TIME SERIES ANALYSIS

Chapter 1: Introduction

A time series is a sequence of observations taken sequentially in time. Many sets of data appear as time series: a monthly sequence of the quantity of goods shipped from a factory, a weekly series of the number of road accidents, hourly observations made on the yield of a chemical process, and so on. An intrinsic feature of a time series is that, typically, adjacent observations are dependent. Time series analysis is concerned with techniques for the analysis of this dependence.

1 Time Series

DEFINITION 1 A **time series** is a sequence of observations over time.

EXAMPLE 1 Records of a person's height:

age:	1	2	3	4	5	6	7
height(m):	0.4	0.5	0.8	1.0	1.1	1.2	1.4
(notation:	y_1	y_2	y_3	y_4	y_5	y_6	y_7)

For this example, we have $n = 7$ observations. We call n the **number of observations** or the **length of a times series**. We denote the observation at time t by y_t (or x_t etc.)

The time series can be then denoted as

$$\{0.4, 0.5, 0.8, 1.0, 1.1, 1.2, 1.4\}$$

or

$$\{y_t : t = 1, 2, ..., n\}$$

Note that the above observations are taken over equally time intervals. We can also observe the variable with unequally time intervals

age:	0.5	1	1.5	2	3	5	7
height(m):	0.35	0.4	0.45	0.5	0.8	1.1	1.4

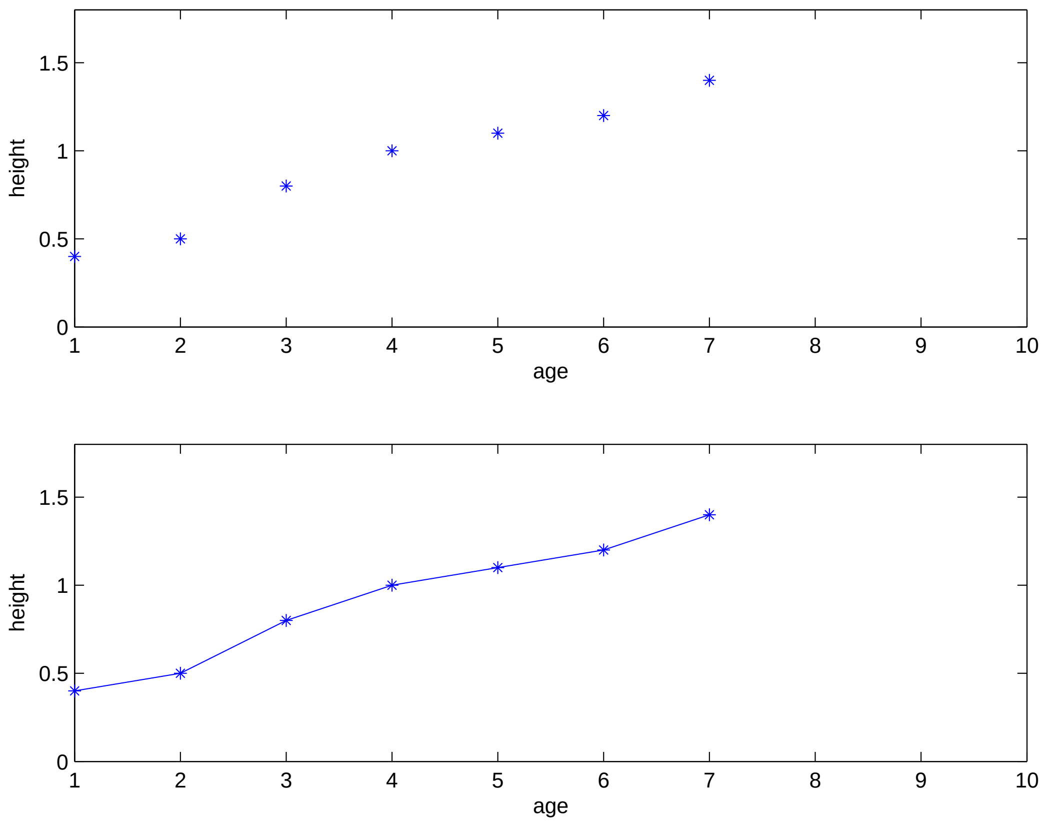


Figure 1: Plot of Example 1

DEFINITION 2 A time series is said to be discrete when the set T_0 of times at which observations are taken is a discrete set. A time series is said to be continuous when observations are recorded continuously over some time interval,e.g. when $T_0 = [0, 1]$.

Remarks: We are mainly interested in discrete-time time series with equally fixed time intervals. e.g. observations made monthly, daily, weekly, etc.

2 Some Representative Time Series

EXAMPLE 2 Unemployment Rate (%) in Singapore

More time series [What can you observe in the time series?]

year	rate	year	rate	year	rate
1973	4.4	1984	2.7	1995	2.7
1974	3.9	1985	4.1	1996	3.0
1975	4.5	1986	6.5	1997	2.4
1976	4.4	1987	4.7	1998	3.2
1977	3.9	1988	3.3	1999	4.6
1978	3.6	1989	2.2	2000	4.4
1979	3.3	1990	1.7	2001	3.4
1980	3.5	1991	1.9	2002	5.2
1981	2.9	1992	2.7	2003	5.4
1982	2.6	1993	2.7		
1983	3.2	1994	2.6		

