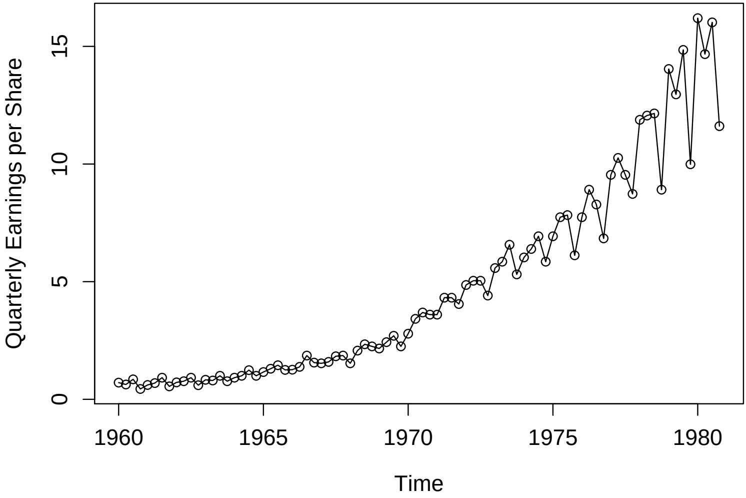
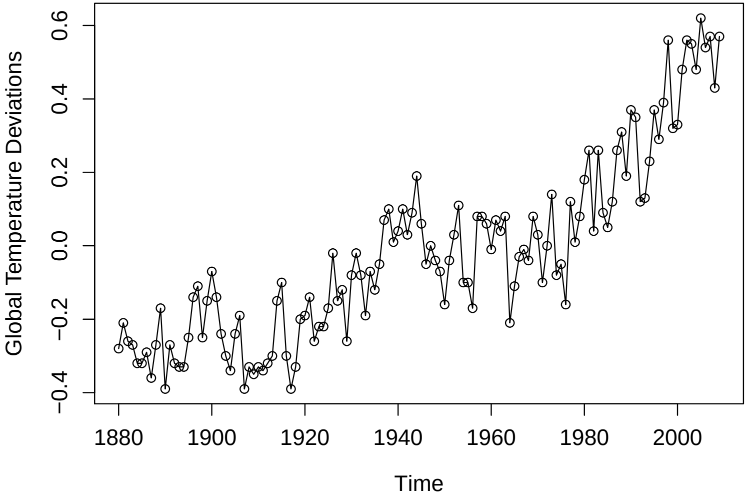
Lecture 18. Time Series

Nan Ye

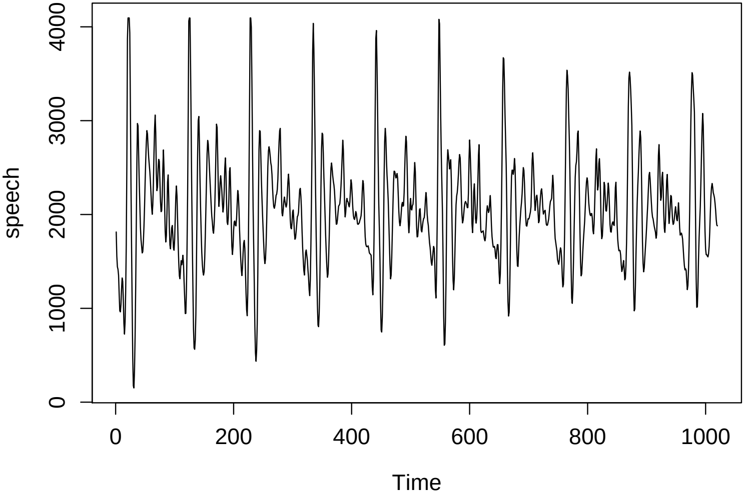
School of Mathematics and Physics
University of Queensland



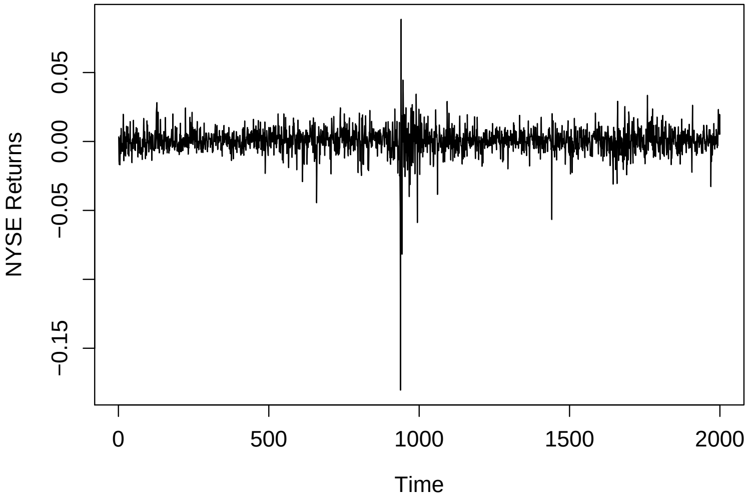
Johnson & Johnson quarterly earnings per share (1960-I to 1984-IV)



Yearly average global temperature (1880-2009)



Speech recording of the syllable aaa...hhh



Returns from NYSE from 2 Feb 1984 to 31 Dec 1991.

