INTRODUCTION

JEN-WEN LIN, PhD

DATE: June 21, 2014

Course information

TA's OH
Thursday 12pm–3pm Room TBA

Time/place:	• Monday and Wednesday 0600-900 pm, SS2118
Office hours:	• 05:20-06:00 pm, SS6025 (Monday and Wednesday) • will ask for your answer before answering the question
Email:	• uofttimeseries@gmail.com (the best way to contact me) • jenwen@utstat.toronto.edu
Textbook:	• Shumway & Stoffer (2010), Time Series Analysis and Its Applications: With R Examples (Springer Texts in Statistics)
Course materials:	• MS PowerPoint slides will be posted on portal • More references will be provided as the class progresses

Time series books

- **C. Chatfield**
  - The Analysis of Time Series: An Introduction, Sixth Edition (Chapman & Hall/CRC Texts in Statistical Science)
- **W. Enders**
  - Applied Econometric Time Series, Wiley
- **W.S. Wei**
  - Time Series Analysis : Univariate and Multivariate Methods
- **Box et al**
  - Time Series Analysis: Forecasting and Control (Wiley Series in Probability and Statistics)
- **J.D. Hamilton**
  - Time Series Analysis
- **Brockwell and Davis**
  - Time Series: Theory and Methods (Springer Series in Statistics)

Marking schemes and rules

1. Midterm test (40%) and final exam (60%).
2. No makeup exam for the midterm test. If students miss the midterm test with a legitimate reason, they may shift the weight of the midterm test to their final exam.
3. No pencil in the exam.

Exams (tentative)

Midterm test:

- EX 200, July 21, 201~~3~~4 with calculator
- 0600-0800pm (Two hours exam)

Final exam:

- TBD, two hours exam
- Past exams will be posted in portal

Course schedule (tentative)

#	July/August	Day	Topic	Remark
1	2	W	Introduction	
2	7	M	Box Jenkins approach	
3	9	W	Box Jenkins approach	
4	14	M	Box Jenkins approach	
5	16	W	ARIMA and forecasts	
6	**21**	**M**	**Midterm test**	
7	23	W	Transfer function noise model	
8	**28**	**M**	**Forecast evaluation**	**Return midterm**
9	30	W	Nonlinear time series/multivariate time series	
10	**4**	**M**	**No class**	**Civic day**
11	6	W	State space model and Kalman filter	
12	11	M	State space model and Kalman filter	

Time series and statistical analysis

Most statistical theories are established based on the assumption of random experiments.

- The order of the occurrences among observations does not matter!!

Time series analysis considers that successive observations are usually not independent, and takes into account the time order of the occurrences among observations.

- A time series is a time-oriented or chronological sequence of observations on a variable of interest
- When successive observations are dependent, future values may be predicted from past observations.

Simulation example

Generate two random processes using R

- set.seed(1234); t<-1:10; E<-rnorm(10)
- a<-t+E; b<-1+t+E (element-wise operation in R)
- plot(a,type="b");lines(b,type="b",col="red")
- t.test(a,b)

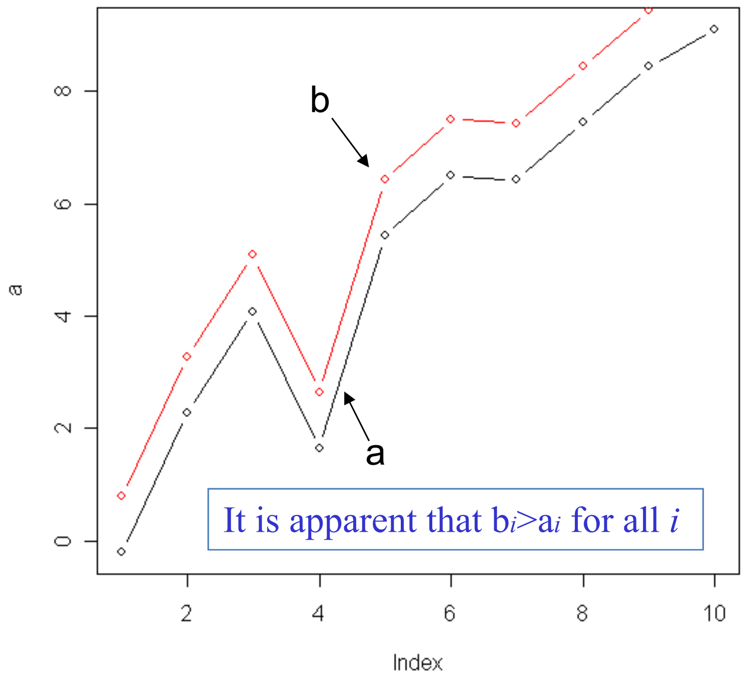
> a

- -0.2070657 2.2774292 4.0844412 1.6543023 5.4291247 6.5060559 6.4252600 7.4533681 8.4355480 9.1099622

