Basic Time Series Analysis

Murthi

Basics: Definitions and Notation

- A time series is a collection of observations made sequentially through time
- Such observations may be denoted by

$$Y_1 \;, Y_2 , Y_3 , \ldots \quad \underset{\uparrow}{Y_t} , \ldots \;, Y_T$$

observation at time t

since data are usually collected at discrete points in time

- The interval between observations can be any time interval (hours within days, days, weeks, months, years, etc).

Examples of time series

- Weekly Sales over time
- Stock price of a company over time
- Malaria incidence or deaths over calendar years
- Daily maximum temperatures
- Hourly records of babies born at a maternity hospital

***Continuous Time series –** temperature, air pollution*
vs
***Discrete Time series –** number of road accidents, sales, market shares*
Time is discrete – week, month etc.

Objectives of a time series

- Description
  - Merely to describe the patterns over time
- Explanation
  - Can the pattern observed over time be explained in terms of other factors or causes?
- Prediction (forecasting)
  - Can past records help us to predict what will happen in the future?
- Improving the past system/behaviour
  - If factors affecting the behaviour of a variable over time can be identified, action may be taken to improve the system, e.g. action over increasing levels of air pollution

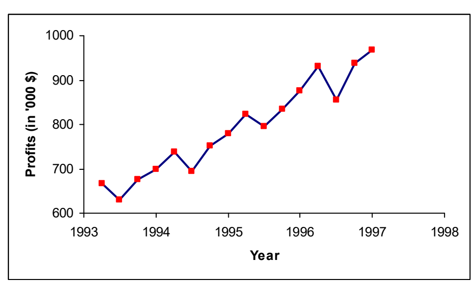
Jumping to conclusions from raw data

Data (interval-scale): Company profits ('000 dollars)
Objective: To study changes in profit figures over consecutive quarters

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	667	631	675	699
2	739	695	751	779
3	823	795	835	875
4	931	855	939	967

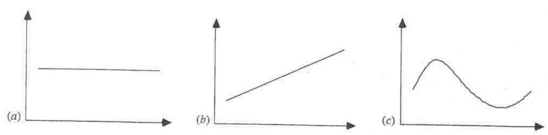
Impression is that the 4th quarter is always higher than the 1st quarter

Take a look again…



Previous impression is largely because there is a general increase over time

Three Aspects of Time Series Behavior
Data: Y_t and t



7

The Moving-Average Model

- The *n*-period **moving average** builds a forecast by averaging the observations in the most recent *n* periods:

$$A_t = (x_t + x_{t-1} + \ldots + x_{t-n+1}) / n$$

- where x_t represents the observation made in period *t*, and A_t denotes the moving average.

Moving Average Example

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Moving Average Example**												
2													
3	*Period*		***4-week Moving Average***						***6-week Moving Average***				
4	*t*	*Observed*	*A(t)*	*Forecast*	*Error*	*Deviation*	*Percent*		*A(t)*	*Forecast*	*Difference*	*Deviation*	*Percent*
5	1	73											
6	2	106											
7	3	76											
8	4	89	86.00										
9	5	106	94.25	86.00	20.00	20.00	19%						
10	6	113	96.00	94.25	18.75	18.75	17%		93.83				
11	7	96	101.00	96.00	0.00	0.00	0%		97.67	93.83	2.17	2.17	2%
12	8	66	95.25	101.00	-35.00	35.00	53%		91.00	97.67	-31.67	31.67	48%
13	9	104	94.75	95.25	8.75	8.75	8%		95.67	91.00	13.00	13.00	13%
14	10	76	85.50	94.75	-18.75	18.75	25%		93.50	95.67	-19.67	19.67	26%
15				85.50						93.50			
16													

9.2 9.3 **9.4**

Measures of Forecast Accuracy

- MSE: the Mean Squared Error between forecast and actual $\quad \text{MSE} = \frac{1}{(u-v+1)}\sum_{t=u}^{v}(F_t - x_t)^2$
- MAD: the Mean Absolute Deviation between forecast and actual $\quad \text{MAD} = \frac{1}{(u-v+1)}\sum_{t=u}^{v}|F_t - x_t|$
- MAPE: the Mean Absolute Percent Error between forecast and actual $\quad \text{MAPE} = \frac{1}{(u-v+1)}\sum_{t=u}^{v}\left|\frac{F_t - x_t}{x_t}\right|$

10

The Exponential Smoothing Model

- **Exponential smoothing** weights recent observations more than older ones.

