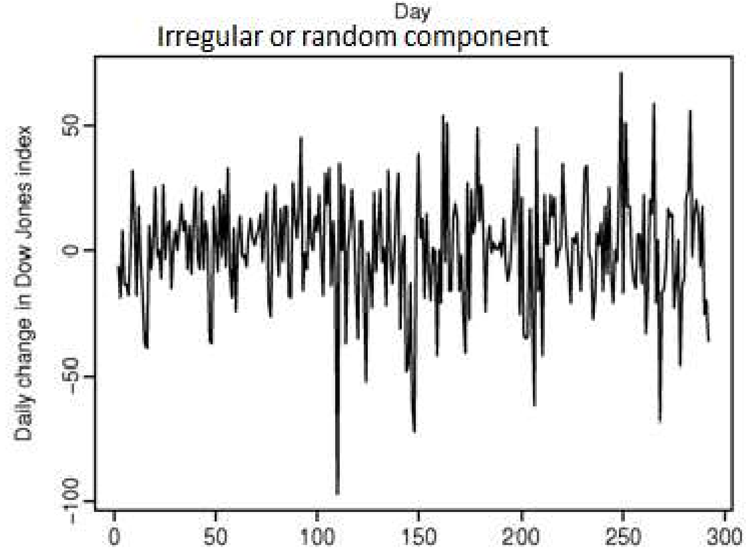
  Purely random, erratic, unforeseen fluctuations due to numerous non-recurring and irregular circumstances such as floods, strikes, earthquakes etc.

# Component of Time series



# Classical models of time series

- Additive model:

$$y_t = T_t + S_t + C_t + \varepsilon_t$$

- Multiplicative model

$$y_t = T_t S_t C_t \varepsilon_t$$

- Mixed models:

$$y_t = T_t C_t + S_t \varepsilon_t$$
$$y_t = T_t + S_t C_t \varepsilon_t$$

$$y_t = T_t + S_t + C_t \varepsilon_t$$

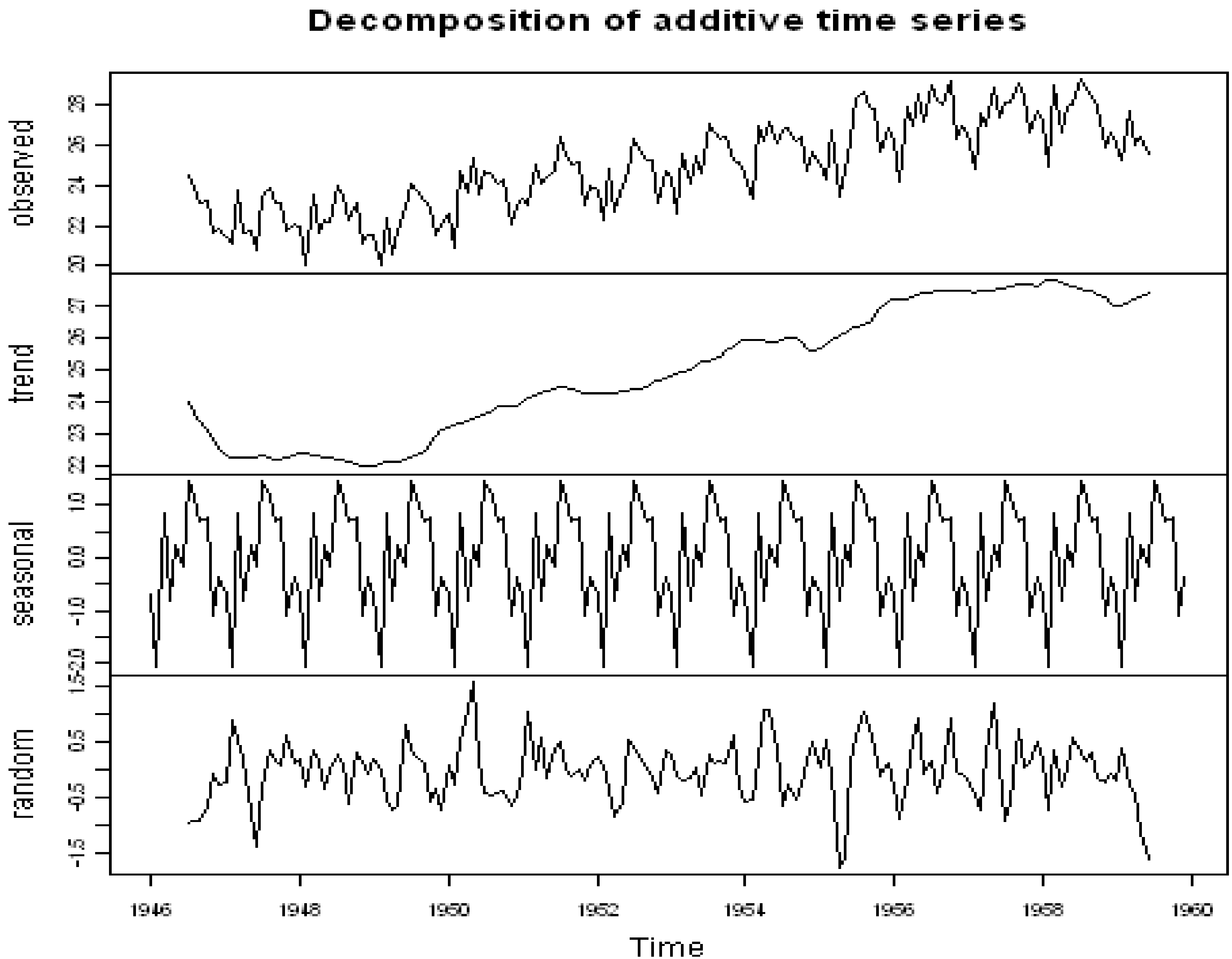
$y_t \to$ Time series at time t

$T_t \to$ Trend at time t

$S_t \to$ Seasonality at time t

$C_t \to$ Cyclicity at time t

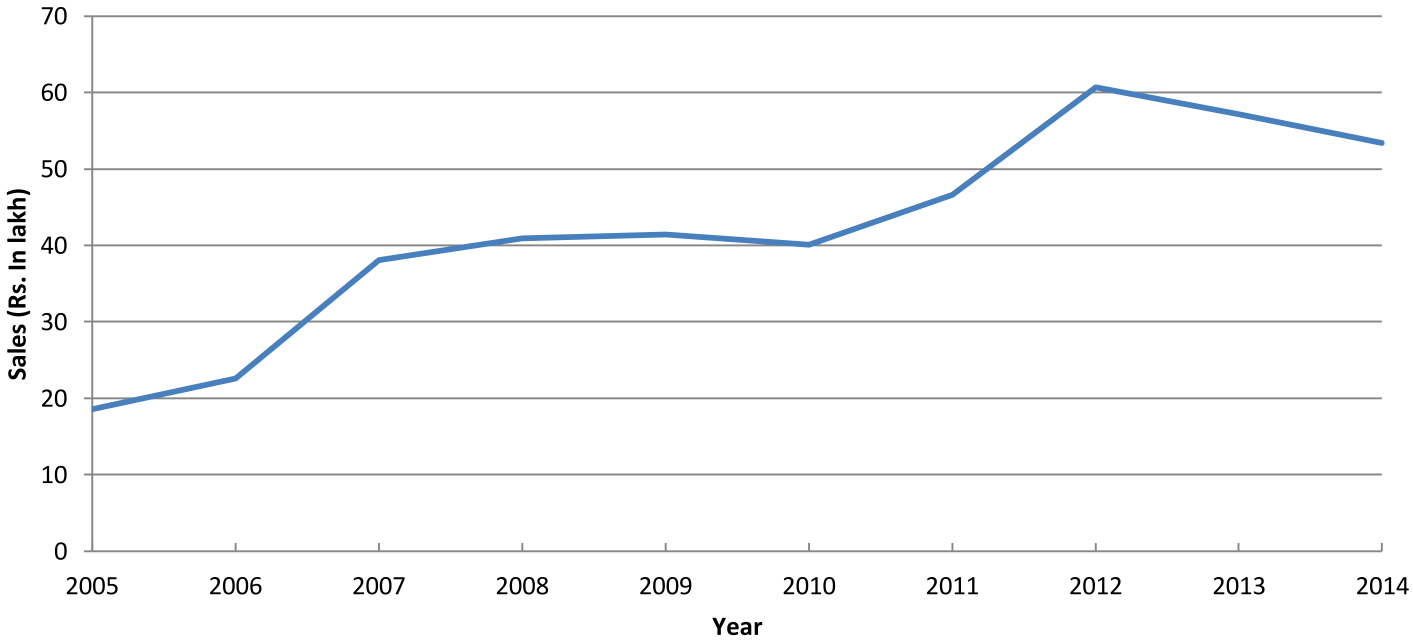
$\varepsilon_t \to$ Random component at time t


Decomposition of additive time series
observed
trend
seasonal
random
Time
1946
1948
1950
1952
1954
1956
1958
1960

# Determination of various components…

## Trend:

- Semi-average method
- Moving average method
- Fitting mathematical curves
- Example: Yearly sales in a shopping mall (2005-2014)



# Trend determination: Semi average method

Increase in trend in 5 years = 51.6-32.32=19.28

Increase in trend in 1 year = 19.28/5 = **3.856**

| Years | Sales (Rs. in lakh) | Semi-Totals | Semi-Average | Trend Values |
|---|---|---|---|---|
| 2005 | 18.6 | | | 28.464 – **3.856** = 24.608 |
| 2006 | 22.6 | | | 32.32 – **3.856** = 28.464 |
| 2007 | 38.1 | 161.6 | 32.32 | 32.32 |
| 2008 | 40.9 | | | 32.32 + **3.856**= 36.176 |
| 2009 | 41.4 | | | 36.176 + **3.856** = 40.032 |
| 2010 | 40.1 | | | 40.032 + **3.856** = 43.888 |
| 2011 | 46.6 | | | 43.888 + **3.856** = 47.744 |
| 2012 | 60.7 | 258 | 51.6 | 51.6 |
| 2013 | 57.2 | | | 51.6 + **3.856** = 55.456 |
| 2014 | 53.4 | | | 55.456 + **3.856** = 59.312 |

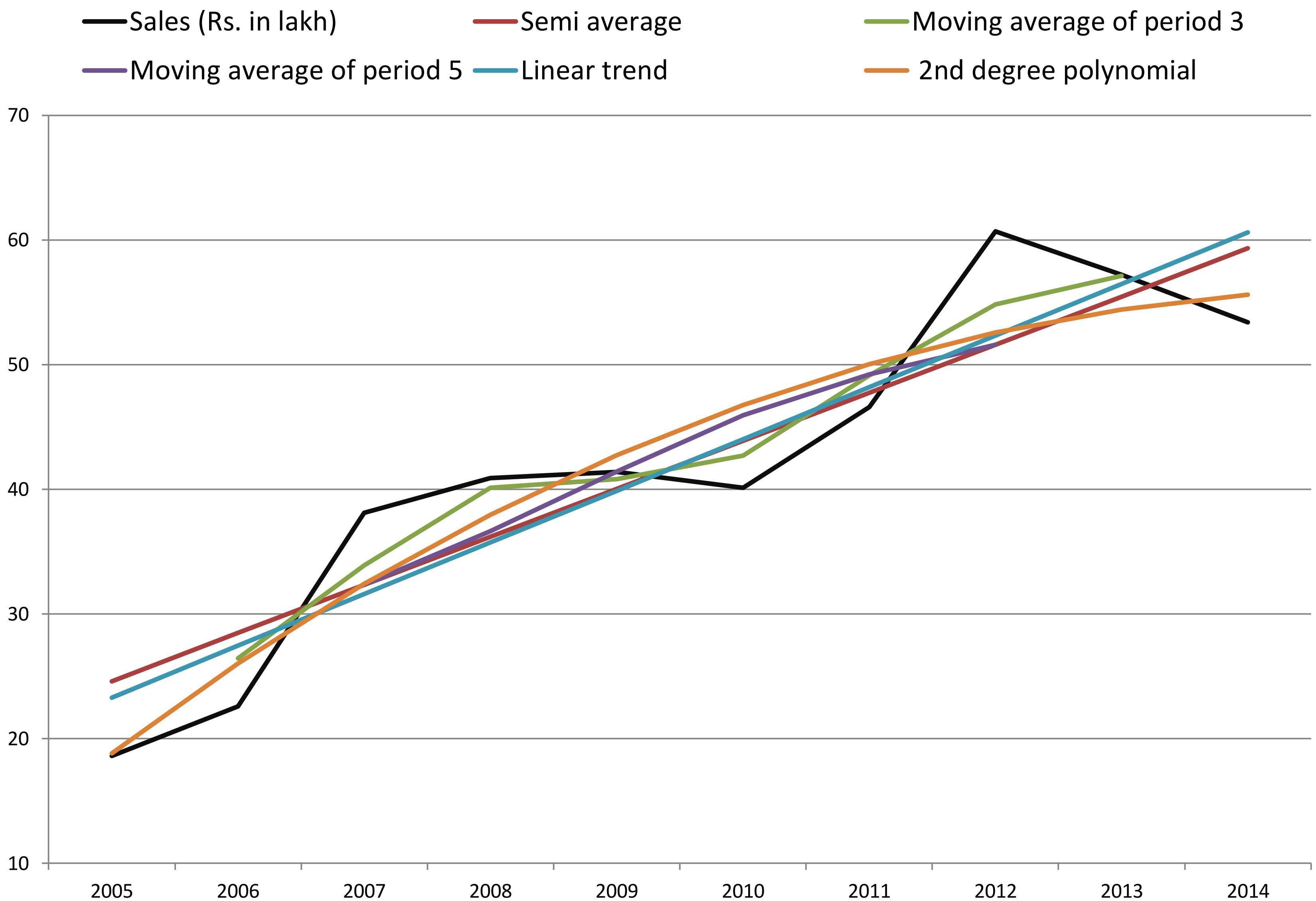
# Trend determination: Moving average method

| Years | Sales (Rs. in lakh) | Moving average of period 3 (Trend value) | Moving average of period 5 (Trend value) |
|---|---|---|---|
| 2005 | 18.6 | | |
| 2006 | 22.6 | 26.43 | |
| 2007 | 38.1 | 33.87 | 32.32 |
| 2008 | 40.9 | 40.13 | 36.62 |
| 2009 | 41.4 | 40.8 | 41.42 |
| 2010 | 40.1 | 42.7 | 45.94 |
| 2011 | 46.6 | 49.13 | 49.2 |
| 2012 | 60.7 | 54.83 | 51.6 |
| 2013 | 57.2 | 57.1 | |
| 2014 | 53.4 | | |

