STAT 6104 - Time Series

Chapter 2 - Probability Models

Agenda

Introduction

- Assume that we have done the preliminary analysis to estimate trend T_t and seasonality S_t
- What remains unexplained is the noise N_t:

$$N_t = X_t - T_t - S_t$$

- We propose statistical models to describe the noise $\{N_t\}$ to extract more information

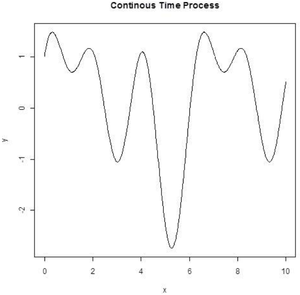
Agenda

Agenda

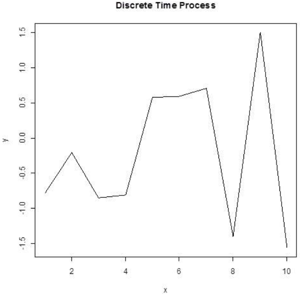
Definition 1– Stochastic Process

- A collection of random variables $\{X_t : t \in R\}$ is called a stochastic process, or random function

- Continuous Time
  $\{X(t) : 0 \leq t < \infty\}$



- Discrete Time
  $\{X_t : t = 1, 2, ..., n\}$



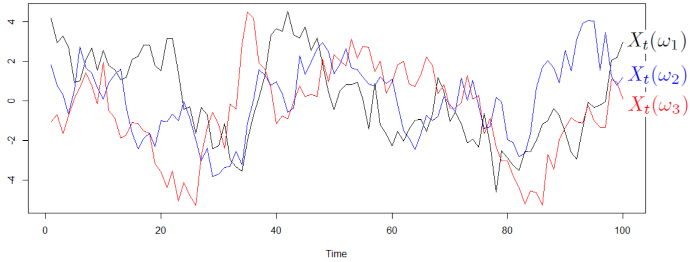
- Discrete Time Stochastic Process is commonly known as Time Series

Definition 1 – Stochastic Process

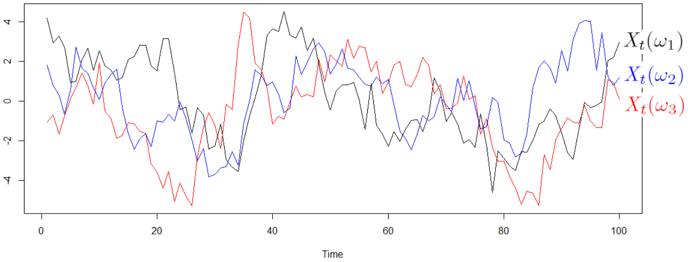
Understand the meaning of stochastic (Random):

- We assume that there is a sample space Ω (a set of all possible scenarios) that generates the randomness
  - $\Omega = (\omega_1, \omega_2, \omega_3, \ldots\ldots)$ (an ∞-face die)

  $\Rightarrow$ $\omega_1 \to \{X_t(\omega_1)\}_{t\geq 0}$; $\omega_2 \to \{X_t(\omega_2)\}_{t\geq 0}$ (each face gives one path)
  - For a fixed ω,
    $\{X_t(\omega)\}_{t\geq 0}$ is called a sample function/ realization/ sample path
  - For a fixed t,
    $X_t(\cdot)$ is a random variable



Definition 1 – Stochastic Process



- In practice, we can only observe one sample path (i.e. the situation is generated by one particular ω)
- To summarize important features from the data, we require the process to be repeating itself in some way (stationarity).