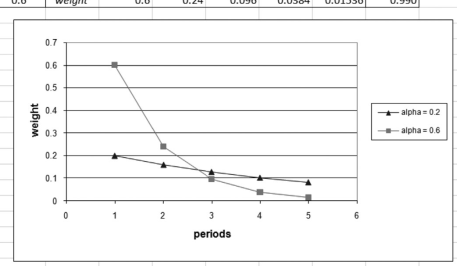
$$S_t = \alpha x_t + (1 - \alpha) S_{t-1}$$

  - Where α (the **smoothing constant)** is some number between zero and one.
  - S_t is the **smoothed value** of the observations (our "best guess" as to the value of the mean)
  - Our forecasting procedure sets the forecast $F_{t+1} = S_t$.

11

Comparison of Weights Placed on *k*-year-old Data

	A	B	C	D	E	F	G	H	I
1	**Smoothing Coefficients**								
2									
3	*Alpha*	*k =*	*1*	*2*	*3*	*4*	*5*	*Total*	
4	0.2	*weight*	0.2	0.16	0.128	0.1024	0.08192	0.672	
5	0.6	*weight*	0.6	0.24	0.096	0.0384	0.01536	0.990	



9.2 9.3 9.4 9.5 **9.6**

12

Summary

- Moving average (MA) and exponential smoothing (ES) models are widely used for routine short-term forecasting.
- They assume that the future will resemble the past. That is, no expected changes.
- However, the exponential smoothing procedure is sophisticated enough to permit representations of a linear trend and a cyclical factor in its calculations.
- Exponential smoothing procedures are adaptive.

13

Why Time Series?

- Modeling an Economic Time Series
  - Observed $y_0, y_1, \ldots, y_t, \ldots$
  - Time domain: A "process"

$$y(t) = ax(t) + by(t-1) + \ldots$$

- Autocorrelation (Serial Correlation)

$$Y_t = b'x_t + \varepsilon_t$$

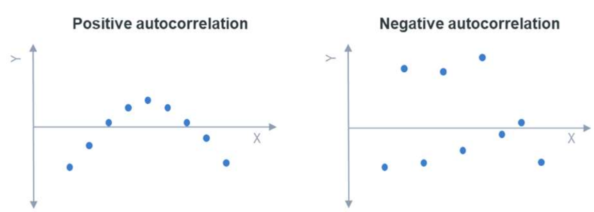
$$Cov(\varepsilon_t, \varepsilon_{t-1}) \neq 0$$

How to detect autocorrelation?

- Durbin-Watson statistic (DW)

```
Proc reg;
model sales = adv comp / DW;
run;
```

- DW ranges from 0-4
  - DW 2-4 negative autocorrelation
  - DW 0-2 positive autocorrelation
  - DW=2 no autocorrelation




```
Proc autoreg;
model sales = adv comp / DWprob;
run;
```

- DWprob gives the p-value associated with DW statistic.