# Trend determination: Method of mathematical curves

- Fitting mathematical curves to the data such as:
  - ✓ Linear curve
  - ✓ Polynomial of order p
  - ✓ Exponential
  - ✓ Power curve

# Trend determination

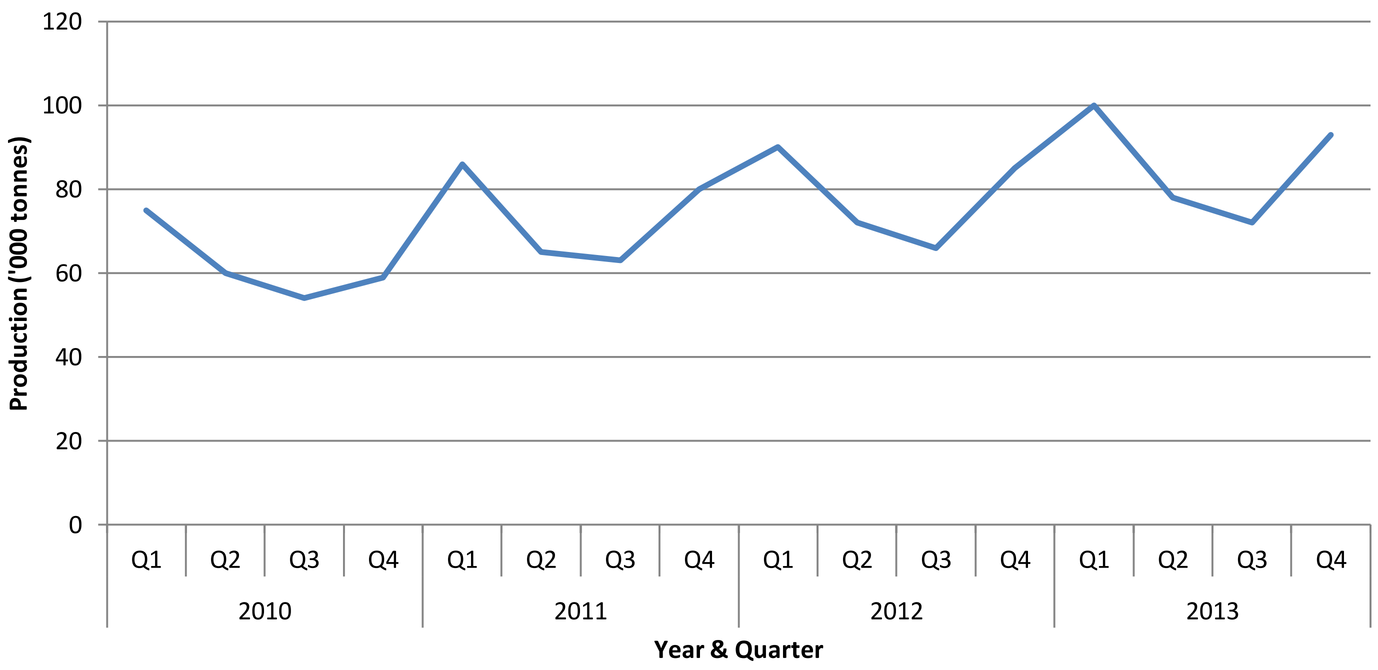


# Determination of various components…

## Seasonal fluctuations:

- Ratio to moving average method
- Ratio to trend method(fitting mathematical curves)

Example: Quarterly production from a factory (2010-2013)



# Ratio to moving average method

- Calculate trend by 4 period moving average (for quarterly data) -> MA(t)

- Trend elimination: R(t) = (y(t)/MA(t)) X 100

  -> Seasonal + Random component

- Seasonal index -> Average the R(t) values across the years

  => Elimination of random component

# Ratio to moving average method

## Computation of seasonal indices:

**Original data:**

| | Quarterly production(in '000 tonnes) | | | |
|---|---|---|---|---|
| | Quarter 1 | Quarter 2 | Quarter 3 | Quarter 4 |
| 2010 | 75 | 60 | 54 | 59 |
| 2011 | 86 | 65 | 63 | 80 |
| 2012 | 90 | 72 | 66 | 85 |
| 2013 | 100 | 78 | 72 | 93 |

**Computation of seasonal indices:**

| | Trend eliminated values | | | | |
|---|---|---|---|---|---|
| | Quarter 1 | Quarter 2 | Quarter 3 | Quarter 4 | |
| 2010 | … | … | 85.21 | 90.25 | |
| 2011 | 128.12 | 91.71 | 85.14 | 106.14 | |
| 2012 | 117.46 | 92.75 | 83.02 | 104.29 | |
| 2013 | 120.48 | 92.04 | … | … | **Total** |
| Average (Elimination of random component) (SI) | 122.02 | 92.17 | 84.45 | 100.23 | 398.87 |
| **Adjusted seasonal indices (SI x k)** | 122.37 | 92.43 | 84.69 | 100.51 | 400.00 |

# Ratio to trend method(fitting mathematical curves)

- Calculate trend by fitting a mathematical curve -> T(t)
- Trend elimination: R(t) = (y(t)/T(t)) X 100

  -> Seasonal + Random component
- Seasonal index -> Average the R(t) values across the years

  => Elimination of random component

# Ratio to trend method

## Computation of seasonal indices:

**Original data:**

| | Quarterly production(in '000 tonnes) | | | |
|---|---|---|---|---|
| | Quarter 1 | Quarter 2 | Quarter 3 | Quarter 4 |
| 2010 | 75 | 60 | 54 | 59 |
| 2011 | 86 | 65 | 63 | 80 |
| 2012 | 90 | 72 | 66 | 85 |
| 2013 | 100 | 78 | 72 | 93 |

**Computation of seasonal indices:**

| | Trend eliminated values | | | | |
|---|---|---|---|---|---|
| | Quarter 1 | Quarter 2 | Quarter 3 | Quarter 4 | |
| 2010 | 119.41 | 93.14 | 81.79 | 87.24 | |
| 2011 | 124.21 | 91.75 | 86.95 | 108.01 | |
| 2012 | 118.93 | 93.17 | 83.66 | 105.60 | |
| 2013 | 121.80 | 93.18 | 84.39 | 106.99 | **Total** |
| Average (Elimination of random component) (SI) | 121.09 | 92.81 | 84.20 | 101.96 | 400.05 |
| **Adjusted seasonal indices (SI x k)** | 121.07 | 92.80 | 84.19 | 101.95 | 400.00 |

# Determination of various components… Cyclical fluctuations:

- Periodogram analysis

Approximate the series by considering it as the superimposition of various periodic curves such as sine and cosine curves at various amplitude and frequency

-A **periodogram** calculates the significance of different frequencies in time-series data to identify any intrinsic periodic signals.