Agenda

Definition 2 – Finite dimensional distribution function

- Finite dimensional distribution functions:

$$F_{\mathbf{t}}(\mathbf{x}) = P\left(X_{t_1} \leq x_1, ..., X_{t_n} \leq x_n\right)$$

for all possible $\mathbf{t} = (t_1, t_2, ..., t_n)$ where, $1 \leq t_1 < t_2 < \cdots < t_n \leq \infty$ and
$\mathbf{x} = (x_1, x_2, ..., x_n) \in R^n, n = 1, 2, ...$

Agenda

Definition 3 – Strictly Stationary

- Strictly stationary - distribution of a process does not change
- $\{X_t\}$ is said to be strictly stationary if
  - for all n,
  - for all $(t_1, t_2, ..., t_n)$,
  - for all h,

$$(X_{t_1}, ..., X_{t_n}) \stackrel{d}{=} (X_{t_1+h}, ..., X_{t_n+h}) \ ,$$

  where "$\stackrel{d}{=}$" means "equal in distribution", i.e.,

$$F_{\mathbf{t}}(\mathbf{x}) = P\left(X_{t_1} \leq x_1, ..., X_{t_n} \leq x_n\right) = P\left(X_{t_1+h} \leq x_1, ..., X_{t_n+h} \leq x_n\right) = F_{\mathbf{t}+h}(\mathbf{x})$$

- May not be easy to check in practice because joint distribution $F_{\mathbf{t}}(\mathbf{x})$ is difficult to compute.

Agenda

Definition 4 – Weakly Stationary

- Weakly Stationary:
  - same mean
  - same variance and same covariance

  across time.
- Formally, $\{X_t\}$ is weakly stationary/ second order stationary/ wide-sense stationary if
  - 1) $\mathrm{E}(X_t) = \mu$
  - 2) $\mathrm{Cov}(X_t, X_{t+h}) = \gamma(h)$ for all t and h.
    (auto-covariance only depends on time lag h , but not time t)
- Easier to verify weakly stationary than strictly stationary

Agenda

Definition 5 – Autocovariance function

- Autocovariance function (ACVF)

$$\gamma(h) = \mathrm{Cov}(X_t, X_{t+h})$$

- Autocorrelation function (ACF)

$$\begin{aligned} \rho(h) &= \frac{\mathrm{Cov}(X_t, X_{t+h})}{\sqrt{\mathrm{Var}(X_t)\mathrm{Var}(X_{t+h})}} \\ &= \frac{\gamma(h)}{\gamma(0)}, \end{aligned}$$

since $\gamma(0) = \mathrm{Var}(X_t) \overset{\text{Stationary}}{=} \mathrm{Var}(X_{t+h})$.

Agenda

Strictly vs Weakly Stationary

- Strictly: distribution of the process unchange
  Weakly: mean/covariance unchange

- Strictly stationary implies weakly stationary whenever $E(X_t^2) < \infty$

- Weakly stationary implies strictly stationary when $\{X_t\}$ is normal ,but in general not true

Strictly vs Weakly Stationary

- Basic Statistics: If $X \sim t_{df}$, then $E(X^k) < \infty$ if and only if $k < df$.

1. $X_1, X_2, ..., X_n \overset{iid}{\sim} t_2$
   - $\{X_t\}$ is strictly stationary
   - But $E(X_t^2) = \infty$,

     $\therefore \{X_t\}$ is **NOT** weakly stationary
2. $X_1, X_2, ..., X_n \overset{iid}{\sim} t_3$
   - Now $E(X_t^2) < \infty$,

     $\therefore \{X_t\}$ is both strictly and weakly stationary
3. $X_1 \sim \exp(1) - 1, X_2, X_3, ... \overset{iid}{\sim} N(0, 1)$
   - $\{X_t\}$ is Weakly stationary:

     $\mathrm{E}(X_t) = 0$, $\mathrm{Var}(X_t) = 1$, $\mathrm{Cov}(X_t, X_{t+k}) = 0$, all t and $k \neq 0$
   - But $\{X_t\}$ is **NOT** Strictly stationary, since $X_1 \overset{d}{\neq} X_2$.

Agenda

Example 1: White Noise

If a time series $\{X_t\}$ is a sequence of uncorrelated random variables with

$$\rho(h) = \begin{cases} 1\,, & h = 0 \\ 0\,, & |h| \geq 1\,, \end{cases} \tag{1}$$

with zero mean and constant variance, then it is called a **white noise** sequence.

