Handout 1

Introduction to Time Series

Class notes for Statistics 451: Applied Time Series
Iowa State University

January 7, 2007
17h 8min

Time Series Data

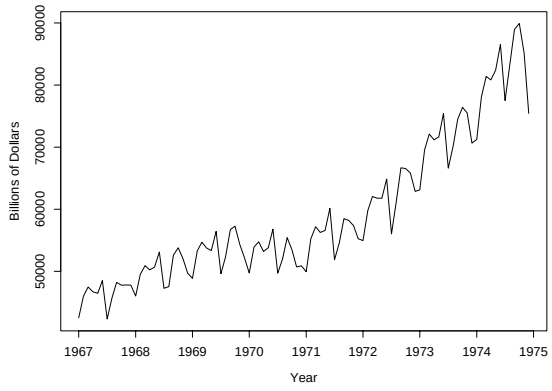
A sequence of observations taken over time (usually equally spaced) $Y_1, Y_2, \ldots, Y_t, \ldots Y_n$ where n is the number of observations in the "realization"

- Univariate – single series, e.g., daily closing price of IBM common stock.
- Multivariate – two or more series (vector time series) (e.g. daily closing price of common stock of IBM, Xerox, Kodak gives 3 values per day).
- Time interval: yearly, monthly, quarterly, weekly, daily, hourly, every minute, ... every k nanoseconds ....
- Population of units → Sample
- Process operating in time → Realization

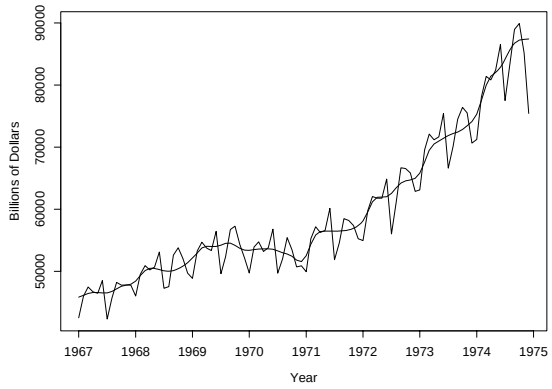
Continuous Time and Discrete Time

- Time is continuous, but data are usually reported at discrete points in time.
- Thus "sampling" a continuous time series leads to a discrete time series.
- Sampling is usually equally spaced in time (sampling sometimes known as "reporting")
- Time series data are usually not independent, especially if sampling interval is small. Observations close together are often more alike than those far apart (e.g. daily temperatures).

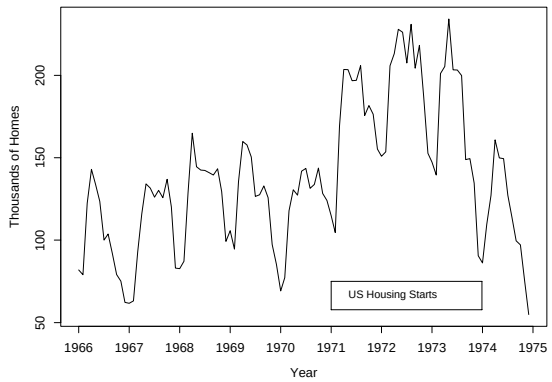
Manufacturer's Shipments 1967-1975



Smoothed Manufacturer's Shipments 1967-1975



Seasonal Data Example
US Housing Starts 1966-1974



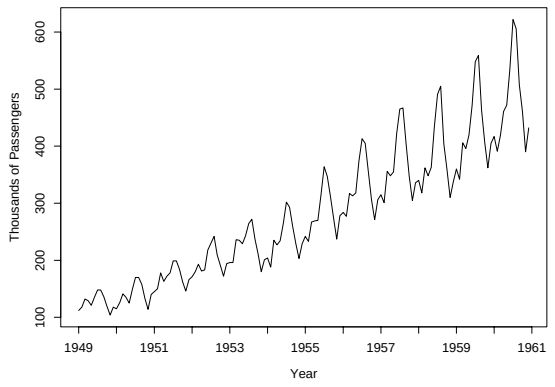
Types of Responses

- Continuous (e.g. temperature, concentration).
- Discrete (e.g., number of people, number manufactured) (often from aggregation). We often approximate discrete responses with a continuous model.
- Binary (e.g., success or failure).