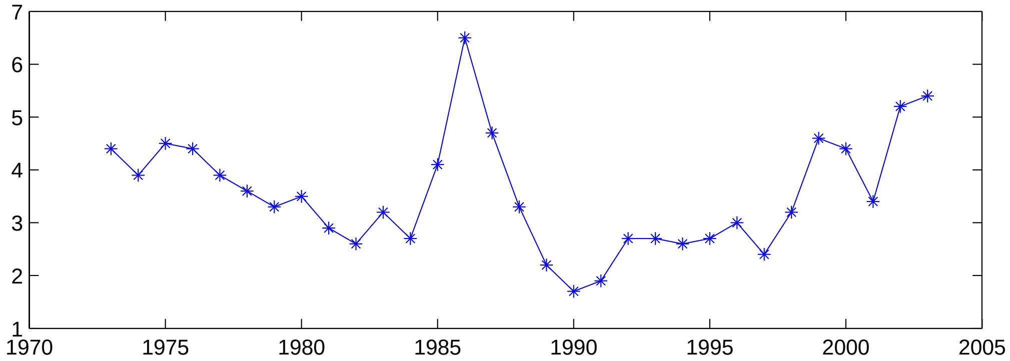


Figure 2: Plot of Example 2

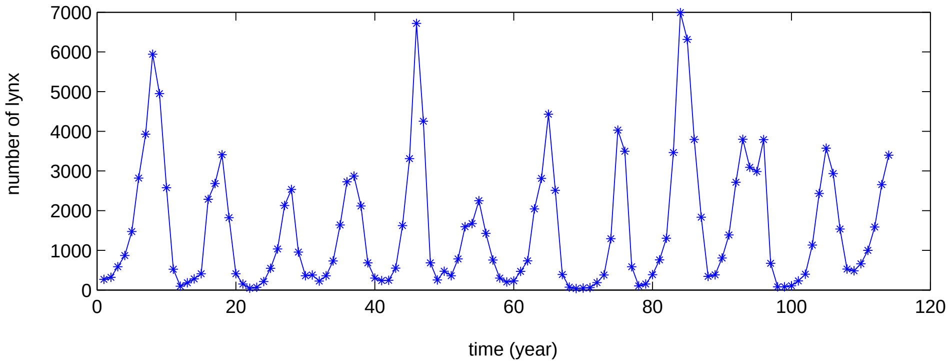


Figure 3: Canadian Lynx captured (1828-1934)

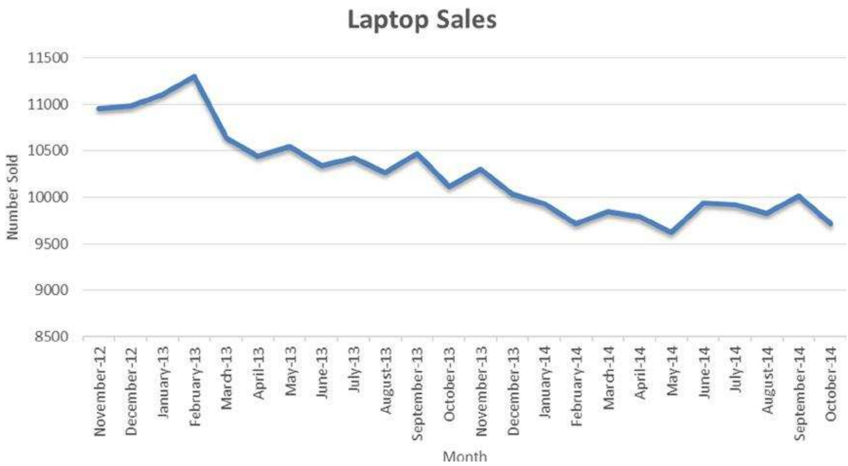


Figure 4: monthly laptop computer sales

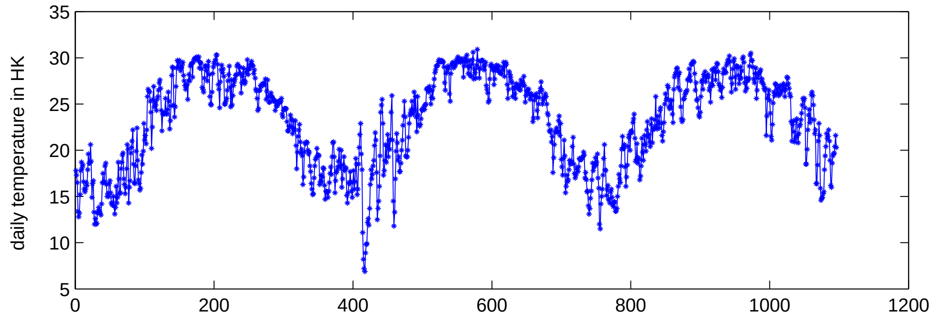


Figure 5: Temperature in Hong Kong (1994-1997)

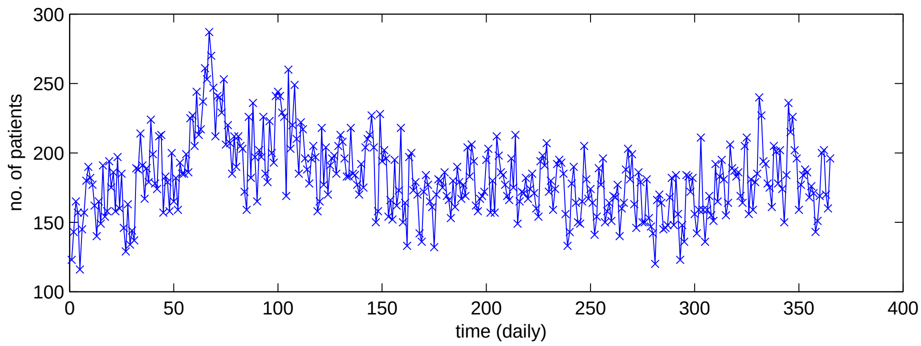


Figure 6: Number of patients with respiratory problems in Hong Kong (1994)