>b

- 0.7929343 3.2774292 5.0844412 2.6543023 6.4291247 7.5060559 7.4252600 8.4533681 9.4355480 10.1099622

Time series plot



Welch two sample t-test

- **Standard R function**

data: a and b

- t = -0.725, df = 18, $\boxed{\text{p-value = 0.4778}}$

H_a: true difference in means is not equal to 0

- 95 percent confidence interval:

-3.897933 1.897933

- sample estimates:

mean of x 5.116843

mean of y 6.116843

Reject H_0 if p-value is small.

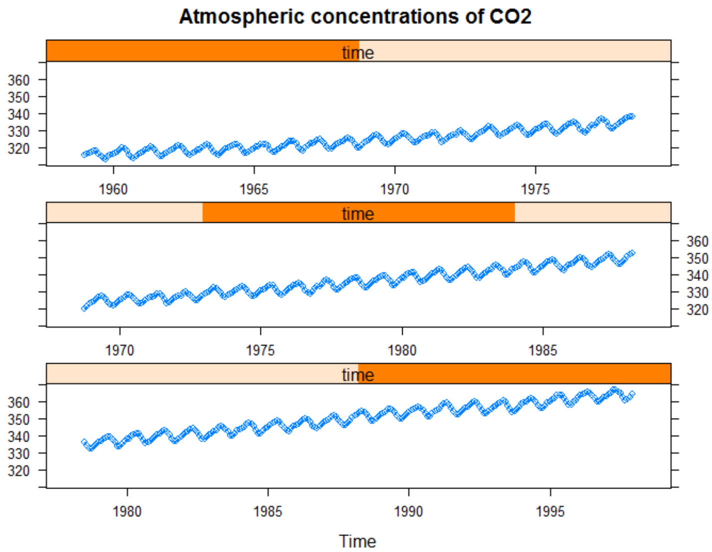
We are not able to reject H_0 that the difference between two series is zero

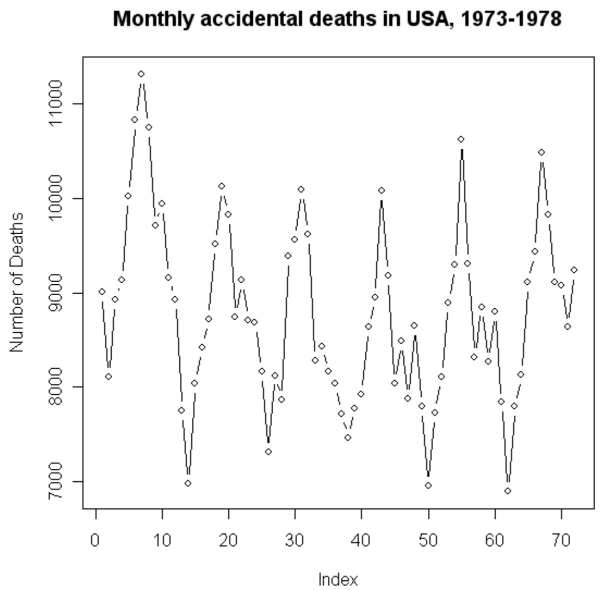
Definition of time series

A time series is usually defined as a collection of random variables indexed according to the order they obtained in time.

- For example, consider a time series as a sequence of random variables, $x_1, x_2, x_3, \cdots$, where the random variable x_1 denotes the value taken by the series at the first time point, the variable x_2 denotes the value for the second time period, and so on.
- In general, a collection of random variables $\{x_t\}$ indexed by t is referred to as a stochastic process, where t in this class is discrete and varies over the integers $t = 0, \pm 1, \pm 2, \ldots$ or some subset of the integers.

Examples of time series data





Examples of time series models (I)

2 building blocks

IID noises with finite second moment

- If $\{X_t\}$ is a sequence of independent and identical random variables with mean zero and second moment equal to $\sigma^2 < \infty$, we write $X_t \sim IID(0, \sigma^2)$.

White noises

- If $\{X_t\}$ is a sequence of random variables with $E(X_t) = 0$, $\sigma^2 < \infty$, and

Covariance $\gamma_X(r, s) = \begin{cases} \sigma^2, r = s \\ 0, \text{otherwise} \end{cases}$.

- Then $\{X_t\}$ is called *white noises* and written as $X_t \sim WN(0, \sigma^2)$.