- zero-mean i.i.d. random variables are white noises if their second moments exist
- white noise *may not be* i.i.d.
  $(\text{e.g.} X_1 \sim \exp(1) - 1, X_k \sim N(0,1), k \geq 2)$
- independence is a stronger assumption than zero-correlation

Example 2: Same Observation

If $\{Y_t\}$ satisfies $Y_1 = Y_2 = \cdots = Y$, then

- $\rho(h) = \mathrm{Corr}(Y, Y) = 1$ for all h
- Essentially there is only one observation

Example 3: Periodic Function

- Suppose a time series $\{X_t\}$ satisfies

$$X_t = A\cos\theta t + B\sin\theta t\,,$$

where A and B are i.i.d. random variables with mean 0 & variance σ^2.

- $\{X_t\}$ is weakly stationary because
  - $\mathrm{E}(X_t) = \mathrm{E}(A)\cos\theta t + \mathrm{E}(B)\sin\theta t = 0$
  - $$\begin{aligned} & \mathrm{Cov}(X_{t+h}, X_t) \\ = \; & \mathrm{Cov}(A\cos\theta(t+h) + B\sin\theta(t+h), A\cos\theta t + B\sin\theta t) \\ = \; & \cos\theta t\ \cos\theta(t+h)\mathrm{Var}(A) + \sin\theta t\ \sin\theta(t+h)\mathrm{Var}(B) \\ = \; & \sigma^2\cos(\theta(t+h) - \theta t) \quad (\cos(A-B) = \cos A\cos B + \sin A\sin B) \\ = \; & \sigma^2\cos\theta h\,, \end{aligned}$$

    independent of t.

Example 4: Moving Average MA(1) Model

Let $\{\epsilon_t\}$ be a white noise sequence and $X_t = \epsilon_t + 0.5\epsilon_{t-1}$. Is $\{X_t\}$ stationary? Find the ACVF and ACF of $\{X_t\}$.

Example 5: Moving Average MA(∞) Model

Let $\{\epsilon_t\}$ be a white noise sequence and $X_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}$ and $\sum_{k=0}^{\infty} \psi_k^2 < \infty$. Is $\{X_t\}$ stationary? Find the ACVF and ACF of $\{X_t\}$.

Agenda

Sample ACVF and ACF

$$\begin{aligned} \text{ACVF} &: \gamma(h) = \mathrm{Cov}(X_t, X_{t+h}) \\ \text{ACF} &: \rho(h) = \frac{\mathrm{Cov}(X_t, X_{t+h})}{\sqrt{\mathrm{Var}(X_t)\mathrm{Var}(X_{t+h})}} = \frac{\gamma(h)}{\gamma(0)} \end{aligned}$$

- Both ACVF and ACF are *population quantities*. It means that they involve the unknown true distribution of the process $\{X_t\}$
- If we have a dataset $X_1, X_2, ..., X_n$ but we do not know the true probability distribution of the time series, we can estimate the ACF/ACVF by
  - Sample ACVF: $C_h = \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \overline{X})(X_{t+h} - \overline{X})$
  - Sample ACF : $r_h = C_h/C_0$

Sample ACVF and ACF

- Sample ACVF: $C_h = \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \overline{X})(X_{t+h} - \overline{X})$
- Why $\frac{1}{n}$ but not $\frac{1}{n-h}$? (Optional)

- To ensure the **non-negative definiteness** of covariance matrix

$$\widehat{\mathrm{Var}}((X_1, \ldots, X_n)) = \left(C_{|i-j|}\right)_{i,j\in\{1,\ldots,n\}} = \frac{1}{n}MM',$$