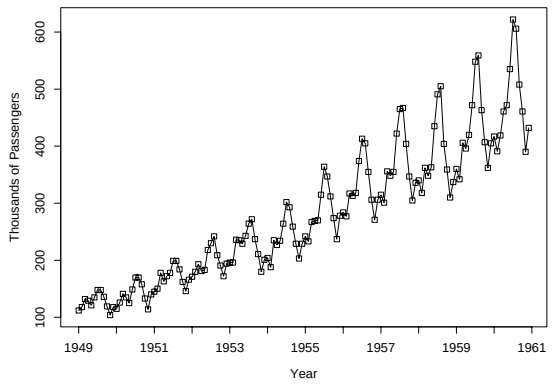
Time Series Applications

- Economics (e.g., GNP, NNP, Unemployment Rate, Interest Rates, Money supply)
- Business (e.g., Inventory, Cash, Sales, Prices, Quality Indices, Stock Price)
- Sociology (e.g., Crime Rates, Divorce Rates)
- Meteorology (e.g., Rainfall, Temperature, Wind Speed)
- Astronomy (e.g., Solar Activity, Sun Spots, Star Brightness)
- Ecology (e.g., Air Pollution, Water Pollution, Wildlife Population)
- Engineering

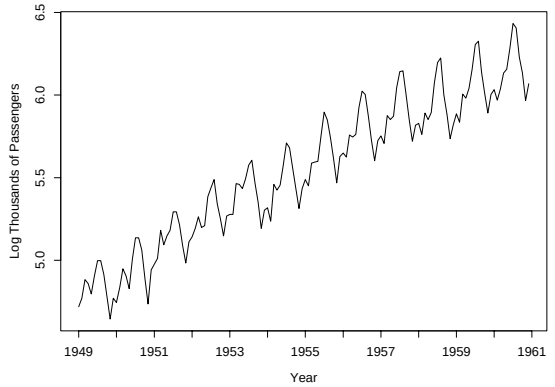
International Airline Passengers 1949-1960



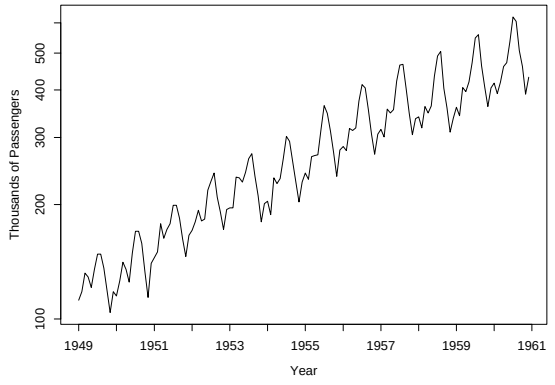
International Airline Passengers 1949-1960



Log International Airline Passengers 1949-1960



International Airline Passengers 1949-1960 on Log Axis



Reasons for Analyzing Time Series

- Description
  Graphical, Numerical, Features
- Explanation
  Explanatory Variables
- Prediction (Forecasting)
  Next Year's Sales
- Control
  Quality of Manufacturing Process Economy

Filters

- Filters are like functions, but for time series. Let $\mathbf{x} = x_1, x_2, \ldots$ and $\mathbf{y} = y_1, y_2, \ldots$. Then $\mathbf{y} = f(\mathbf{x})$:

- Linear filter:

$$y_t = \sum_{r=-q}^{+s} a_r x_{t+r} = a_{-q}x_{t-q} + a_{-q+1}x_{t-q+1} + \ldots a_0 x_t + \ldots a_r x_{t+r}.$$

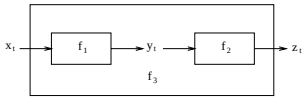
For example with $q = 2$, $s = 2$, and $a_r = 1/5$,

$$y_t = \sum_{r=-2}^{2} a_r x_{t+r} = (x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2})/5$$

which is a "moving average" filter.