# Stochastic model

- After removing the systematic/deterministic component, the remaining series has been modeled using various stochastic model

For additive model: residual series→

$$r_t = (T_t + S_t + C_t + \varepsilon_t) - (T_t^{'} + S_t^{'} + C_t^{'})$$

For multiplicative model: residual series→

$$r_t = \frac{T_t S_t C_t \varepsilon_t}{T_t^{'} S_t^{'} C_t^{'}}$$

- Conditional mean model
  - **Linear models:**
  - Autoregressive model of order p(AR(p))

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$$

  - Moving average model of order q(MA(q))

$$y_t = \phi_0 + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots + \phi_q \varepsilon_{t-q} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$$

  - ARMA(p, q) model

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + \gamma_1 \varepsilon_{t-1} + \gamma_2 \varepsilon_{t-2} + \ldots + \gamma_q \varepsilon_{t-q} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$$

# Stationary process

- Strong stationarity

  - Joint probability distribution does not change when shifted in time

$$F_Y(y_{t_1}, y_{t_2}, \ldots, y_{t_k}) = F_Y(y_{t_1+h}, y_{t_2+h}, \ldots, y_{t_k+h})$$

# Stationary process…

- Weak stationarity:

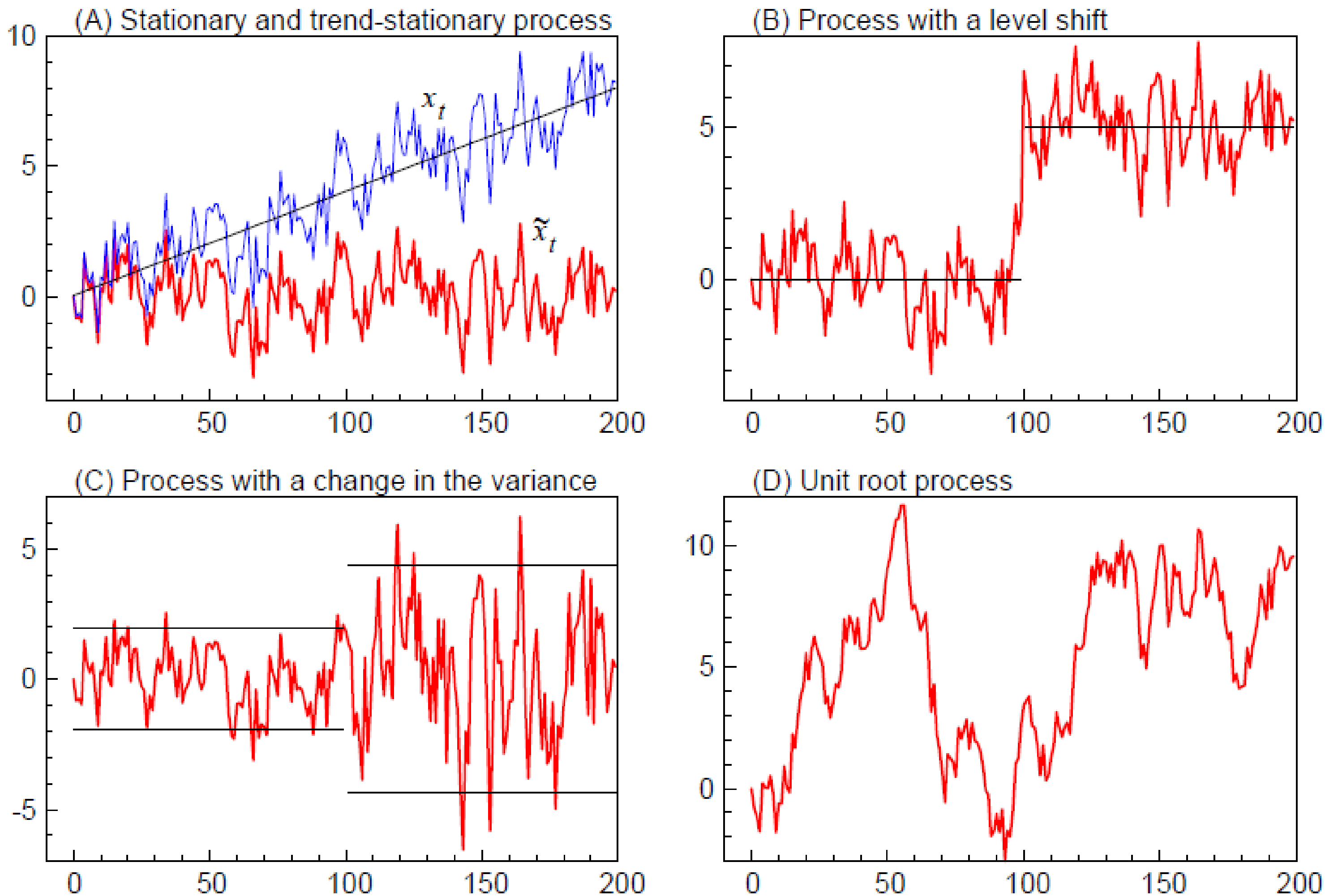
  - $E(Y_t) = \text{Constant}$

  - $Cov(Y_t, Y_{t+h}) = f(h) = \text{Independent of } t$

**Examples:**

1) White Noise is stationary
2) Linear trend process is not stationary
3) Random walk is not stationary
4) ARCH process is not stationary

# Non – stationary time series



# Stochastic model

## Model Identification:

Autocorrelation function(ACF) of order k:

$$r(k) = Corr(y(t), y(t+k))$$

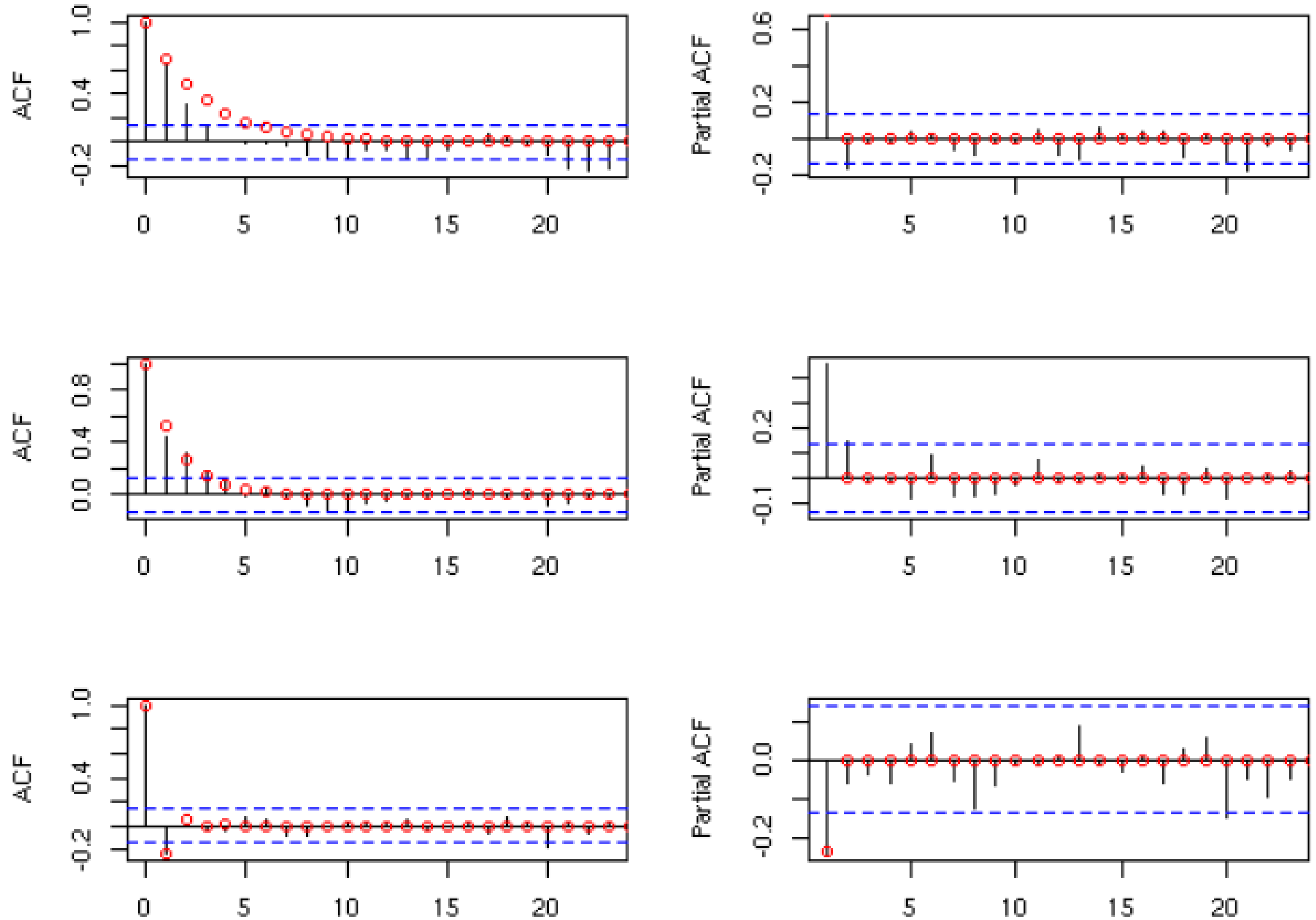
Partial autocorrelation function(PACF) of order k:

$$\alpha(k) = Corr(y(t), y(t+k) \mid y(t+1), \ldots, y(t+k-1))$$

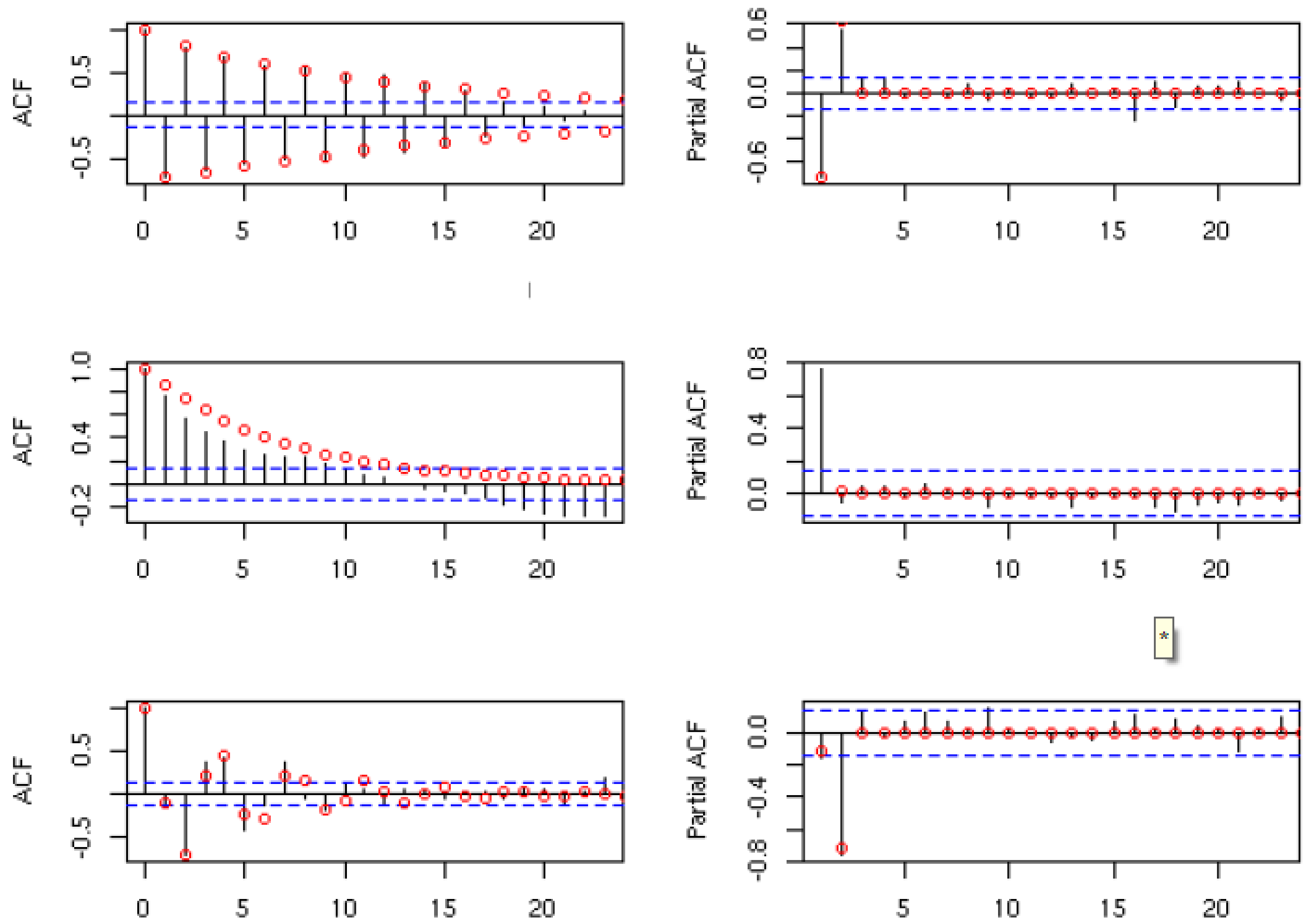
**Behavior of ACF and PACF for ARMA model:**

| Model | ACF | PACF |
|---|---|---|
| White Noise | All zeros | All zeros |
| AR(p) | Exponential Decay | p significant lags before dropping to zero |
| MA(q) | q significant lags before dropping to zero | Exponential Decay |
| ARMA(p,q) | Decay after qth lag | Decay after pth lag |