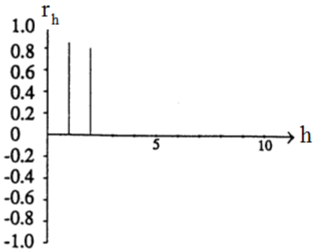
where with $\tilde{X}_k = X_k - \overline{X}$,

$$M = \begin{pmatrix} \tilde{X}_1 & \tilde{X}_2 & \tilde{X}_3 & \cdots & \tilde{X}_n & 0 & 0 & 0 \\ 0 & \tilde{X}_1 & \tilde{X}_2 & & \tilde{X}_{n-1} & \tilde{X}_n & 0 & 0 \\ 0 & 0 & \tilde{X}_1 & & \tilde{X}_{n-2} & \tilde{X}_{n-1} & \tilde{X}_n & 0 \\ & & & \ddots & & & & \\ & & & & \tilde{X}_1 & \tilde{X}_2 & & \tilde{X}_n \end{pmatrix}, M' = \begin{pmatrix} \tilde{X}_1 & 0 & 0 & 0 \\ \tilde{X}_2 & \tilde{X}_1 & 0 & 0 \\ \tilde{X}_3 & \tilde{X}_2 & \tilde{X}_1 & 0 \\ \vdots & & \ddots & \vdots \\ \tilde{X}_n & \tilde{X}_{n-1} & \tilde{X}_{n-2} & \tilde{X}_1 \\ 0 & \tilde{X}_n & \tilde{X}_{n-1} & \tilde{X}_2 \\ & & & \vdots \\ 0 & 0 & \tilde{X}_n & \vdots \\ 0 & 0 & 0 & \tilde{X}_n \end{pmatrix}.$$

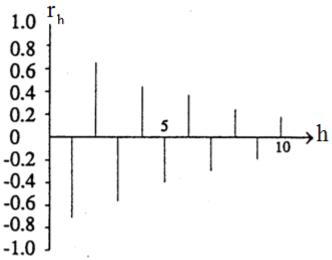
- Denote $\boldsymbol{X} = (X_1, \ldots, X_n)$, $\widehat{\mathrm{Var}}(\boldsymbol{X})$ is non-negative definite guarantees that $\widehat{\mathrm{Var}}(\boldsymbol{a}'\boldsymbol{X}) = \boldsymbol{a}'\widehat{\mathrm{Var}}(\boldsymbol{X})\boldsymbol{a} = \frac{1}{n}\boldsymbol{a}'MM'\boldsymbol{a} = \frac{1}{n}(\boldsymbol{a}'M)(\boldsymbol{a}'M)' \geq 0$.

ACF plot

ACF plot: A graph to summarize r_h (`acf(x)` in `R`)



Example 1



Example 2

- By definition, $r_0 = \frac{C_0}{C_0} = 1$
- Trust r_h up to $h = n/3$, otherwise r_h is not accurate

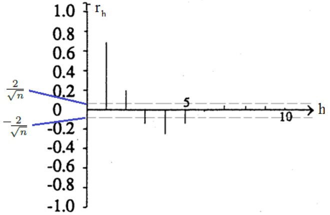
Sample Correlation Function

- Test for time dependence:
  - $H_0 : \rho(h) = \mathrm{Corr}(X_t, X_{t+h}) = 0$
  - Idea: check if r_h is significantly different from 0

Theorem 1

If $\{X_t\}$ is white noise, then $r_h \sim N(0, \frac{1}{n})$ for each $h \geq 1$ when n is large.

- Proof of this result is complicated
- Conclude time dependence exists if r_hs exceed horizontal lines at levels $\pm\frac{2}{\sqrt{n}}$

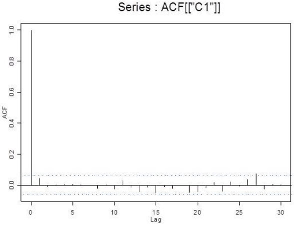


Example-ACF plot

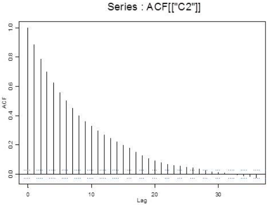
Sketch the sample acf plot for the dataset $(2, 5, 3, 5, 6, 3, 4)$. Is this time series a white noise sequence?