Examples of time series models (II)

3 basic models

Random walk

- If $X_t = X_{t-1} + a_t$, $a_t \sim NID(0, \sigma^2)$, then $\{X_t\}$ is called a random walk.

Indep identical normal distribution.

First-order autoregressive process

- $X_t - \phi X_{t-1} = a_t$, $a_t \sim WN(0, \sigma^2)$ and a_t is uncorrelated with X_s $\forall\ s < t$.
  - X_t is partially related with X_{t-1}
  - When $\phi = 1$, we get a special case: random walk

Moving averaging

- $X_t = \frac{1}{3}(a_t + a_{t-1} + a_{t-2}), a_t \sim WN(0, \sigma^2)$.

Approaches to time series analysis

Time domain analysis

we are going to focus on this approach in this class.

- Focuses on modeling some future values of a time series as a parametric function of the current and past values
- very sensitive to extreme values.

Frequency domain analysis

- Assumes that the primary characteristics of interest in time series analyses relate to periodic or systematic sinusoidal variations found naturally in most data

In many cases, the two approaches may produce similar answers for long series, but the comparative performance over short samples is better done in the time domain.

Decomposition of time series

Statisticians usually decompose a time series into components representing

- trend
- seasonal variation
- other cyclic changes
- irregular fluctuations.

This decomposition is sometimes referred to as the classical decomposition model in time series analysis.

Seasonal variation

- Time series exhibit variation that is annual in period (or every 12 units of time).
- For example, the sales of electronic companies in the second quarter are typically the lowest.

Cyclic variation

- Time series exhibit variation at a fixed period due to some other physical cause.
- Examples are daily variation in temperature and business cycles.

Trend

- This may be loosely defined as ‘long-term change in the mean level’.

Classical decomposition model

trend component
seasonal component
random noise component

Classical decomposition model: $X_t = m_t + s_t + Y_t$, where m_t is a slowly changing function known as a "trend component", s_t is a function with known period d referred to as a "seasonal component", and Y_t is a random noise (irregular) component, where $EY_t = 0$, $s_{t+d} = s_t$ and $\sum_{j=1}^{d} s_j = 0$.

expect noise to be zero
s_t: seasonal
sum of a "season" is zero.

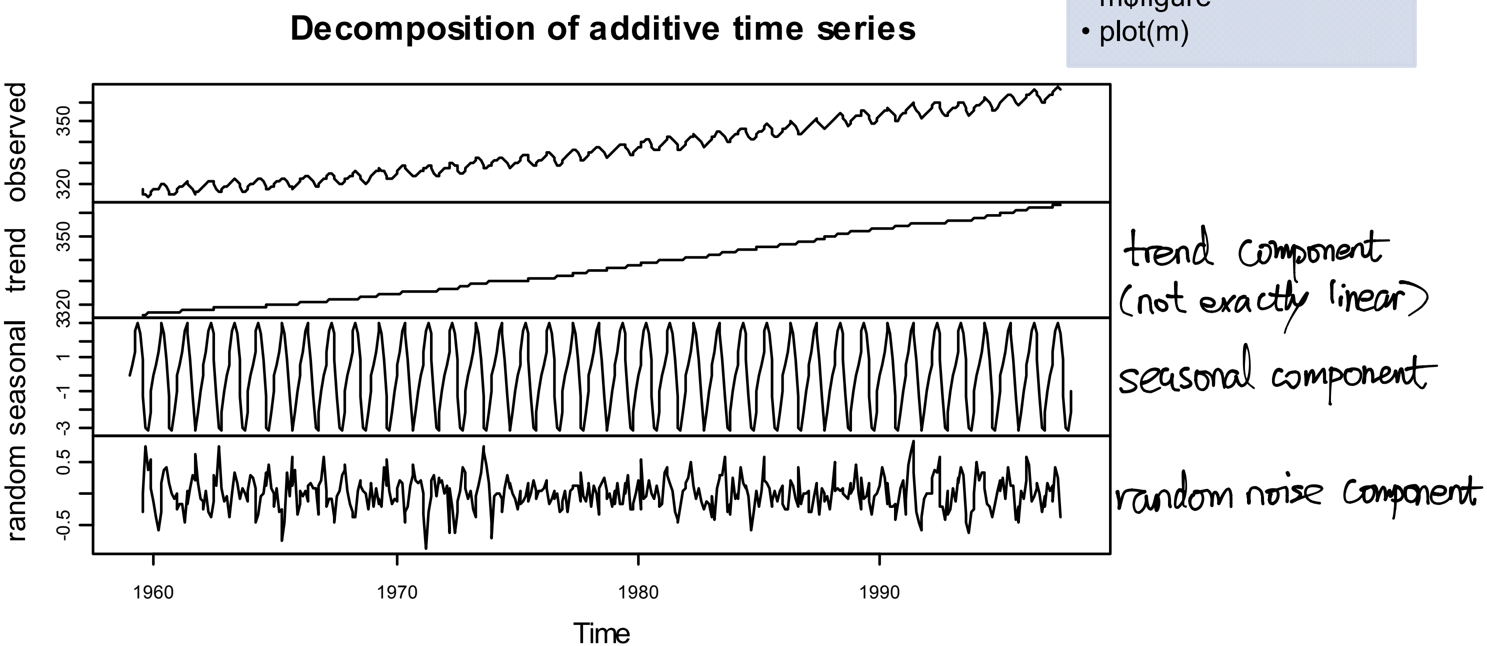
- Its aim is to estimate and extract the deterministic components m_t and s_t in the hope that $\{Y_t\}$ will turn out to be (weakly) stationary.