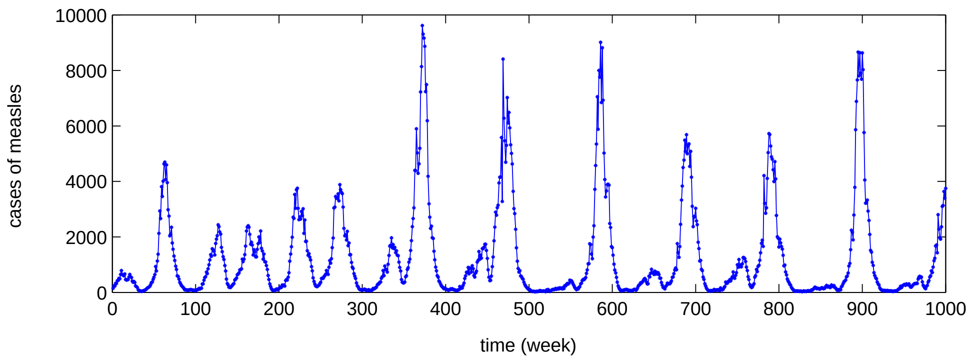


Figure 7: Measles cases in London (1944-1978)

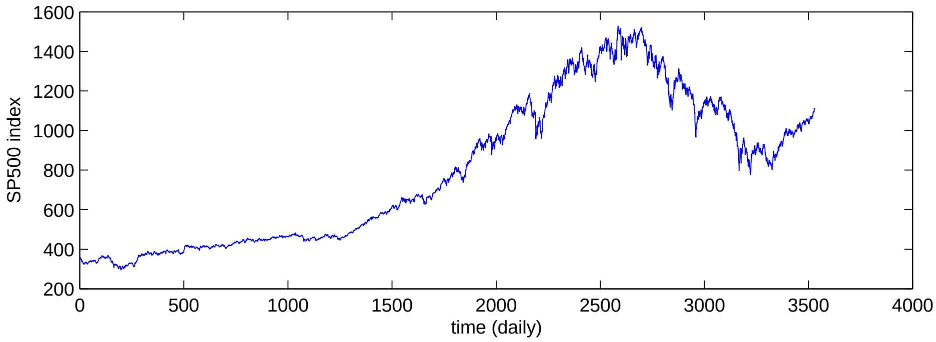


Figure 8: S & P500 stock index (1990-2004)

3 Purpose of Time Series Analysis

Time series data are often examined in hopes of discovering a historical pattern that can be exploited in the preparation of a forecast.

Specifically speaking, the first objective of time series analysis is to understand or model the stochastic mechanism that gives rise to the observed series. When presented with a time series, **the first step** in the analysis is usually to plot the observations against time to give what is called **a time plot**, and then to obtain simple descriptive measures of the main properties of the series. The power of the time plot is illustrated in Fig 4, which clearly shows that there is a regular seasonal effect.

After an appropriate family of models has been chosen and estimated, the next major objective of time series analysis is to **forecast** or **predict** future values of the series.

DEFINITION 3 Prediction of future events and conditions are called forecasts, and the act of making such predictions is called forecasting.

For example,

- what will be the unemployment rate next year?
- Is there a trend in global temperature?
- what is the seasonal effect?
- what is the relationship between GDP and interest rate?

Broadly speaking, forecasting methods can be divided into two basic types:

- Qualitative forecasting methods: use the opinions of experts to predict future events subjectively.

▷ Quantitative forecasting methods: Based the historical data, use statistical methods to predict future values of a variable.

The readers may be expecting the first section to deal with summary statistics. Indeed, in most areas of statistics, a typical analysis begins by computing the sample mean (or median or mode) and the standard deviation to measure "location" and "dispersion". However, time series analysis is different ! If a time series contains trend, seasonality or some other systematic component, the usual summary statistics can be seriously misleading and should not be calculated. Moreover, even when a series does not contain any systematic components, the summary statistics do not have their usual properties.

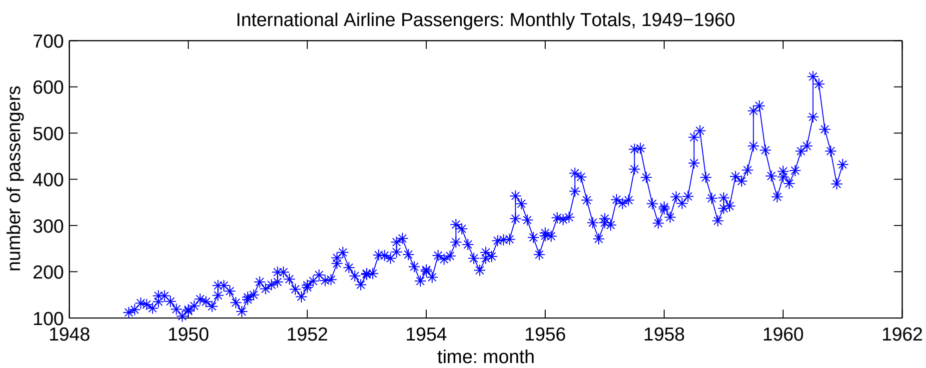
4 Components of a time series

In order to identify the pattern of time series data, it is often convenient to think of a time series as consisting of several components.