Filters in Series

Two (or more) filters in series form an overall filter:

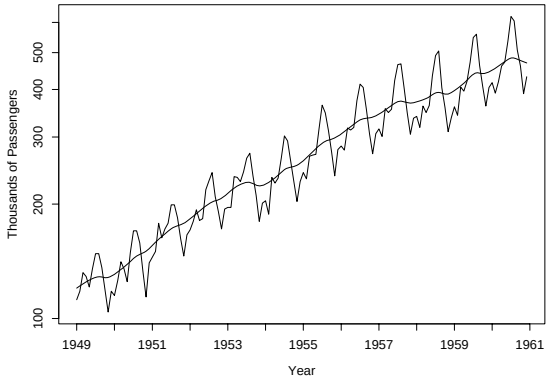


	f_1	$f_3 = f_2 \odot f_1$
x_1		
	$(x_1 + x_2)/2$	
x_2		$(x_1 + 2x_2 + x_3)/4$
	$(x_2 + x_3)/2$	
x_3		$(x_2 + 2x_3 + x_4)/4$
	$(x_3 + x_4)/2$	
x_4		
$\vdots$	$\vdots$	$\vdots$

$$z_t = \sum_{r=-1}^{1} a_r x_{t+r}$$

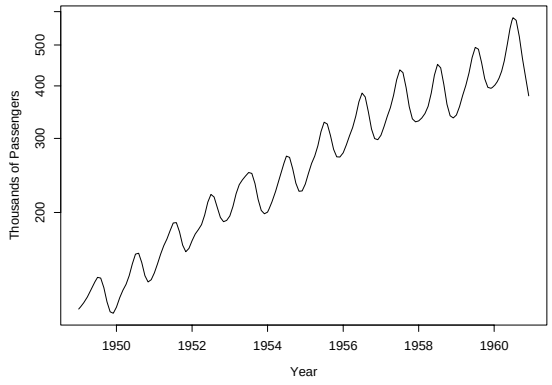
where $a_{-1} = 1/4$, $a_0 = 1/2$, and $a_1 = 1/4$. This particular weighted moving average is also known as a "Hanning filter."

Smoothed International Airline Passengers 1949-1960



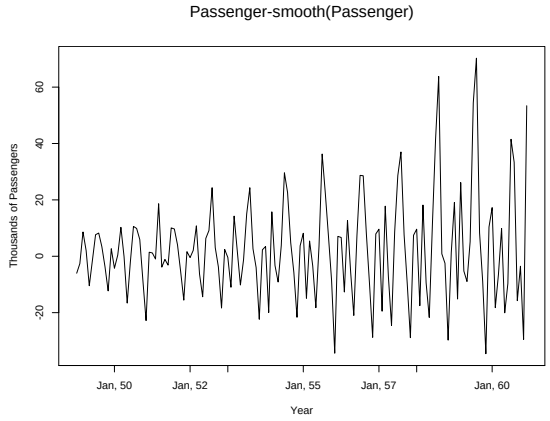
Smoothed International Airline Passengers 1949-1960

`tsplot(smooth(passengers.ts))`



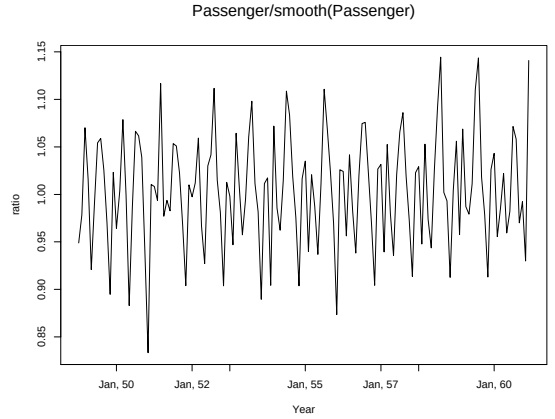
Difference Noise
International Airline Passengers 1949-1960