Classical decomposition in R

Atmospheric concentrations of CO_2 are expressed in parts per million (ppm) and reported in the preliminary 1997 SIO manometric mole fraction scale.

R code:

- library(TSA)
- m<-decompose(co2)
- m$figure
- plot(m)



Irregular components as stationary time series

Two types of stationary time series:

informal definition

- The strictly stationary time series is defined by the distributional property of a stochastic process (time series),
- The weakly stationary time series is defined via the first two moments of a process.

strictly implies weakly.

under normal distribution, strictly = weakly.

Strictly stationary time series

A strictly stationary time series is one for which the probabilistic behavior of every collection of values $(X_{t+1}, \dots, X_{t+n})$ is identical to that of the time shifted set $(X_{t+h+1}, \dots, X_{t+h+n})$. That is,

$$P\{X_{t+1} \leq c_1, \dots, X_{t+n} \leq c_n\} = P\{X_{t+h+1} \leq c_1, \dots, X_{t+h+n} \leq c_n\}$$

for all $t, n \in Z$ and $n > 0$, $c_1, \dots, c_n \in R$, and all time shifts $h = 0, \pm 1, \pm 2, \dots$.

- A time series is said to be strictly stationary if its properties are not affected by a change in the time origin.

Definitions

Consider two stochastic processes (or time series) $X_t = (x_t, x_{t+1}, x_{t+2}, \dots)^T$ and $Y_t = (y_t, y_{t+1}, y_{t+2}, \dots)^T$

- $\mu_t^x = E(\cancel{y}_t)$ and $\mu_t^y = E(y_t)$ for all $t \in Z$
- $\gamma(t,s) = E[(x_t - \mu_t)(x_t - \mu_s)]$ for all $t, s \in Z$
- $\rho(t,s) = \gamma(t,s)/\sqrt{\gamma(t,t)\gamma(s,s)}$ (autocorrelation)
- $\gamma_{xy}(t,s) = E[(x_t - \mu_t^x)\left(y_s - \mu_s^y\right)]$
- $\rho_{xy}(t,s) = \gamma_{xy}(t,s)/\sqrt{\gamma_x(t,t) \cdot \gamma_y(s,s)}$ (cross correlation)
- For stationary time series, $\gamma(t,s)$ may be denoted as $\gamma(t, t+h)$, where $s = t + h, \forall t, s, h \in Z$

Remarks

- *Autocorrelation* refers to the correlation of a time series with its own past and future values. Autocorrelation is also sometimes called "*lagged correlation*" or "*serial correlation*", which refers to the correlation between members of a series of numbers arranged in time.
- Positive autocorrelation might be considered a specific form of "*persistence*", a tendency for a system to remain in the same state from one observation to the next.
- Consequences of the presence of autocorrelation

Weakly stationary time series

$E(X_t) = \mu, var(X_t) = \gamma(0) = \sigma^2$ are constant and does not depend on time t for all $t \in Z$.

The autocovariance function, $\gamma(t, t+h)$, is a function of h and is independent of time for all $t, h \in Z$.

- A weakly stationary time series is also known as covariance-stationary or stationary in the wide sense
- For stationary time series, $\gamma(t, t+h)$ and $\rho(t, t+h)$ can be denoted as $\gamma(h)$ and $\rho(h)$, respectively
- $\gamma(0) = var(X_t)$

strictly: $f(X_1, X_2) = f(X_2, X_3) = \cdots = f(X_t, X_{t+1}) = \cdots$

weakly: $cov(X_1, X_2) = cov(X_2, X_3) = \cdots = cov(X_t, X_{t+1}) = \cdots$

Examples of Autocovariance and stationary time series

Random walk

- For simplicity and without loss of generality, let $a_t = 0$ for all $t < 0$.
- $X_t = X_{t-1} + a_t = \underbrace{X_{t-2} + a_{t-1}}_{X_{t-1}} + a_t = \cdots = \sum_{k=0}^{t} a_{t-k}$
- $var(X_t) = var(\sum_{k=0}^{t} a_{t-k}) = (t+1) \cdot \sigma^2$. That is, $\gamma(0)$ is a function of time and X_t is not weakly stationary.