This Lecture

- Nature of time series data
- Time series modelling
- Time series decomposition
- Stationarity
- Time domain models

Nature of Time Series Data

Characteristics

- A time series is often viewed as a collection of random variables $\{X_t\}$ indexed by time.
- In a time series, measurements at adjacent time points are usually correlated.
- As compared to other types of correlated data, such as clustered or longitudinal data, observations in a time series may explicitly depend on previous observations and/or errors.

Probabilistic description

- We can describe a time series using the distribution of the random variables $\{X_t\}$.
- Frequently, we look at some summary statistics

$$\begin{aligned} \text{Mean function} \quad & \mu_X(t) = \mathbb{E}(X_t) \\ \text{Autocovariance function} \quad & \gamma_X(s,t) = \mathrm{cov}(X_s, X_t) \\ \text{Autocorrelation function (ACF)} \quad & \rho_X(s,t) = \frac{\gamma_X(s,t)}{\sqrt{\gamma_X(s,s)\gamma_X(t,t)}}. \end{aligned}$$

- We often drop X from μ_X, γ_X and ρ_X when there is no ambiguity.

Time Series Modelling

Chasing after stationarity

- The objective of time series modelling is to develop compact representation of the time series, to facilitate tasks including interpretation, prediction, control, hypothesis testing and simulation.
- Some form of time invariance is required to find regularity in data and extrapolate into future.
- *Stationarity* is a basic form of time invariance, and much of time series modelling is about transforming times series so that the transformed time series is stationary.

Exploratory data analysis

- Plotting the time series should be the first step before any formal modelling attempt.
- This is useful for identifying important features for choosing an appropriate model.
- For example, use the plots to look for the trends, presence of seasonal components or outliers.

Modelling paradigms

- There are two main modelling paradigms.
- The time domain approach views a time series as the description of the evolution of an entity, and focuses on capturing the dependence of current value on history.
- The frequency domain approach views a time series as the superposition (addition) of periodic variations.

Time Series Decomposition

An additive decomposition

- A classical decomposition of a time series $\{X_t\}$ is

$$X_t = T_t + S_t + E_t,$$

  where T_t is the trend component, S_t is the seasonal component (recurring variation with fixed period), E_t is the error component.
- The trend component and seasonal component can be deterministic or stochastic.
- Sometimes, a cyclical component (recurring variation with non-fixed period) is included in the systematic component.

A multiplicative decomposition

- A common multiplicative decomposition is

$$X_t = T_t S_t E_t.$$

- This is equivalent to first converting X_t to the log scale and then perform an additive decomposition

$$\ln X_t = \ln T_t + \ln S_t + \ln E_t$$

Stationarity

Strict stationarity

- A time series $\{X_t\}$ is strictly stationary if its probabilistic behavior is invariant to time shift.
- To be precise, for any k, for any time points $t_1, \ldots, t_k$, for any $x_1, \ldots, x_k$, and for any δ, we have

$$P(X_{t_1} \leq x_1, \ldots, X_{t_k} \leq x_k) = P(X_{t_1+\delta} \leq x_1, \ldots, X_{t_k+\delta} \leq x_k)$$

- Strict stationarity implies that the mean function $\mu(t) = \mathbb{E}(X_t)$ and the autocovariance function $\gamma(t, t+h) = \text{cov}(X_t, X_{t+h})$ do not depend on t.
- Strict stationarity is often too much to ask for, and usually not necessary for learning a model.

Stationarity

- A time series $\{X_t\}$ is said to be (weakly) stationary if $\mu(t)$ and $\gamma(t, t+h)$ do not depend on t.
- The autocovariance and autocorrelation functions of a stationary time series can be more compactly written as

$$
\begin{aligned}
\gamma(h) &= \gamma(t, t+h), \\
\rho(h) &= \rho(t, t+h) = \gamma(h)/\gamma(0).
\end{aligned}
$$

- Randomness in a stationary time series is sufficiently constrained for finding out regularity in data.
- A stationary time series has a trivial system component (constant mean).
- Stationary time series are used as an important building block for the random component of more sophisticated models.