Modeling trends

- Linear trend model
  - $Y_t = c_1 + c_2t$
- Quadratic trend model
  - $Y_t = c_1 + c_2t + c_2t^2$
- Exponential growth curve
  - $Y_t = A*\exp(c_2t)$
  - $\mathrm{Log}(Y_t) = \log(A) + c_2t$
- Autoregressive trend model
  - $Y_t = c_1 + c_2Y_{t-1}$

Modeling seasonality

- Linear trend model with seasonal dummy variables
  - $Y_t = c_1 + c_2 t + c_3 Season_1 + c_4 Season_2 + c_5 Season_3$

Modeling lags of X variables

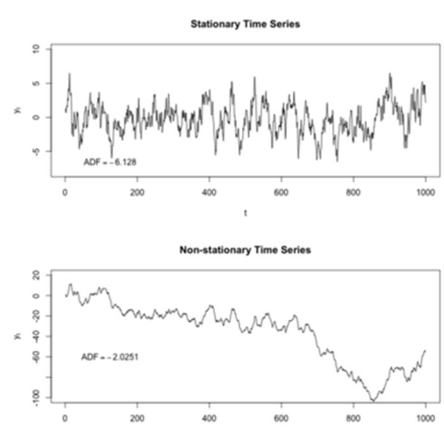
- $S_t = a_0 + a_1 A_t + a_2 A_{t-1} + a_3 A_{t-2} + \varepsilon_t$
- S_t = sales at time t
- A_t = Ad expenditure at time t
- Advertising effects last for a long period of time.
- Above model is called distributed lag model.

Modeling lags of Y variable

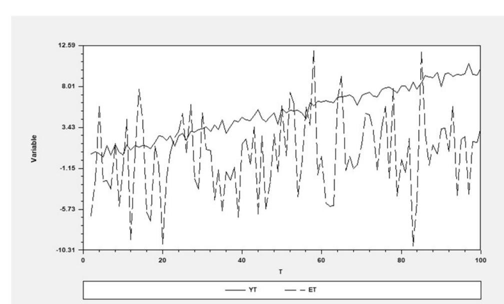
- $Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + a_3 Y_{t-3} + \varepsilon_t$
- This is called a autoregressive model (AR)
- $Y_t = a_0 + a_1 \varepsilon_{t-1} + a_2 \varepsilon_{t-2} + a_3 \varepsilon_{t-3} + \varepsilon_t$
- This is called a moving average model (MA)
- ARMA model has both components.
- PROC ARMA

Stationary Time Series

- Is the underlying process invariant w.r.t. time?
- Yes – then the time series is said to be stationary
- Otherwise it is non-stationary.
- In a stationary process the mean, variance and covariance are all stationary
- If a series is non-stationary, it can be made stationary by first differencing.
- i.e. compute $y_t - y_{t-1}$. Test if this variable is stationary.



Stationary vs. Nonstationary Series



Time Series Analysis - Part 2

murthi

Distributed Lag Models

- Modeling lags of X variables
- $S_t = a_0 + a_1 A_t + a_2 A_{t-1} + a_3 A_{t-2} + \varepsilon_t$
- Example: Advertising effects last for along period of time.

9.2 Finite Distributed Lags

Okun's Law: Change in unemployment is related to growth in GDP

- We can write it as:

$$DU_t = \alpha + \beta_0 G_t + \beta_1 G_{t-1} + \beta_2 G_{t-2} + \cdots + \beta_q G_{t-q} + e_t$$

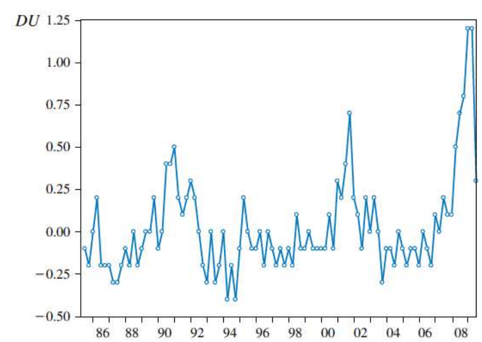
- DU = difference in unemployment levels
- We can calculate the growth in output, G, as:

$$G_t = \frac{GDP_t - GDP_{t-1}}{GDP_{t-1}} \times 100$$

9.2 Finite Distributed Lags

9.2.2 An Example: Okun's Law

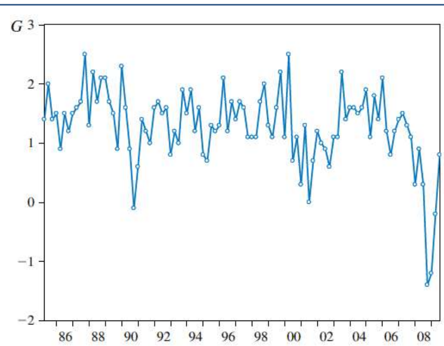
Change in the U.S. unemployment rate 1985Q3 to 2009Q3