First order autoregressive processes

- $X_t = \phi X_{t-1} + a_t = \phi \underbrace{(\phi X_{t-2} + a_{t-1})}_{X_{t-1}} + a_t = \cdots = \sum_{k=0}^{\infty} \phi^k a_{t-k}$
- $var(X_t) = \sum_0^{\infty} \phi^{2k} var(a_t) = \sigma^2/(1-\phi^2)$
- $var(X_t)$ converges to a constant if $\phi \neq 1$. That is, X_t is not weakly stationary if $\phi = 1$.

Sample autocorrelation functions (SACF)

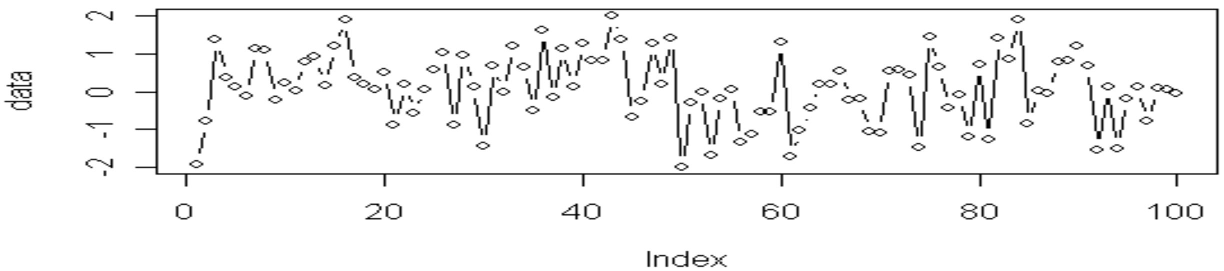
Plotting the sample autocorrelation functions (SACF) $\hat{\rho}(h)$ against the lag h for $h = 1,2,\ldots,M$ are useful tool in interpreting time series, where M is usually much less than the series length.

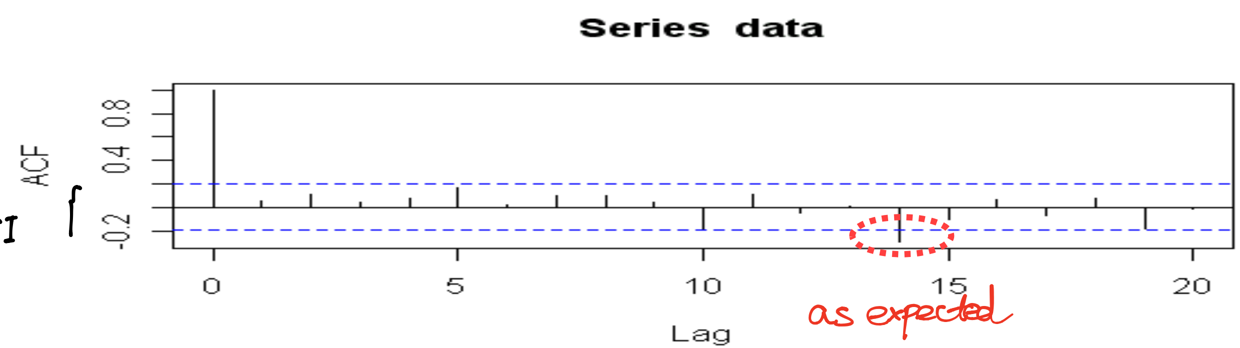
- $\hat{\gamma}(h) = n^{-1}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x})$ with $\hat{\gamma}(-h) = \hat{\gamma}(h)$ for $h = 0,1,\ldots,n-1$
- $\hat{\rho}(h) = \hat{\gamma}(h)/\hat{\gamma}(0)$

For a random time series, $\hat{\rho}_k$ is approximately $N\left(0, \frac{1}{N}\right)$ *for* $k \neq 0$, where N is the length of the series. Thus, if a time series is random, we can expect 19 out of 20 of the values of $\hat{\rho}_k$ to lie between $\pm 2/\sqrt{N}$.