DEFINITION 4 The components of a time series are: Trend, Cycle, Seasonal variations and Irregular fluctuations.

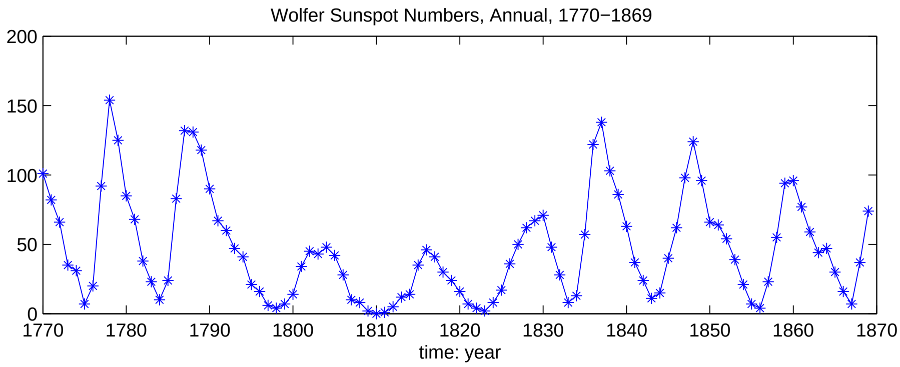
(i) **Trend** refers to the upward or downward movement that characterizes a time series over a period of time. Trend reflects the long-run growth or decline in the time series.

Reasons for trend: technology improvement; changes in consumer tastes; increases of per capita income; increase of total population; market growth; inflation or deflation.



(ii) **Cycle** refers to recurring up and down movements around trend levels. These fluctuations can have a duration of anywhere from two to ten years or even longer measured from peak to peak or trough to trough.

Typical examples for cycle: “business cycle (Karl Mark’s explanation)”; “nature Phenomena”; “interaction of two variables”. Explanations of cycle is very difficult.



(iii) **Seasonal variations or Seasonality** are periodic patterns in a time series that complete themselves within a calendar year and are then repeated on a yearly basis.

Typical examples are the average monthly temperature, the number of monthly housing starts and department store sales. Reasons for seasonalilty: weather; customs.

(iv) **Irregular fluctuations** represent what is "left over" in a time series after tread, cycle and seasonal variations have been accounted for. It is random movement in a time series. Some irregular fluctuations in a time series are caused by "unusual" events that can not be forecasted such as earthquakes, accidents, hurricanes, wars and the like.

These components may be combined in different ways. It is usually assumed that they are multiplied or added. This leads to possible structures for a time series

(i) Additive

$$Y_t = T_t + C_t + S_t + R_t$$

(ii) Multiplicative

$$Y_t = T_t \times C_t \times S_t \times R_t$$

where T_t is the trend component (or factor) in time period (or point) t; S_t is the seasonal component (or factor) in time period (or point) t; C_t is the cyclical component (or factor) in time period (or point) t; R_t is the irregular component (or factor) in time period (or point) t;

The second type can be changed into an additive model by taking the logarithms.

5 Stationary time series

A mathematical definition of stationary time series will be given later. However it may be helpful to introduce here the idea of stationary from an intuitive point of view.

A time series is said to be stationary if there is no systematic change in mean (no trend), if there is no systematic change in variance and if strictly periodic variations have been removed. Intuitively, the properties of one section of the data are much like those of any other section.

6 Transformation

The first, and most important, step in any time series analysis is to plot the observations against time. This graph, called a **time plot**, will show up important features of series such as trend, seasonality, outliers and discontinuities.

Plotting the data may suggest that it is sensible to consider transforming them, for example, by taking logarithms or square roots. Two main reasons for making a transformation are as follows.

- ▷ Stabilize the variance: If there is a trend in the series and the variance appears to increase with the mean, then it may be advisable to transform the data. In particular, if the standard deviation is directly proportional to the mean, a logarithmic transformation is indicated. On the other hand, if the variance changes through time without a trend being present, then a transformation will not help. Instead, a model that allows for changing variance should be considered.

- ▷ Make the seasonal effect additive: If there is a trend in the series and the size of the seasonal effect appears to increase with the mean, then it may be advisable to transform the data so as to make the seasonal effect constant from year to year. The seasonal effect is then said to be additive.

7 Time Series with a trend and seasonality

Inspection of a graph may also suggest the possibility of representing the data as a realization of the process,

$$X_t = T_t + S_t + e_t, \tag{7.1}$$

where T_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 e_t is a random noise component. The error term e_t represents random fluctuations that cause the X_t values to deviate from the average level EX_t.

If the seasonal and noise fluctuations appear to increase with the level of the process then a preliminary transformation of the data is often used to make the transformed data compatible with model (7.1).

Our aim is to estimate and extract the deterministic components T_t and S_t in the hope that the residual or noise component e_t will turn out to be a stationary random process. We can then use the theory of such processes to find a satisfactory probabilistic model for the process e_t , to analyze its properties, and to use it in conjunction with T_t and S_t for purposes of prediction and control of X_t.

An alternative approach is to apply difference operators repeatedly to the data set X_t until the differenced observations resemble a realization of some stationary process Z_t. We can then use the theory of stationary processes for the modelling, analysis and prediction of Z_t and hence of the original process.

The two approaches to trend and seasonality removal, (a) by estimation of T_t and S_t in (7.1) and (b) by differencing the data $\{X_t\}$, will be discussed in some detail.

7.1 Time Series with a trend

In the absence of a seasonal component model (7.1) becomes the following.

Definition 5 A trend model is

$$X_t = T_t + e_t$$

where x_t is the time series in period t, T_t is the trend in time period t, e_t is the error term in time period t.