More examples of ACF function

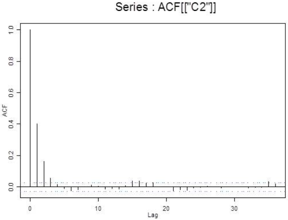
- $Y_t = Z_t, \{Z_t\}$ are i.i.d.



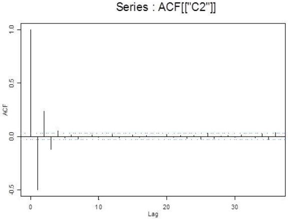
- $Y_t = 0.9Y_{t-1} + Z_t$



- $Y_t = 0.4Y_{t-1} + Z_t$

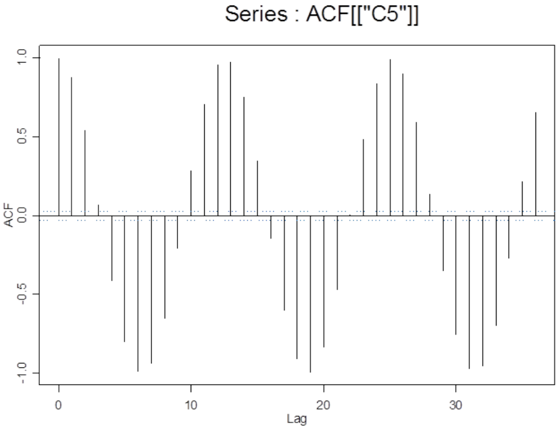


- $Y_t = -0.5Y_{t-1} + Z_t$



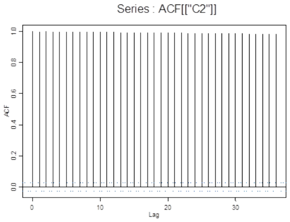
More examples of ACF function

- $Y_t = a\cos(t\omega)$, where a, ω are constants

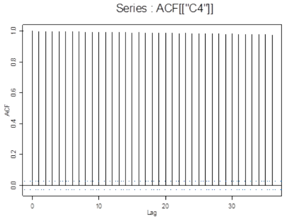


More examples of ACF function

- $Y_t = Y_{t-1} + Z_t$



- $Y_t = at + Z_t$



1. In both cases, ACF decay very slowly
2. Commonly observed in **non-stationary** process
3. In practice, perform detrending/filtering and study the resulting stationary noises

Agenda

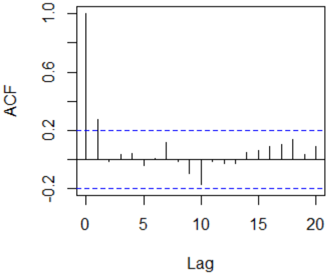
Summary of Chapter 2

- Strictly Stationary: Joint distribution remain unchanged
- Weakly Stationary: Mean and covariance structure remain unchanged
- White Noise: $E(X_t) = 0$, $\mathrm{Cov}(X_t, X_k) = \sigma^2 1_{\{t=k\}}$

$$
\begin{aligned}
\text{ACVF} \quad &: \gamma(h) &&= \mathrm{Cov}(X_t, X_{t+h}) \\
\text{ACF} \quad &: \rho(h) &&= \frac{\mathrm{Cov}(X_t, X_{t+h})}{\sqrt{\mathrm{Var}(X_t)\mathrm{Var}(X_{t+h})}} = \frac{\gamma(h)}{\gamma(0)} \\
\text{Sample ACVF} \quad &: C_h &&= \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \overline{X})(X_{t+h} - \overline{X}) \\
\text{Sample ACF} \quad &: r_h &&= \frac{C_h}{C_0}
\end{aligned}
$$

Summary of Chapter 2

ACF plot (ACF r_h against lag h):



Check if $\rho(h) = \mathrm{Corr}(X_t, X_{t+h}) = 0$:

- $\{X_t\}$ is white noise $\Rightarrow r_h \sim N(0, \frac{1}{n})$ for all $h \geq 1$
- Draw horizontal lines at levels $\pm\frac{2}{\sqrt{n}}$ on acf and check for exceedance.

- Stationary
- Non-stationary