R code:

- data<-rnorm(100)
- par(mfrow=c(2,1))
- plot(data,type="b")
- acf(data)





95% CI

as expected

type I error

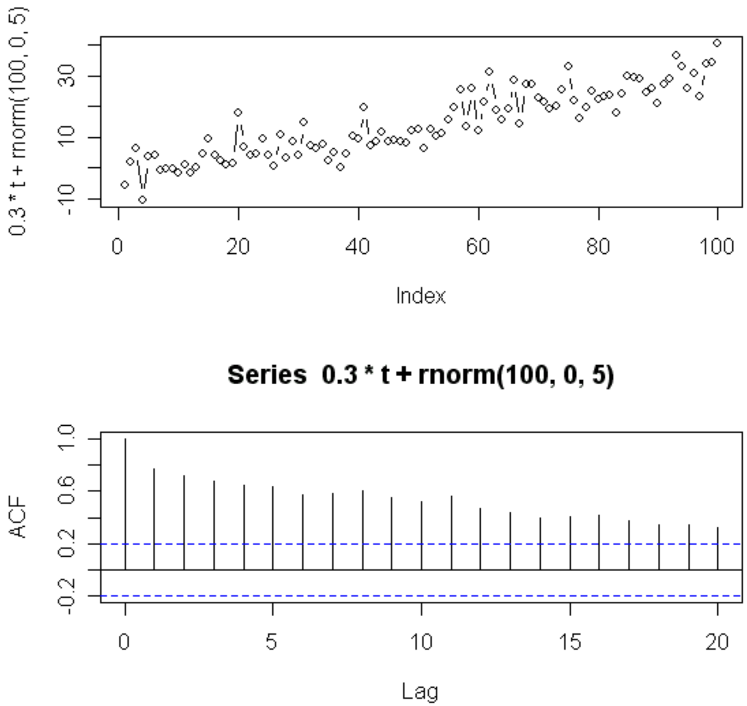
SACF of time series with linear time trends

If a time series contains a trend, then *the values of $\hat{\rho}_k$ will not come down to zero except for very large values of the lag.*

- This is because an observation on one side of the overall mean tends to be followed by a large number of further observations on the same side of the mean because of the trend
- See the theoretical justification in the practice questions

R code:

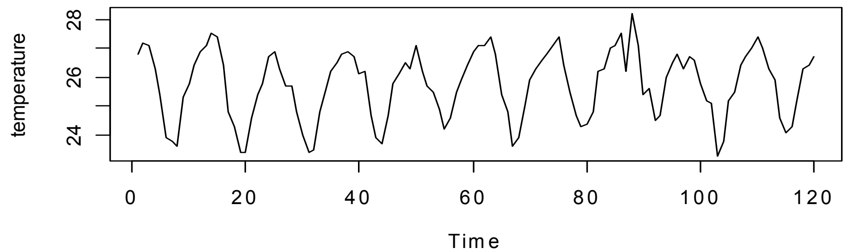
```
par(mfrow=c(2,1))
set.seed(1234)
plot(0.3*t+rnorm(100,0,5),type="b")
acf(0.3*t+rnorm(100,0,5))
```