Consider a straight-line trend

$$T_t = \beta_0 + \beta_1 t, \quad \beta_1 \neq 0.$$

Method 1 (Least squares estimation of T_t). In this procedure we attempt to fit a parametric family of functions, e.g.

$$T_t = \beta_0 + \beta_1 t \tag{7.2}$$

to the data by choosing the parameters, in this illustration a_0 , a_1 and a_2 , to minimize

$$\sum_t (Y_t - T_t)^2.$$

This method needs to assume that the error term e_t satisfies the constant variance and independence assumptions.

Method 2 (Differencing to generate stationary data). As an alternative, we now attempt to eliminate the trend term by differencing.

DEFINITION 6 We define the first difference operator ∇ by

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t,$$

where B is the backward shift operator,

$$BX_t = X_{t-1}, B^i X_t = X_{t-i}, \quad \text{for} \quad i = 1, 2, \ldots.$$

Similarly,

$$\nabla^i X_t = \nabla(\nabla^{i-1} X_t)$$

with $\nabla^0 X_t = X_t$.

Polynomials in B and ∇ are manipulated in precisely the same way as polynomial functions of real functions. For example,

$$\nabla^2 X_t = \nabla(\nabla Xt) = (1 - B)(1 - B)X_t = (1 - 2B + B^2)X_t = X_t - 2X_{t-1} + X_{t-2}.$$

If the operator ∇ is applied to a linear trend function

$$T_t = \beta_1 t + \beta_0,$$

then we obtain the constant function

$$\nabla T_t = \beta_0.$$

In the same way any polynomial trend of degree k can be reduced to a constant by application of the operator ∇^k.

7.2 Time Series with a trend and seasonality

We now consider time series that display seasonal variation and hence model (7.1)

$$X_t = T_t + S_t + e_t.$$

There are two types of seasonal variation.

DEFINITION 7 We say that the time series exhibits constant seasonal variation if the magnitude of the seasonal swing does not depend on the level of the time series.

We say that the time series exhibits varying (increasing) seasonal variation if the magnitude of the seasonal swing depends on the level of the time series.

When a time series displays varying (increasing) seasonal variation, it is common practice to apply a transformation to the data in order to a transformed series that displays constant seasonal variation. A transformation of the form is

DEFINITION 8 Box-Cox transformation: for some $\lambda \geq 0$,

$$z_t = \frac{x_t^\lambda - 1}{\lambda}, \text{if } \lambda > 0$$

It can be proved that as the power λ approaches zero, the transformation is

$$\log(x_t).$$

In fact this is common way to make data display constant seasonal variation. In addition, one may also take the square root transformation ($\lambda = \frac{1}{2}$).

7.3 Modelling seasonality using Dummy variables

The relation between the number of seasons and dummy variables is as follows.

DEFINITION 9 A dummy variable is a variable takes values 0 and 1. Suppose that there are **L** seasons. We need to introduce **L − 1** dummy variables to describe the seasons.

seasons		$D_{1,t}$	$D_{2,t}$	$\cdots$	$D_{L-1,t}$
1	$\longleftrightarrow$	1	0	$\cdots$	0
2	$\longleftrightarrow$	0	1	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
L-1	$\longleftrightarrow$	0	0	$\cdots$	1
L	$\longleftrightarrow$	0	0	$\cdots$	0

or

$$D_{k,t} = \begin{cases} 1, & \text{if time period } t \text{ is season } k \\ 0, & \text{otherwise} \end{cases}$$

For example, (normal) seasons Spring, Summer, Autumn, Winter can be described by

seasons	$D_{1,t}$	$D_{2,t}$	$D_{3,t}$
Spring	1	0	0
Summer	0	1	0
Autumn	0	0	1
Winter	0	0	0

DEFINITION 10 The seasonal factor expressed using dummy variables is

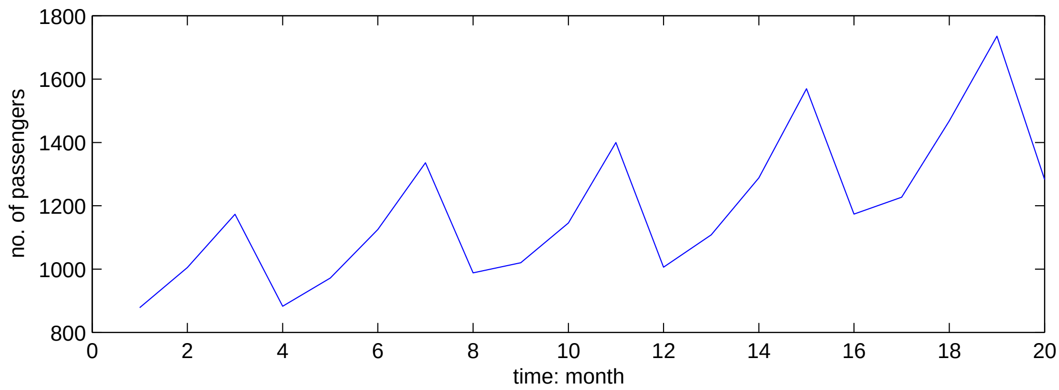
$$S_t = \beta_{s,1} D_{1,t} + \cdots + \beta_{s,L-1} D_{L-1,t}.$$

EXAMPLE 3 (International Airline Passengers: Quarterly Totals, 1956-1960). The time series are

$$878, 1005, 1173, 883, 972, 1125, 1336, 988, 1020, 1146,$$

$$1400, 1006, 1108, 1288, 1570, 1174, 1227, 1468, 1736, 1283$$

.