`tsplot(passengers.ts-smooth(passengers.ts))`



Ratio Noise
International Airline Passengers 1949-1960

`tsplot(passengers.ts/smooth(passengers.ts))`



Applications of "Classical" Decomposition Model

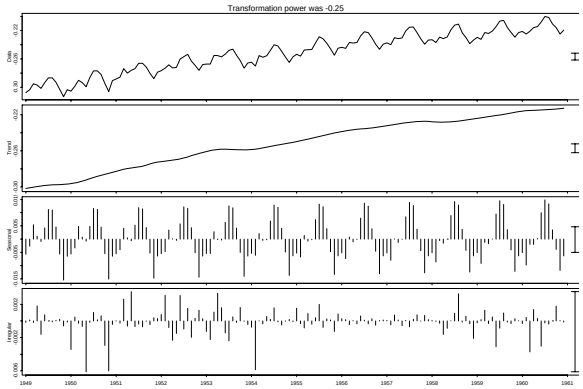
Model:

$$Z_t = T_t \times S_t \times C_t \times I_t$$

- Decompose Z_t into T_t, S_t, C_t and $I_t, t = 1, \ldots, n$.
- Forecast components into future and combine to forecast Z_t.
  - Forecast $T_{101}, S_{101}, C_{101}, I_{101}$.
  - $\hat{Z}_{101} = \hat{T}_{101}\hat{S}_{101}\hat{C}_{101}\hat{I}_{101}$
- Compute a smoothed series: $\text{Smooth}_t = T_t S_t C_t$ (no I_t)
- Compute seasonally adjusted series (also known as deseasonalized): $D_t = T_t C_t I_t$ (no S_t)

Sabl Plot for the
International Airline Passengers 1949-1960

`sablplot(sabl(passengers.ts))`

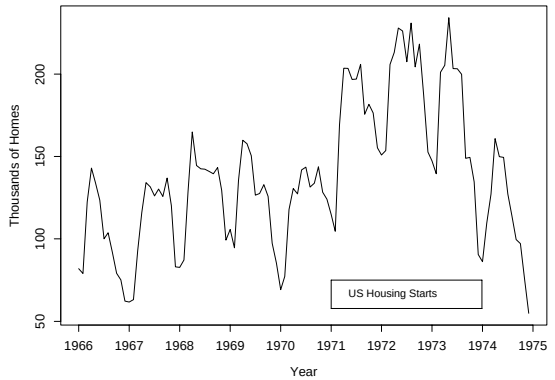


For more information on decomposition methods, see

- Some elementary business statistics textbooks
- SAS ETS PROC X-11
- Splus function `sabl()`