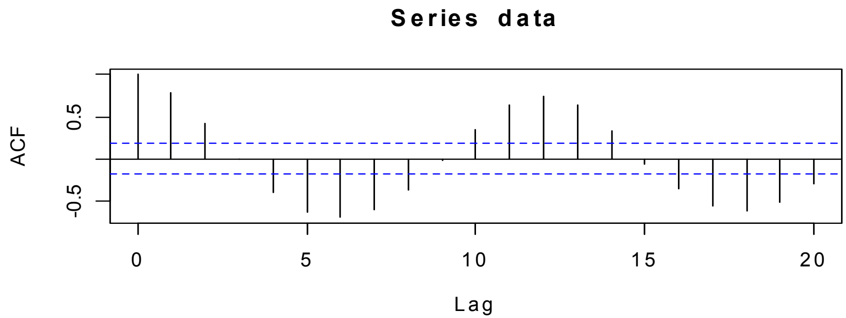


Interpreting SACF of seasonal variation

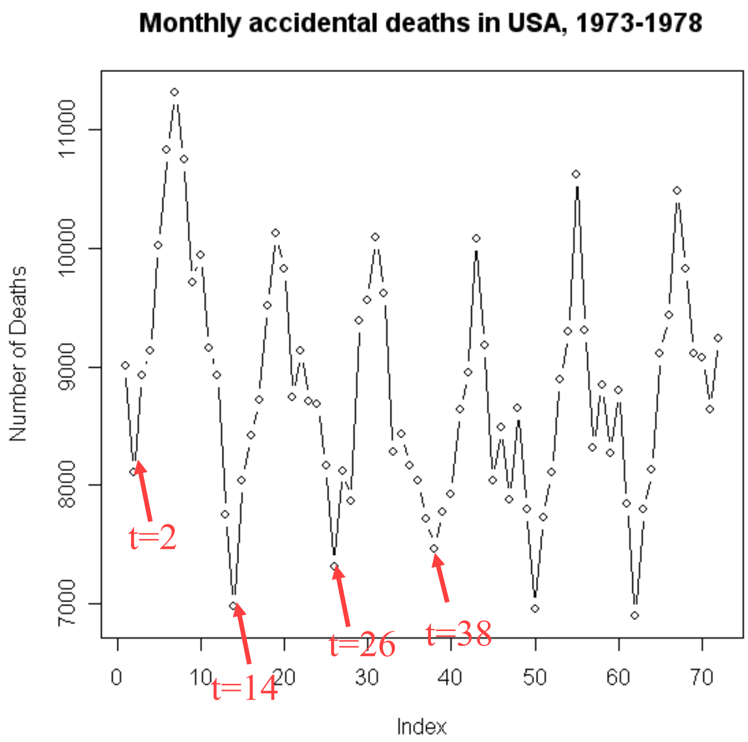
SACF

If a time series contains seasonal variation, then the sample autocorrelation function will also exhibit oscillation at the same frequency.





Seasonal variation



"The length of time" between consecutive peaks or troughs is fixed

Steps to time series modeling

Plot the time series and check for

- Trend, seasonal and other cyclic components, any apparent sharp changes in behavior, as well as any outlying observations

Remove trend and seasonal components to get residuals

Choose a model to fit the residuals

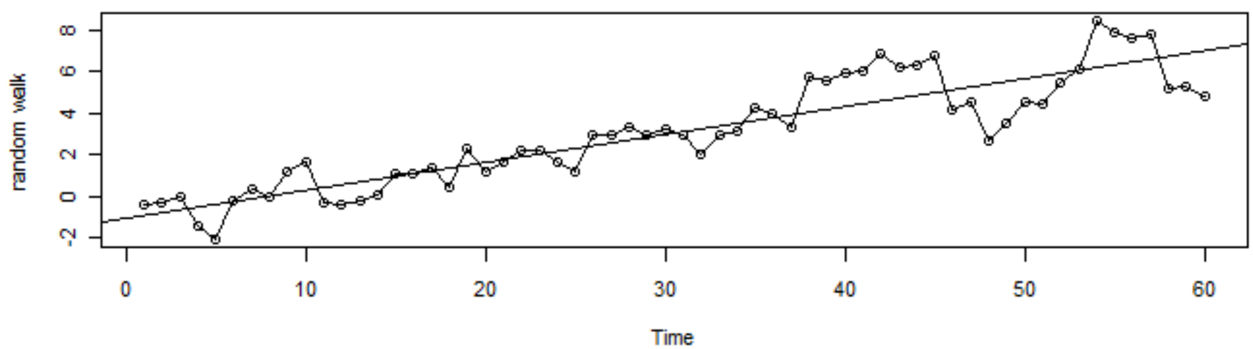
Forecasting can be carried out by forecasting residual and then inverting the transformation carried out in Step 2.

Estimation of linear time trend

- Model: $Y_t = \mu_t + X_t$, $\mu_t = \beta_0 + \beta_1 t$, $t = 1, \ldots, n$
- Least squares estimation: $Q(\beta_0, \beta_1) = \sum_{t=1}^{n} [Y_t - (\beta_0 + \beta_1 t)]^2$
- Estimator: $\hat{\beta}_1 = \sum_{t=1}^{n} (Y_t - \bar{Y})(t - \bar{t}) / \sum_1^n (t - \bar{t})$, and $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{t}$, where $\bar{t} = \frac{n+1}{2}$.

R code:

- library(TSA)
- data(rwalk)
- mod_timetr<-lm(rwalk~time(rwalk))
- summary(mod_timetr)
- win.graph(height=2.5, pointsize=8)
- plot(rwalk, type='o', ylab="random walk")
- abline(mod_timetr) # add the fitted regression line



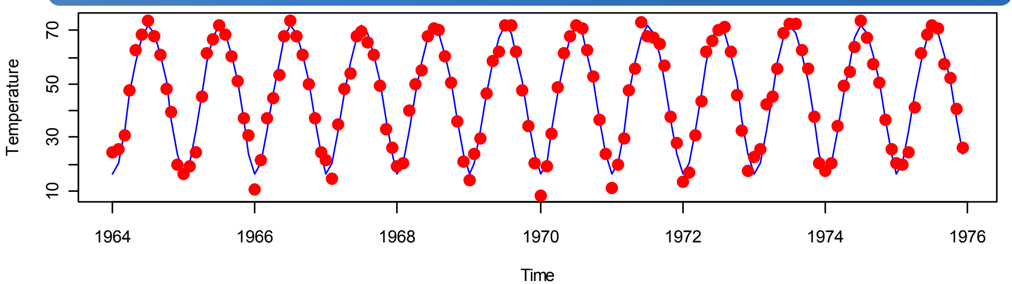
Estimation of cyclical or seasonal time trends (I)