Our observations: 1. there is trend; 2. the trend is linear; 3. there is seasonality ; 4. the seasonality is increasing.

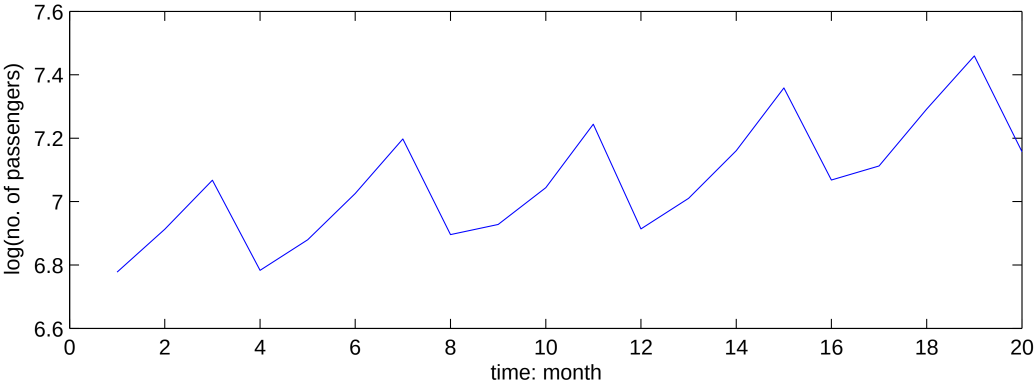
Method 1. Take a transformation of x_t:

$$y_t = \log(x_t).$$

The plot says that taking logarithms of data may equalize the seasonal variation reasonably well.

Fit the following model

$$y_t = T_t + S_t + e_t$$



where $T_t = \beta_0 + \beta_1 t$.

Using dummy variables for the seasonality gives the following regression model

$$y_t = \beta_0 + \beta_1 t + \beta_{s,1} D_{1,t} + \beta_{s,2} D_{2,t} + \beta_{s,3} D_{3,t} + e_t$$

where

$$D_{k,t} = \begin{cases} 1, & \text{if time period } t \text{ is season k} \\ 0, & \text{otherwise} \end{cases}$$

Below we assume that the constant variance and independence regarding the error e_t are satisfied.

The data are listed below

time	x_t	$y_t = \log(x_t)$	const	t	D_1	D_2	D_3	pred	e_t
1	878	6.78	1	1	1	0	0	6.7635	0.0141
2	1005	6.91	1	2	0	1	0	6.909	0.0037
3	1173	7.07	1	3	0	0	1	7.0875	-0.0202
4	883	6.78	1	4	0	0	0	6.7856	-0.0023
5	972	6.88	1	5	1	0	0	6.8525	0.0269
6	1125	7.03	1	6	0	1	0	6.998	0.0275
7	1336	7.20	1	7	0	0	1	7.1765	0.021
8	988	6.90	1	8	0	0	0	6.8746	0.0211
9	1020	6.93	1	9	1	0	0	6.9414	-0.0139
10	1146	7.04	1	10	0	1	0	7.087	-0.0429
11	1400	7.24	1	11	0	0	1	7.2654	-0.0212
12	1006	6.91	1	12	0	0	0	6.9636	-0.0498
13	1108	7.01	1	13	1	0	0	7.0304	-0.0201
14	1288	7.16	1	14	0	1	0	7.1759	-0.0151
15	1570	7.36	1	15	0	0	1	7.3544	0.0044
16	1174	7.07	1	16	0	0	0	7.0525	0.0156
17	1227	7.11	1	17	1	0	0	7.1194	-0.007
18	1468	7.29	1	18	0	1	0	7.2649	0.0268
19	1736	7.46	1	19	0	0	1	7.4434	0.016
20	1283	7.16	1	20	0	0	0	7.1415	0.0154
			1	21	1	0	0	7.2083	
			1	22	0	1	0	7.3539	
			1	23	0	0	1	7.5323	
			1	24	0	0	0	7.2305	

where

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ \cdots & & & & \\ 1 & 20 & 0 & 0 & 0 \end{pmatrix} \qquad Y = \begin{pmatrix} 6.78 \\ 6.91 \\ \cdots \\ 7.16 \end{pmatrix}$$

We have the following calculations

$$(X^TX)^{-1} =$$

$$\begin{pmatrix} 0.4250 & -0.0187 & -0.2562 & -0.2375 & -0.2187 \\ -0.0187 & 0.0016 & 0.0047 & 0.0031 & 0.0016 \\ -0.2562 & 0.0047 & 0.4141 & 0.2094 & 0.2047 \\ -0.2375 & 0.0031 & 0.2094 & 0.4063 & 0.2031 \\ -0.2187 & 0.0016 & 0.2047 & 0.2031 & 0.4016 \end{pmatrix}$$

$$X^TY = \begin{pmatrix} 141.29 \\ 1498.36 \\ 34.71 \\ 35.43 \\ 36.33 \end{pmatrix}.$$

The estimator of β's are

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_{s,1} \\ \hat{\beta}_{s,2} \\ \hat{\beta}_{s,3} \end{pmatrix} = (X^TX)^{-1}X^TY = \begin{pmatrix} 6.6967 \\ 0.0222 \\ 0.0446 \\ 0.1679 \\ 0.3241 \end{pmatrix}.$$