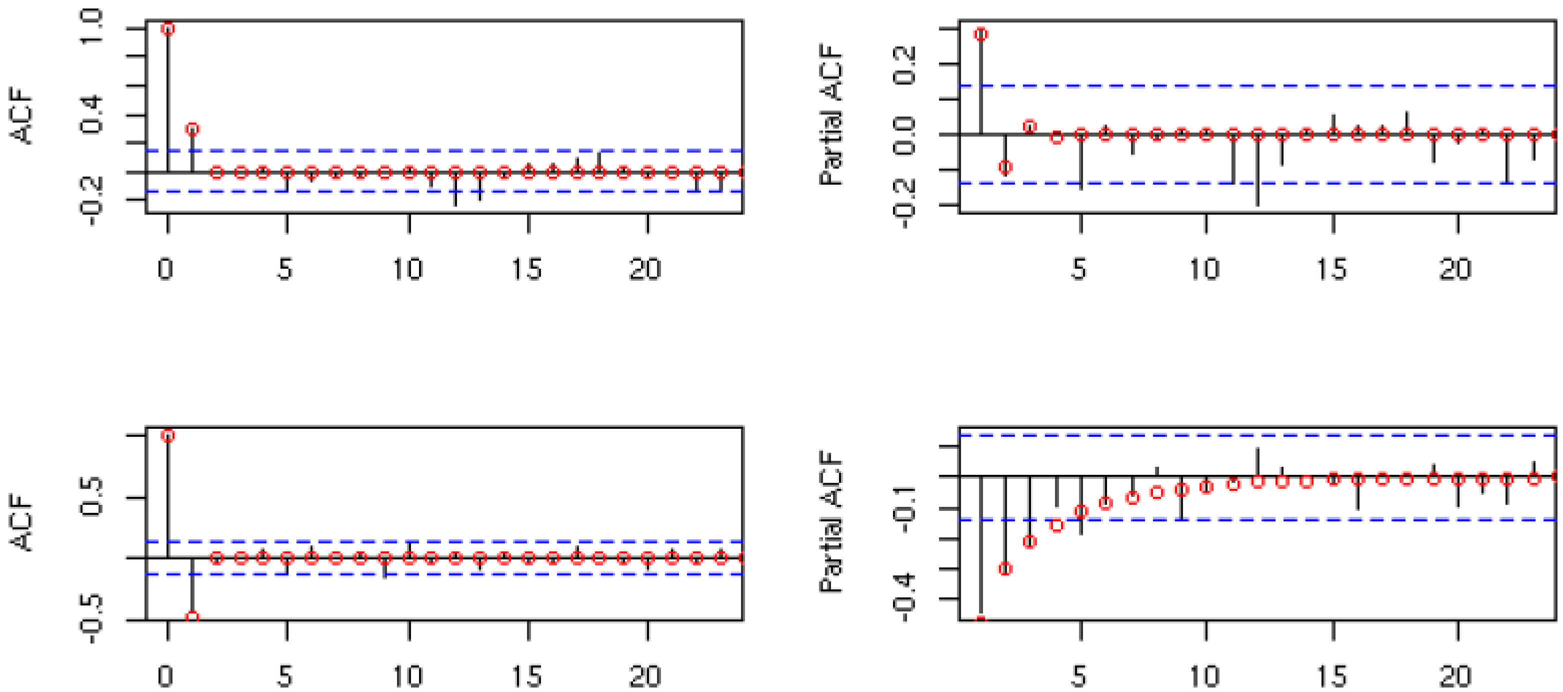
# ACF and PACF for an AR(1) process



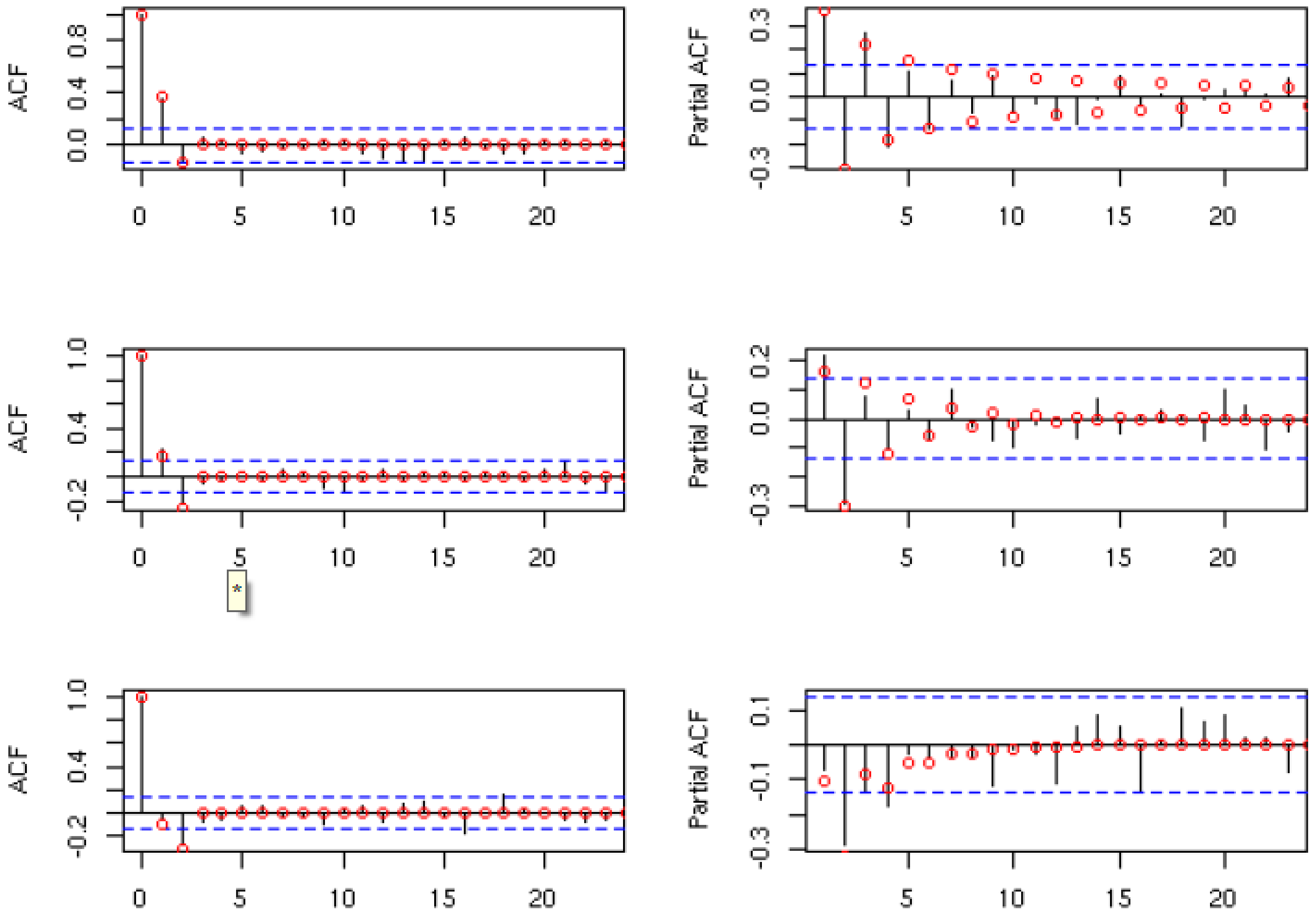
# ACF and PACF for an AR(2) process



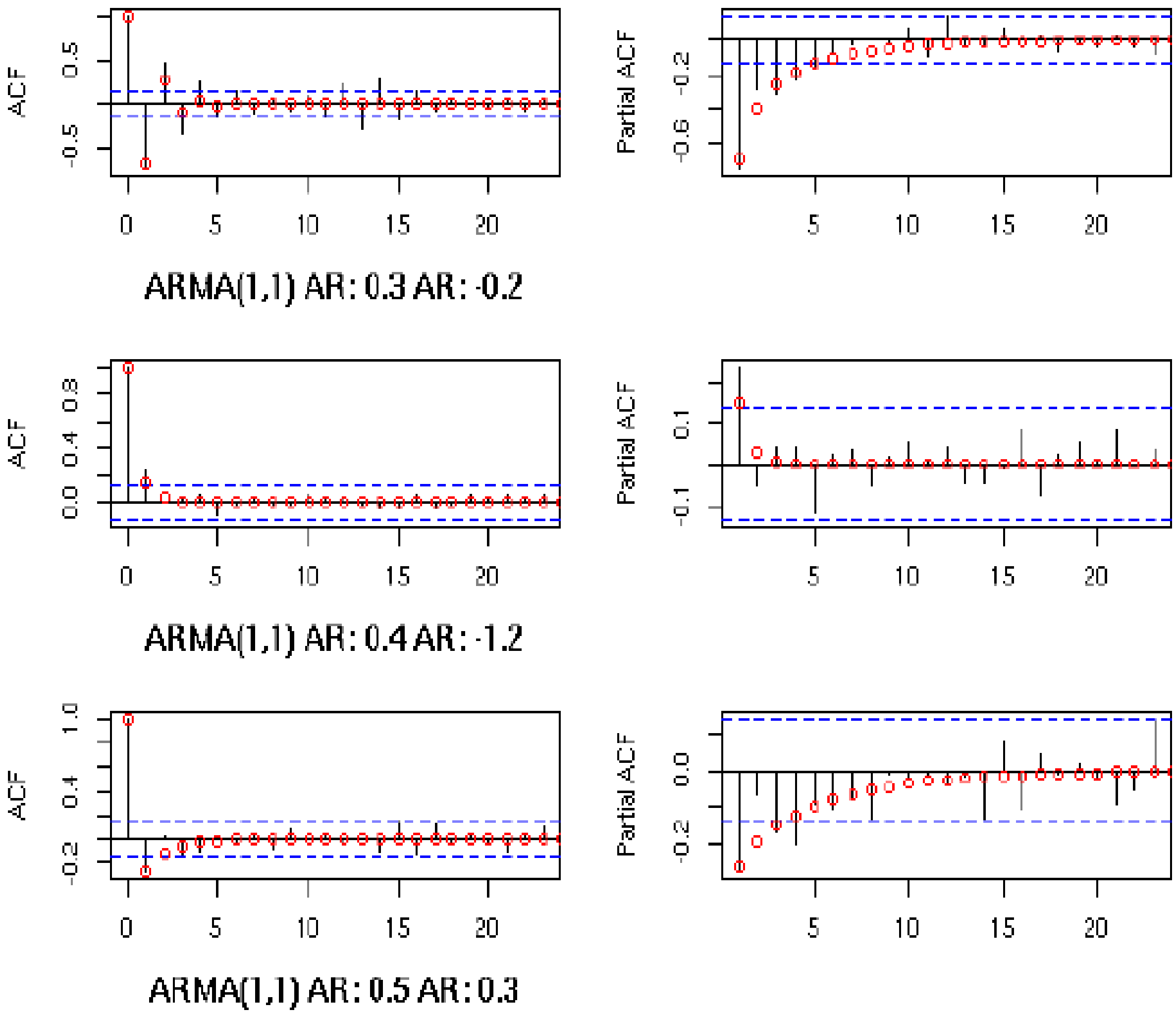
# ACF and PACF for an MA(1) process



# ACF