9.2 Finite Distributed Lags

9.2.2 An Example: Okun's Law

U.S. GDP growth: 1985Q2 to 2009Q3



9.2 Finite Distributed Lags

An Example: Okun's Law

Spreadsheet of Observations for Distributed Lag Model

t	Quarter	U_t	U_{t-1}	DU_t	G_t	G_{t-1}	G_{t-2}	G_{t-3}
1	1985Q2	7.3	•	•	1.4	•	•	•
2	1985Q3	7.2	7.3	−0.1	2.0	1.4	•	•
3	1985Q4	7.0	7.2	−0.2	1.4	2.0	1.4	•
4	1986Q1	7.0	7.0	0.0	1.5	1.4	2.0	1.4
5	1986Q2	7.2	7.0	0.2	0.9	1.5	1.4	2.0
96	2009Q1	8.1	6.9	1.2	−1.2	−1.4	0.3	0.9
97	2009Q2	9.3	8.1	1.2	−0.2	−1.2	−1.4	0.3
98	2009Q3	9.6	9.3	0.3	0.8	−0.2	−1.2	−1.4

9.2 Finite Distributed Lags

9.2.2 An Example: Okun's Law

Estimates for Okun's Law Finite Distributed Lag Model

Lag Length $q = 3$

Variable	Coefficient	Std. Error	t-value	p-value
Constant	0.5810	0.0539	10.781	0.0000
G_t	−0.2021	0.0330	6.120	0.0000
G_{t-1}	−0.1645	0.0358	−4.549	0.0000
G_{t-2}	−0.0716	0.0353	−2.027	0.0456
G_{t-3}	0.0033	0.0363	0.091	0.9276
Observations = 95		$R^2 = 0.652$		$\hat{\sigma} = 0.1743$

Lag Length $q = 2$

Variable	Coefficient	Std. Error	t-value	p-value
Constant	0.5836	0.0472	12.360	0.0000
G_t	−0.2020	0.0324	−6.238	0.0000
G_{t-1}	−0.1653	0.0335	−4.930	0.0000
G_{t-2}	−0.0700	0.0331	−2.115	0.0371
Observations = 96		$R^2 = 0.654$		$\hat{\sigma} = 0.1726$

Stationary process

- Is the underlying process invariant with respect to time?
- Yes – then the time series is said to be stationary
- Otherwise it is non-stationary.
- In a stationary process the mean, variance and covariance are all stationary
- If a series is non-stationary, it can be made stationary by first differencing.
- i.e. compute $y_t - y_{t-1}$. Check if this variable is stationary.

9.1 Introduction

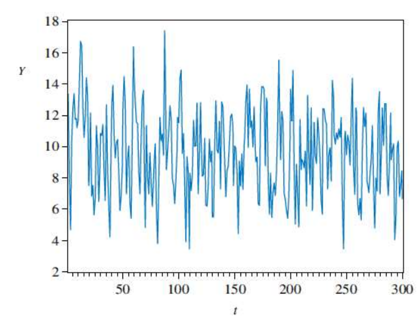
Stationarity

- A stationary variable is one that is not explosive, nor trending, and nor wandering aimlessly without returning to its mean (i.e. not a random walk)
- $\text{Mean}(Y_t) = \text{Mean}(Y_{t+m})$
- $\text{Var}(Y_t) = \text{Var}(Y_{t+m})$
- $\text{Cov}(Y_t, Y_{t+k}) = \text{Cov}(Y_{t+m}, Y_{t+k+m})$
- for any value of m