The prediction errors are then

$$\hat{e}_t = y_t - (1, t, D_{1,t}, D_{2,t}, D_{3,t})\hat{\beta}$$

The estimator of $\sigma^2 = Var(\varepsilon_t)$ is

$$\hat{\sigma}^2 = \sum_{t=1}^{n} \hat{e}_t^2/(n - n_p) = 6.7435 \times 10^{-4}.$$

(with $n = 20$ and $n_p = 5$)

Total sum of squares: we have $\bar{y} = 7.0644$ and

$$S_{yy} = \sum_{t=1}^{n} (y_t - \bar{y})^2 = 0.6577$$

Sum squares of prediction errors

$$SSE = \sum_{t=1}^{n} \hat{e}_t^2 = 0.0101$$

Our estimated model is then

$$\begin{aligned} y_t &= 6.70 + 0.02t + 0.04D_{1,t} + 0.17D_{2,t} + 0.32D_{3,t} \\ &\quad (0.0169) \quad (0.0010) \quad (0.0167) \quad (0.0166) \quad (0.0165) \\ R^2 &= 0.9846, \qquad DW = 0.8346 \qquad F = 240.09 \end{aligned}$$

$$\hat{\sigma}^2 = 6.7435 \times 10^{-4}.$$

(i) point Predictions: $\hat{y}_t = X_t\hat{\beta}$: first quarter in year 61: $X_t = (1, 21, 1, 0, 0)$,

$$\hat{y}_{21} = 6.70 + 0.02 * 21 + 0.04 * 1 + 0.17 * 0 + 0.32 * 0 = 7.2$$

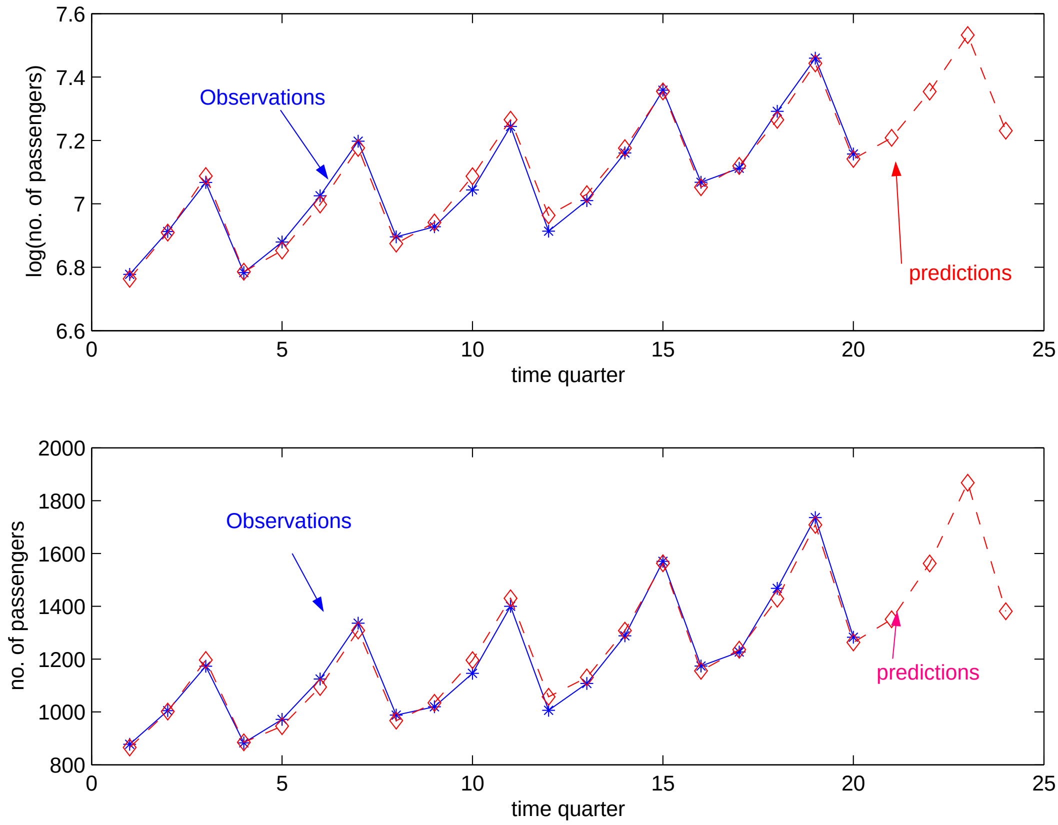
$$\hat{y}_{22} = ??, \hat{z}_{23} \cdots$$

(ii) prediction intervals with 95% confidence:

$$\hat{y}_t \pm 1.96 s \sqrt{1 + X_t (X^T X)^{-1} X_t^T}.$$

first quarter in 61: $7.2 \pm 0.0608 = [7.14, 7.26]$
or y_t: $[e^{7.14}, e^{7.26}] = [1261, 1422]$
second quarter in 61: ??
thrid quarter in 61: ??
fourth quarter in 61: ??

The predictions are shown in the following figure.



Seasonal differencing will be discussed in the future instead of estimating the parameters associated with dummy variables.

8 Autocorrelation and correlogram

For any two variables x and y with N observations, say $x_1, x_2, \cdots, x_N$ and $y_1, \cdots, y_N$, the sample correlation coefficient is given by

$$\hat{\rho} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}.$$

This quantity lies in the range [-1,1] and measures the strength of the linear association between the two variables.

For time series we can define the following.

Definition 11 The sample mean is

$$\hat{\mu}_x = \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t.$$

The sample autocovariance function (SACVF) is

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x}).$$

The sample autocorrelation function (SACF) at lag h is

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}.$$

Definition 12 Correlogram: The plot of sample autocorrelation coefficients $\hat{\rho}(h)$ against k for k =0,1,2,3,$\cdots$. Correlogram is also called sample AutoCorrelation Function (ac.f.)