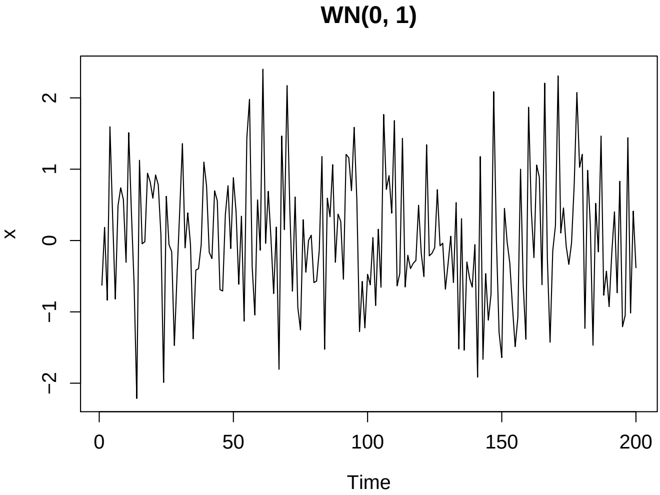
Time Domain Models

White noise processes

- A white noise process $\{\epsilon_t\}$ is a collection of uncorrelated random variables with mean 0 and finite variance σ^2.
- This is often denoted as $\epsilon_t \sim \mathrm{WN}(0, \sigma^2)$.
- The mean, autocovariance and autocorrelation functions are

$$\begin{aligned} \mu(t) &= 0 \\ \gamma(t, t+h) &= \mathrm{cov}(\epsilon_t, \epsilon_{t+h}) = \begin{cases} \sigma^2, & h = 0, \\ 0, & h \neq 0. \end{cases} \\ \rho(t, t+h) &= \begin{cases} 1, & h = 0, \\ 0, & h \neq 0. \end{cases} \end{aligned}$$

- White noise processes are thus stationary, and they serve as an important building block for more sophisticated time series models.



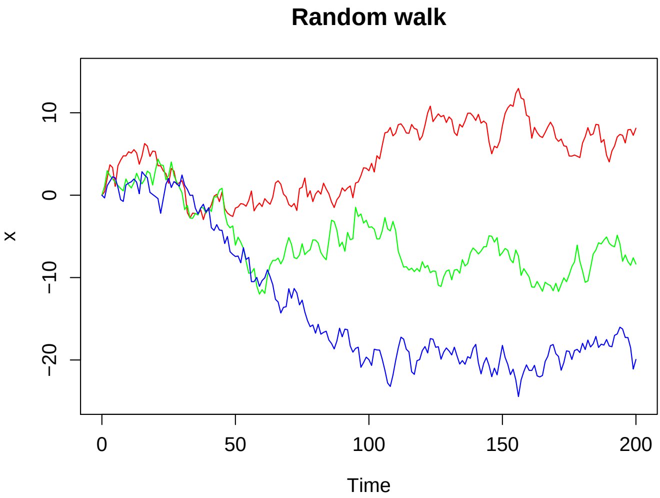
An example white noise series.

Random Walk

- Consider the random walk $X_t = \sum_{i=1}^{t} \epsilon_i$, where $\epsilon_t \sim \text{WN}(0, \sigma^2)$.
- The mean, autocovariance, and autocorrelation functions are

$$\begin{aligned} \mu(t) &= 0, \\ \gamma(t, t+h) &= t\sigma^2, \\ \rho(t, t+h) &= \frac{t}{\sqrt{t(t+h)}}. \end{aligned}$$

- $\{X_t\}$ is not stationary.



Three random walk series from the same model.

Moving average process

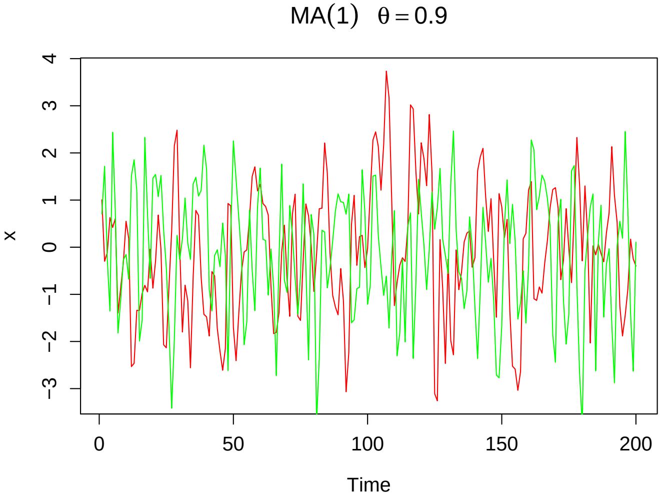
- $\{X_t\}$ is a moving average process of order 1, or MA(1), if $X_t = \epsilon_t + \theta\epsilon_{t-1}$, where $\epsilon_t \sim \mathrm{WN}(0, \sigma^2)$.
- The mean, autocovariance, and autocorrelation functions are