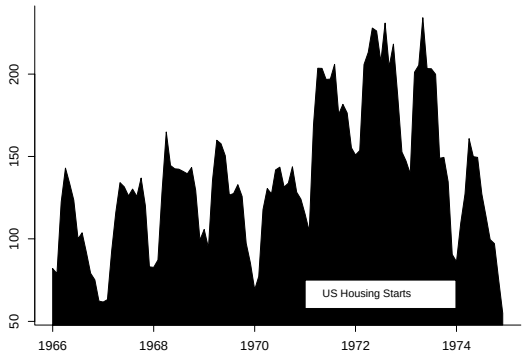
US Housing Starts 1966-1974

`tsplot(hstart)`



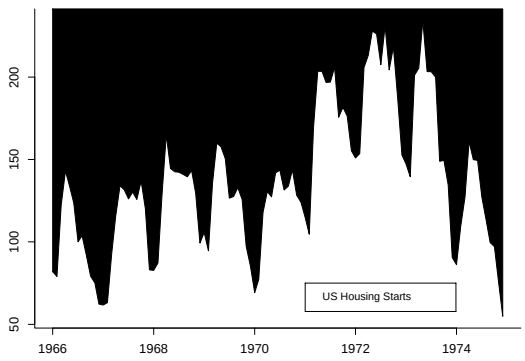
US Housing Starts 1966-1974

`shaded.tsplot(hstart,top=F)`

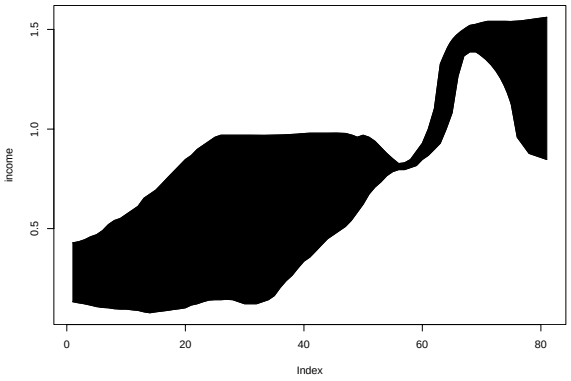


US Housing Starts 1966-1974

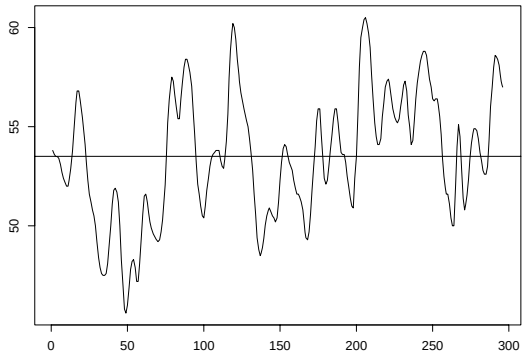
`shaded.tsplot(hstart,top=T)`



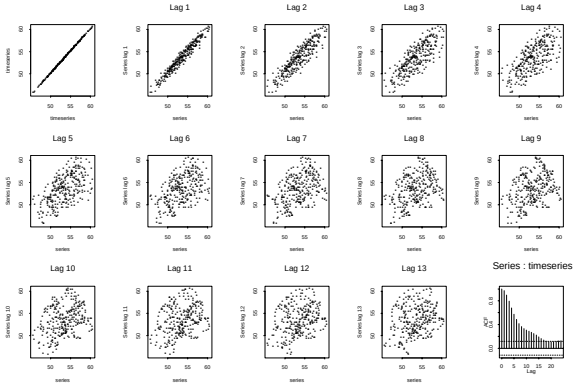
Balance of Trade in England 1700-1780 (Playfair)



Percent CO_2 Outlet Gas (sampling interval 9 seconds)

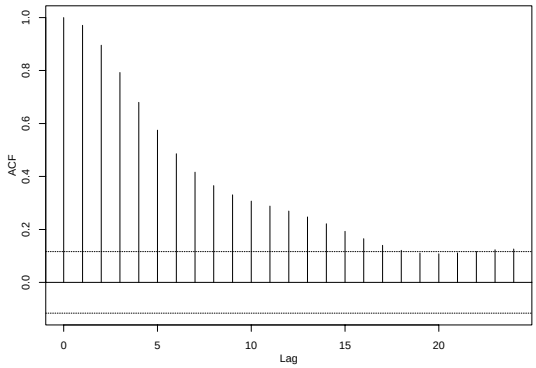


Visualization of the Autocorrelation Function for the Percent CO_2 Outlet Gas Data



Autocorrelation Function for the Percent CO_2 Outlet Gas Data

Series : gasry.d$ts



SplusTS `iden` Command Output for the Percent CO_2 Outlet Gas Data