9 Errors in Forecasting

Example 4 Prediction of unemployment

Measuring Forecasting error

All forecasting situations involve some degree of uncertainty. We recognize this fact by including an irregular component in the description of a time series. The presence of this irregular component, which represents unexplained or unpredictable fluctuations in the data, means that some error in forecasting must be expected.

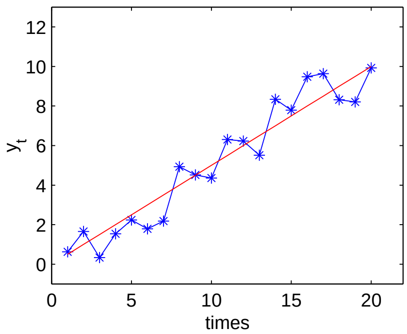
We now consider the problem of measuring forecasting errors. Denote the actual value of the variable of interest in time t as y_t and the predicted value of y_t by $\hat{y}_t$. We then introduce the following concepts.

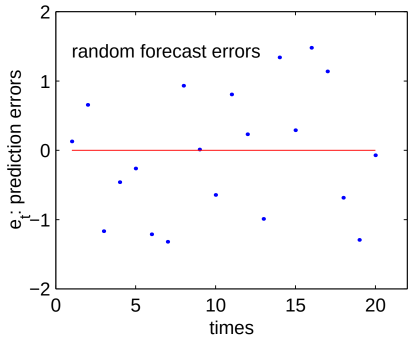
Definition 13

forecast error:	$e_t = y_t - \hat{y}_t$
Plots of forecast error	plot e_t against t
absolute deviation:	$\lvert e_t \rvert = \lvert y_t - \hat{y}_t \rvert$,
squared error:	$e_t^2 = (y_t - \hat{y}_t)^2$.

An examination of forecast errors over time can often indicate whether the forecasting technique being used does or does not match the pattern of the data. For example, if a forecasting technique is accurately forecasting the trend, seasonal, or cyclical components that are present in a time series, the forecast errors should reflect only the irregular component. In such a case, the forecast errors should be purely random. See the following figures and we will demonstrate more in the subsequent lectures.

EXAMPLE 5 $\{y_t$: 0.1001, 1.6900, 2.3156, 2.7119, 3.7902, 3.6686, 4.6908, 2.7975, 4.4802, 4.8433, 3.8959, 6.2573, 5.4435, 8.4151, 6.6949, 8.5287, 8.7193, 8.0781, 7.3293, 9.9408$\}$





If the forecasting errors over time indicate that the forecasting methodology is appropriate (random distribution of errors), it is important to measure the magnitude of the errors so that we can determine whether accurate forecasting is possible.

(i) mean of estimation errors:

$$n^{-1}\sum_{t=1}^{n} e_t;$$

(ii) mean absolute deviation (MAD):

$$n^{-1}\sum_{t=1}^{n} |e_t|;$$

(iii) mean squared errors (MSE):

$$n^{-1}\sum_{t=1}^{n} e_t^2.$$

Example:

Actual value y_t	Predicted value $\hat{y}_t$	Error $e_t = y_t - \hat{y}_t$	Absolute deviation $\lvert e_t \rvert$	squared error e_t^2
25	22	3	3	9
28	30	-2	2	4
29	30	-1	1	1
mean		0	2	4.67

Therefore, for the above three predictions,

$$\begin{aligned}
\text{mean prediction error} &= 0 \\
\text{mean absolute deviation} &= 2 \\
\text{mean squared error} &= 4.67.
\end{aligned}$$

"Mean prediction error" can not be used to assess a prediction method because the positive and negative errors, no matter how large or small, will cancel each other out.

MAD and MSE can be used as criteria for assessing a prediction method. An assessment can be carried out by observations.

EXAMPLE 6 One can propose the following two simple methods for prediction of y_t.

Method A:

$$\hat{y}_t = y_{t-1}$$

Method B:

$$\hat{y}_t = \frac{1}{2}(y_{t-1} + y_{t-2})$$

Actual value y_t	Predicted value of A $\hat{y}_{A,t}$	Predicted value of B $\hat{y}_{B,t}$	Absolute deviation $\lvert e_{A,t}\rvert$	$\lvert e_{B,t}\rvert$
25	–	–	–	–
28	25	–	3	–
29	28	26.5	1	2.5
30	29	28.5	1	1.5
27	30	29.5	3	3
23	27	28.5	4	5.5
20	23	25	3	5

we have

$$\begin{aligned}
MAD_A &= (3 + 1 + 1 + 3 + 4 + 3)/6 = 2.5; \\
MSE_A &= (3^2 + 1^2 + 1^2 + 3^2 + 4^2 + 3^2)/6 = 7.5
\end{aligned}$$

and

$$\begin{aligned}
MAD_B &= (2.5 + 1.5 + 3 + 5.5 + 5)/5 = 3.5; \\
MSE_B &= (2.5^2 + 1.5^2 + 3^2 + 5.5^2 + 5^2)/5 = 14.55.
\end{aligned}$$

Both criteria suggest that method A is more accurate than method B. ■

10 A Scientific view towards statistical forecasting

(i) Statistical forecasting must be based on very strong assumptions: the future behavior of the times series has the same kind "statistics properties" as the observations.

(ii) Usually, statistical methods can only be applied to "short term" forecasts.

(iii) Statistical forecasting must be combined with other prediction methods in practice.