- Monthly mean model: $Y_t = \mu_t + X_t,\ E(X_t) = 0, \forall t$, where μ_t is monthly data with 12 constants (parameters) which gives the expected value for each of the 12 months.
- We may write $\mu_t = \begin{cases} \beta_1, & t = 1,13,25,\ldots \\ \beta_2, & t = 2,14,26,\ldots \\ & \vdots \\ \beta_{12}, & t = 12,24,36,\ldots \end{cases}$.

R code:

```
• data(tempdub)
• month.<-season(tempdub)
• mod_cyctr<-lm(tempdub~month.); temp<-fitted(mod_cyctr)
• win.graph(height=2.5, pointsize=8)
• plot(ts(temp,freq=12,start=c(1964,1)), ylab='Temperature', type="l" ,col=4,
  ylim=range(c(temp,tempdub)))
• points(tempdub,col=2, lwd=4)
```

Monthly average temperature (in degrees Fahrenheit) recorded in Dubuque 1/1964 - 12/1975.

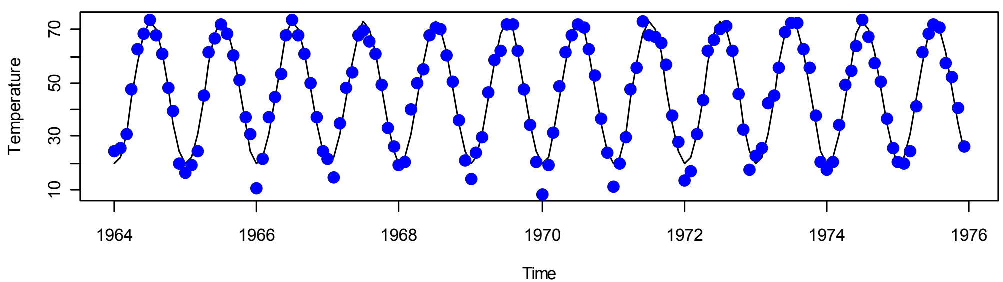


Estimation of cyclical or seasonal time trends (II)

- Model: $\mu_t = \beta_1 \cos(2\pi f t) + \beta_2 \sin(2\pi f t)$, where $1/f$ is called the period.
- For example, monthly data with time index as $1, 2, \ldots$, has $f = 1/12$ because such sinusoidal function will repeat itself every 12 months. In this case, the period is 12.
- Least square estimation: use $\cos(2\pi f t)$ and $\sin(2\pi f t)$ as predictor variables.

R code:

- har.<-harmonic(tempdub,1)
- mod_costr<-lm(tempdub~har.); temp_<-fitted(mod_costr)
- win.graph(height=2.5, pointsize=8)
- plot(ts(temp_,freq=12,start=c(1964,1)),
- ylab='Temperature', type="l" ,ylim=range(c(temp_,tempdub)))
- points(tempdub,col=2, lwd=4)



Moving average methods for trend estimation

Two-sided moving average:

- Let q be nonnegative integer and assume that m_t is approximately linear over the interval $[t-q, t+q]$ and the average of the error terms over this interval is close to zero
- $\hat{m}_t = \sum_{-q}^{q} X_{t+j}/(2q+1), q+1 \le t \le n-q$

One-sided moving average:

- $\hat{m}_t = \sum_{0}^{n-t} \alpha(1-\alpha)^j X_{t+j}, t = 1, \ldots, q$
- $\hat{m}_t = \sum_{0}^{t-1} \alpha(1-\alpha)^j X_{t+j}, t = n-q+1, \ldots, n.$
- Empirically, it has found that values of α between 0.1 and 0.3 give reasonable estimates.

R code

```
phi<-0.6; sigma2<-1; gamma0<-sigma2/(1-phi*phi)
set.seed(1234)
x0<-rnorm(1,0,gamma0); z<-rnorm(59); x<-numeric(60); x[1]<-x0
for(i in 2:60) x[i]<-phi*x[i-1]+z[i-1]
ma<-filter(x,rep(1/13,13))
ts.plot(x,type="b", main="two-sided moving-average", lwd=1)
lines(ma,col=2,lwd=2)
```