$$\mu(t) = 0,$$

$$\gamma(t, t+h) = \begin{cases} \sigma^2(1+\theta^2), & h = 0, \\ \sigma^2\theta, & h = \pm 1, \\ 0, & \text{otherwise}, \end{cases},$$

$$\rho(t, t+h) = \begin{cases} 1, & h = 0, \\ \theta/(1+\theta^2), & h = \pm 1, \\ 0, & \text{otherwise}, \end{cases}.$$

- MA(1) is stationary.

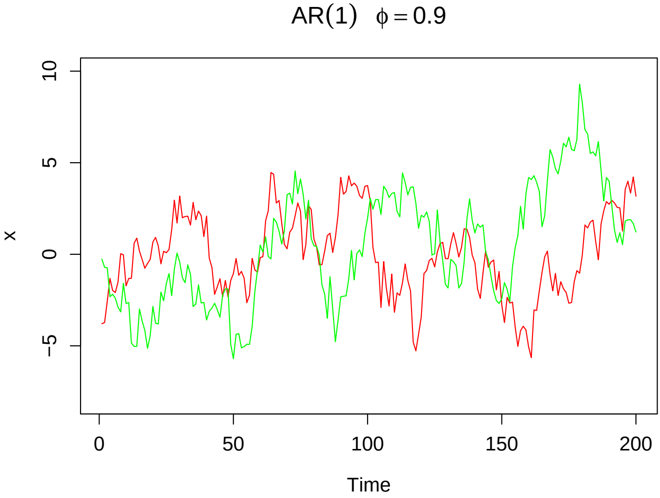


Two MA(1) series from the same model.

Autoregressive process

- $\{X_t\}$ is an autoregressive process of order 1, or AR(1), if $X_t = \phi X_{t-1} + \epsilon_t$, where $\epsilon_t \sim \mathrm{WN}(0, \sigma^2)$.
- When AR(1) is stationary, the mean, autocovariance and autocorrelation functions are

$$\mu(t) = 0,$$
$$\gamma(t, t+h) = \frac{\phi^{|h|}\sigma^2}{1-\phi^2},$$
$$\rho(t, t+h) = \phi^{|h|}.$$



Two AR(1) series from the same model.

Linear processes

- A linear process $\{X_t\}$ is a linear combination white noise variates ϵ_t, that is,

$$X_t = \mu + \sum_{i=-\infty}^{+\infty} \psi_i \epsilon_{t-i},$$

  where $\epsilon_t \sim \mathrm{WN}(0, \sigma^2)$.

- The mean and covariance functions are

$$\mu(t) = \mu,$$

$$\gamma(t, t+h) = \sigma^2 \sum_{i=-\infty}^{\infty} \psi_i \psi_{i+h}.$$

- White noise is a linear process with $\mu = 0$ and $\psi_i = \begin{cases} 1, & i = 0, \\ 0, & i \neq 0. \end{cases}$.
- MA(1) is a linear process with $\mu = 0$, and $\psi_i = \begin{cases} 1, & i = 0, \\ \theta, & i = 1, \\ 0, & \text{otherwise.} \end{cases}$.
- AR(1) is a linear process with $\mu = 0$, and $\psi_i = \begin{cases} \phi^i, & i \geq 0, \\ 0, & \text{otherwise.} \end{cases}$.

ARMA

- $\{X_t\}$ is ARMA(p, q) if it is stationary and

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_q \epsilon_{t-q},$$

  where $\phi_p \neq 0$, $\theta_q \neq 0$ and $\epsilon_t \sim \mathrm{WN}(0, \sigma^2)$.
- p and q are called the autoregressive and the moving average orders respectively.
- AR$(1) =$ ARMA$(1, 0)$, and MA$(1) =$ ARMA$(0, 1)$.

What You Need to Know

- Time series exhibits correlation between measurements and serial dependence.
- Time series modelling requires some form of time invariance.
- Time series decomposition is helpful for understanding the underlying patterns.
- Stationarity is a basic form of time invariance.
- Time domain models.