9.1 Introduction

Stationarity

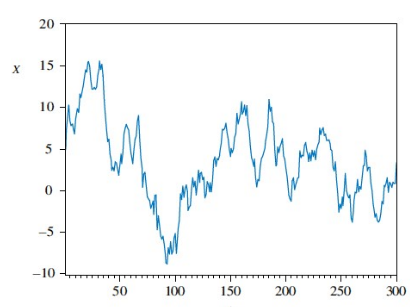
Time series of a stationary variable



9.1 Introduction

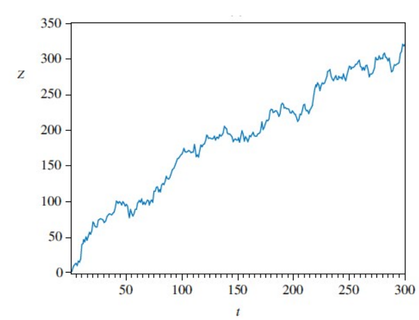
9.1.2a Stationarity

Time series of a nonstationary variable that is "slow-turning" or "wandering"



9.1.2a Stationarity

Time series of a nonstationary variable: Trend



Random walk Model

- $Y_t = Y_{t-1} + \varepsilon_t$ $\quad E(\varepsilon_t) = 0, E(\varepsilon_t, \varepsilon_s) = 0$
- Generated for example by flips of a coin.
- If heads, value = 1, if tails, value = -1
- Forecast: $Y_{t+1} = Y_t$
  $Y_{t+2} = Y_t$

But the standard error of the forecast increases over time

If we know a series follows a random walk model, then there is no point in fitting a model.

Checking for Random walk

$Y_t = a + bt + \rho Y_{t-1} + \varepsilon_t$ $\quad E(\varepsilon_t) = 0, E(\varepsilon_t, \varepsilon_s) = 0$

$Y_t - Y_{t-1} = a + bt + (\rho-1)Y_{t-1} + \varepsilon_t$

If $\rho=1$, then series is random walk and non-stationary

- Dickey Fuller tests (DF) – unit root tests
- Augmented Dickey Fuller tests (ADF)

Regression one random walk against another will give spurious results.

If a series follows random walk, differencing might remove that.

Checking for white noise

- Bartlett's test: to test $\rho_k = 0$
- Sample autocorrelation coefficients are approximately N(0,1/sqrt(T))
- Q-test : joint test that all $\rho_k = 0$

9.3 Serial Correlation

Computing Autocorrelation

- Recall that the population correlation between two variables x and y is given by:

$$\rho_{xy} = \frac{\operatorname{cov}(x,y)}{\sqrt{\operatorname{var}(x)\operatorname{var}(y)}}$$

- For autoregressive terms

$$\rho_1 = \frac{\operatorname{cov}(G_t,G_{t-1})}{\sqrt{\operatorname{var}(G_t)\operatorname{var}(G_{t-1})}} = \frac{\operatorname{cov}(G_t,G_{t-1})}{\operatorname{var}(G_t)}$$

$$\rho_k = \operatorname{cov}(G_t, G_{t-k}) / \operatorname{var}(G_t)$$

9.3 Serial Correlation

- Applying this to our problem, we get for the first four autocorrelations:

$$r_1 = 0.494 \quad r_2 = 0.411 \quad r_3 = 0.154 \quad r_4 = 0.200$$

9.3 Serial Correlation

- How do we test whether an autocorrelation is significantly different from zero?
  - The null hypothesis is H_0: $\rho_k = 0$
  - A suitable test statistic is:

$$Z = \frac{r_k - 0}{\sqrt{1/T}} = \sqrt{T}\, r_k \sim N(0,1)$$

9.3 Serial Correlation

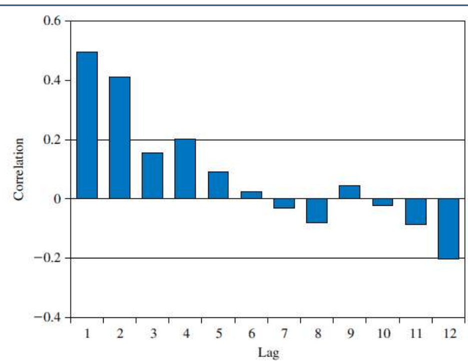
- For our problem, we have:

$$Z_1 = \sqrt{98} \times 0.494 = 4.89, \quad Z_2 = \sqrt{98} \times 0.414 = 4.10$$
$$Z_3 = \sqrt{98} \times 0.154 = 1.52, \quad Z_4 = \sqrt{98} \times 0.200 = 1.98$$

  - We conclude that G, the quarterly growth rate in U.S. GDP, exhibits significant serial correlation at lags one and two

9.3 Serial Correlation

FIGURE 9.6 Correlogram for *G*



Lagrange Multiplier test

If e_t and e_{t-1} are correlated, then one way to model the relationship between them is to write:

$$e_t = \rho e_{t-1} + v_t$$

We can substitute this into a simple regression equation:

$$y_t = \beta_1 + \beta_2 x_t + \rho e_{t-1} + v_t$$

9.4 Other Tests for Serially Correlated Errors

- To derive the relevant auxiliary regression for the autocorrelation LM test, we write the test equation as:

$$y_t = \beta_1 + \beta_2 x_t + \rho \hat{e}_{t-1} + v_t$$

  - But since we know that $y_t = b_1 + b_2 x_t + \hat{e}_t$, we get:

$$b_1 + b_2 x_t + \hat{e}_t = \beta_1 + \beta_2 x_t + \rho \hat{e}_{t-1} + v_t$$

9.4 Other Tests for Serially Correlated Errors

- Rearranging, we get:

$$\begin{aligned}\hat{e}_t &= (\beta_1 - b_1) + (\beta_2 - b_2) x_t + \rho \hat{e}_{t-1} + v_t \\ &= \gamma_1 + \gamma_2 x_t + \rho \hat{e}_{t-1} + v\end{aligned}$$

  - If H_0: $\rho = 0$ is true, then LM = T x R^2 has an approximate $\chi^2_{(1)}$ distribution
    - T and R^2 are the sample size and goodness-of-fit statistic, respectively, from least squares estimation of Eq. 9.26

9A
The Durbin-Watson Test

Durbin Watson test

- We can now write:

$$d \approx 2(1 - r_1)$$

  - If the estimated value of ρ is $r_1 = 0$, then the Durbin-Watson statistic $d \approx 2$
    - This is taken as an indication that the model errors are not autocorrelated
  - If the estimate of ρ happened to be $r_1 = 1$ then $d \approx 0$
    - A low value for the Durbin-Watson statistic implies that the model errors are correlated, and $\rho > 0$

Stochastic time series models

- Autoregressive model AR(3)
- $S_t = a_0 + a_1 S_{t-1} + a_2 S_{t-2} + a_3 S_{t-3} + \varepsilon_t$

- Moving average model MA(3)
- $S_t = a_0 + a_1 \varepsilon_{t-1} + a_2 \varepsilon_{t-2} + a_3 \varepsilon_{t-3} + \varepsilon_t$

- A model with both components is called a ARMA(1,1) model.
- $S_t = a_0 + a_1 S_{t-1} + a_2 \varepsilon_{t-1} + \varepsilon_t$

- First de-trend and de-seasonalize the Y_t variable
- $Y_t - Y_{t-12} = a + b\,t + \varepsilon_t$
- Check if Y_t follows a random walk.
- Look at ACF plots: if ACF drops to zero after q lags, then it indicates a MA(q) model.
- Look at PACF plots: if PACF drops to zero after p lags, then it indicates a AR(p) model.
- If both ACF and PACF do not become zero, it indicates ARMA (p,q) model
- If ACF and PACF are zero at all t periods, it indicates a white noise process.