and PACF for an MA(2) process



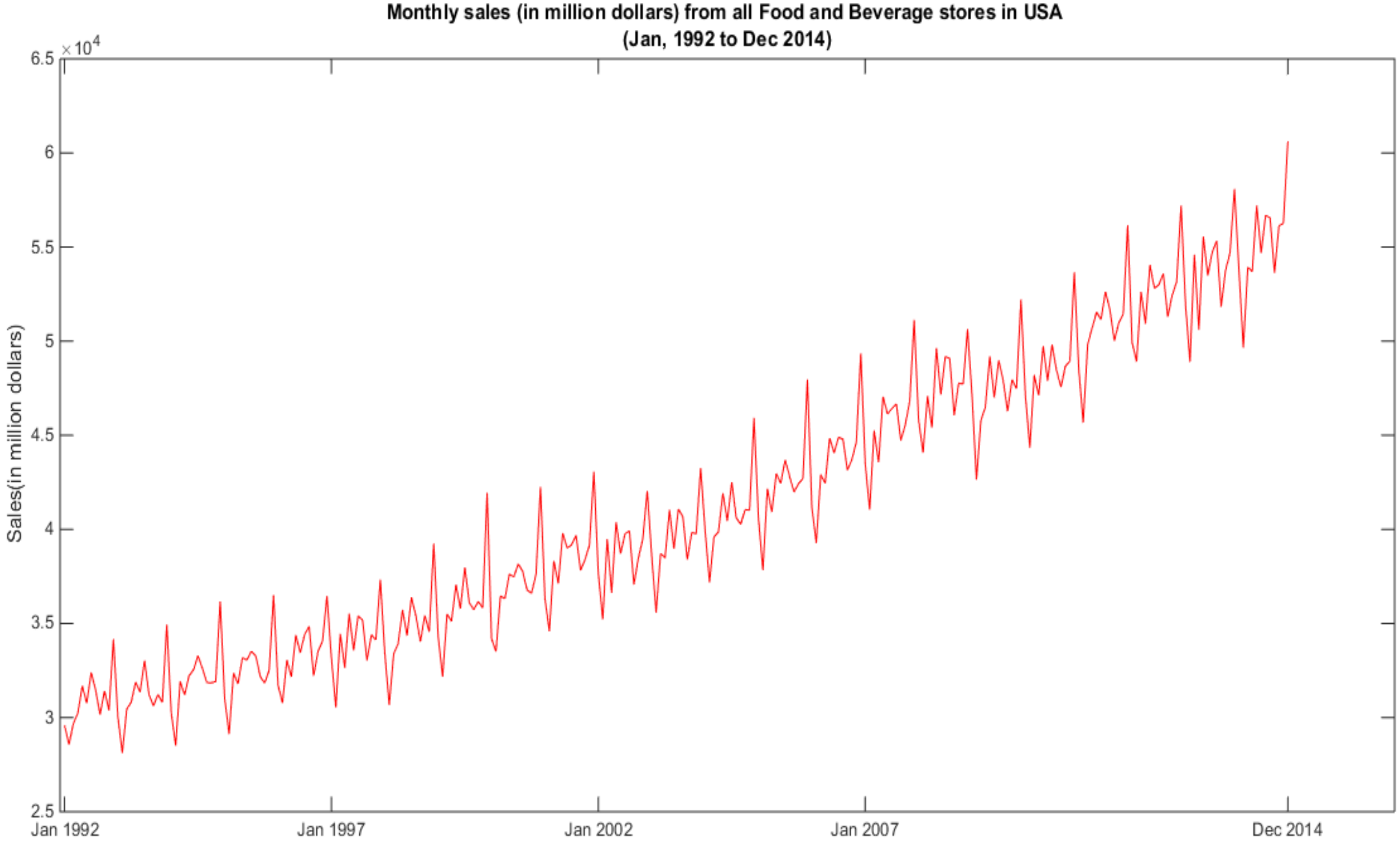
# ACF and PACF for an ARMA(1,1) process



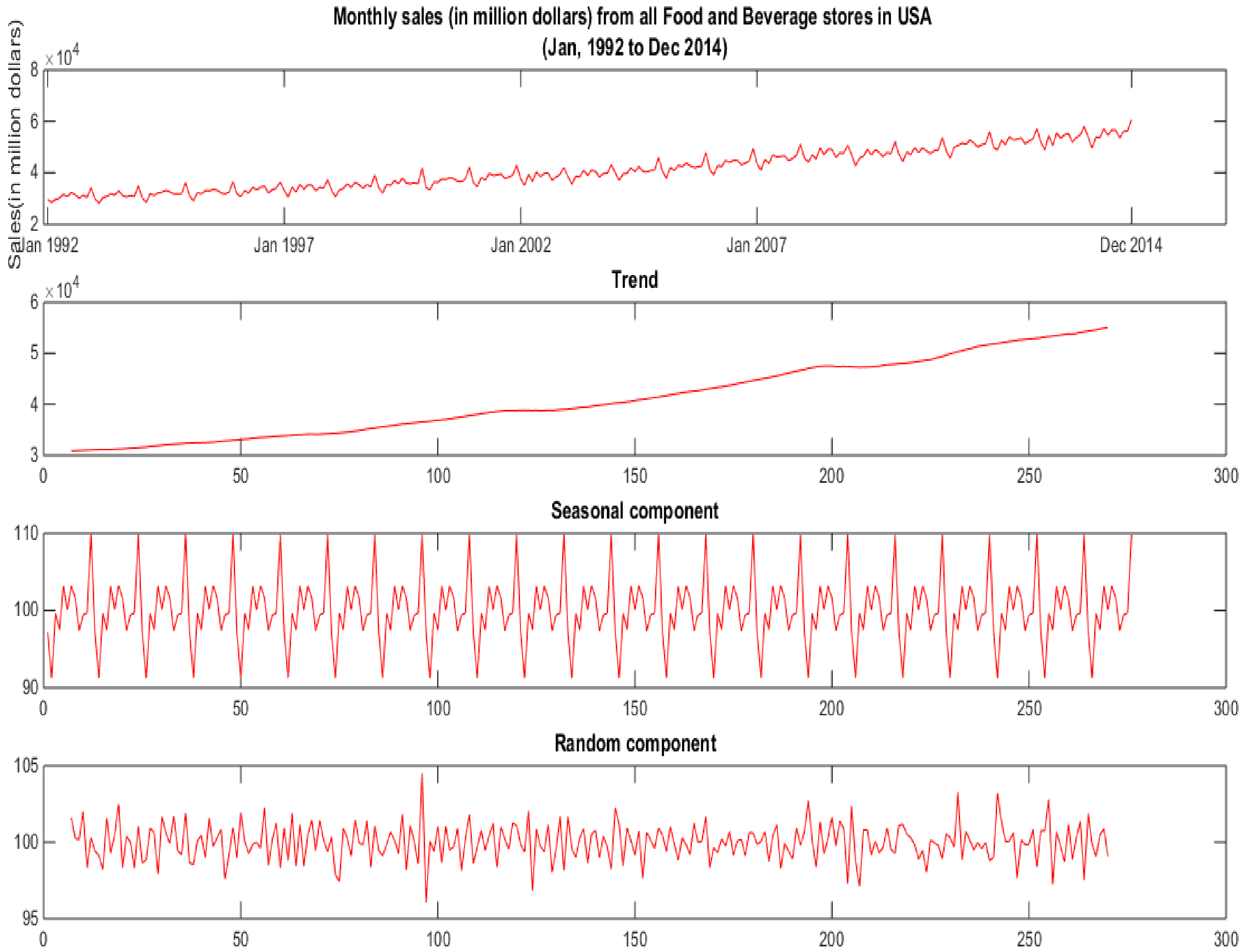
# Time series modelling: An Example

## Dataset used

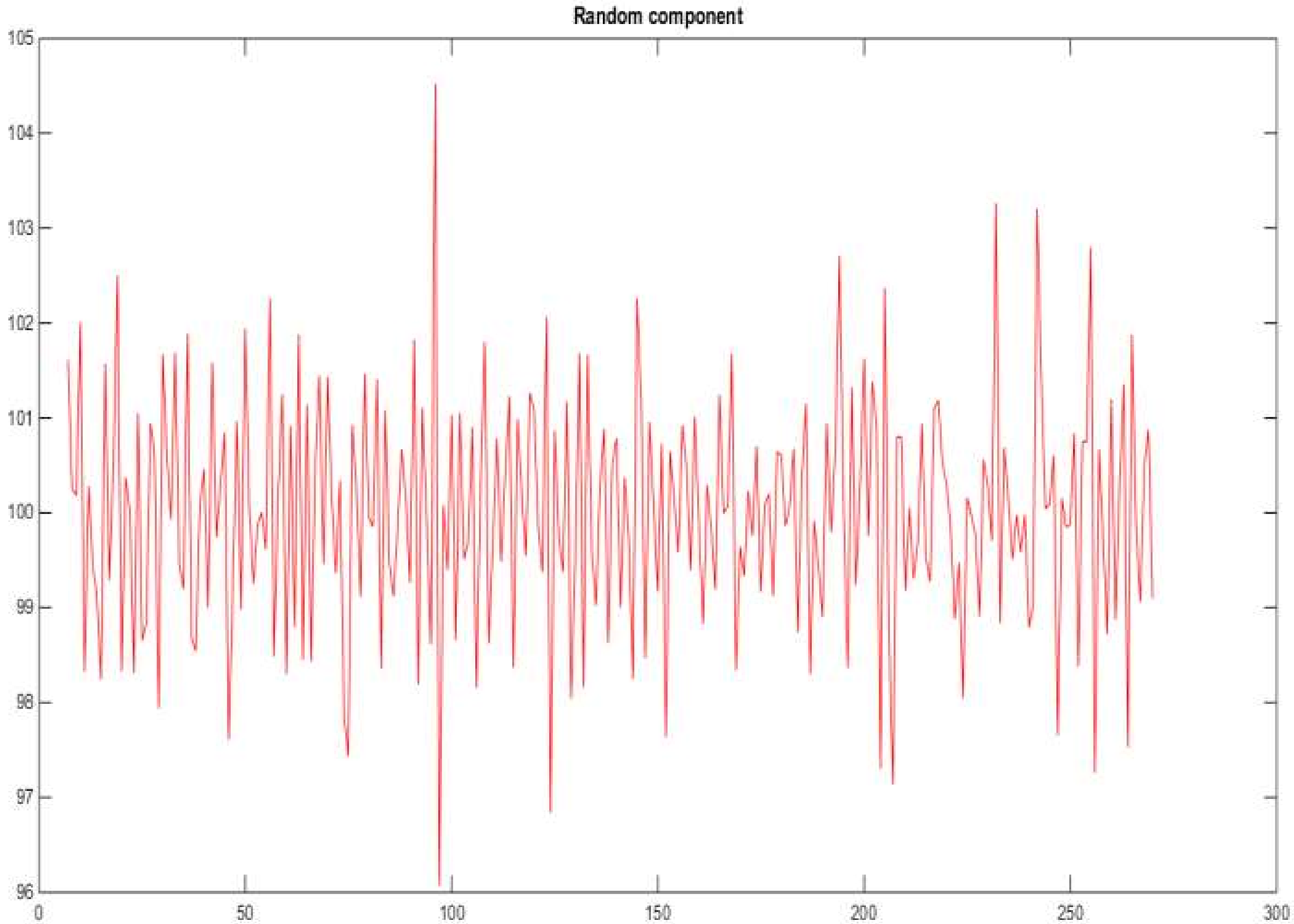
Monthly sales (in million dollars) from all Food and Beverage stores in USA from Jan, 1992 to Dec 2014



# Sales data: Decomposition in different components



# Sales data: Work with the random component



# Random component: ACF & PACF