ARMA Models in Time Series Lecture

Dr. Goutam Chakraborty

1

Topics in this session

- From exponential smoothing to more complex models for univariate time series
- What is stationarity?
- What is ARMA?
- How are these models formulated and estimated?

2

From Exponential Smoothing to More Complex (Parameterized) Models

- All exponential smoothing (simple, double, seasonal..) forecasts can be constructed by simple algorithm
  - In fact, most could be done by hand!
  - These often work well and sometimes perform better than more complex models
- Complex (Parameterized) models include:
  - ARMA (**AR**: Auto regressive, **MA**: Moving average)
  - AR**I**MA (with differencing)
  - ARIMA**X** (with **X**: exogenous variable)
  - UCM (Unobserved Components Model)
  - State Space Models
- Other approaches: Spectral Analysis

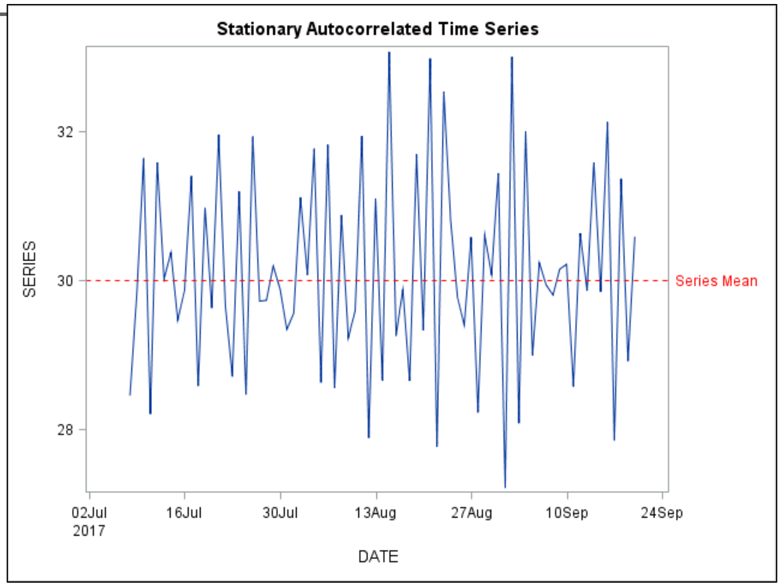
3

Conceptual Issues

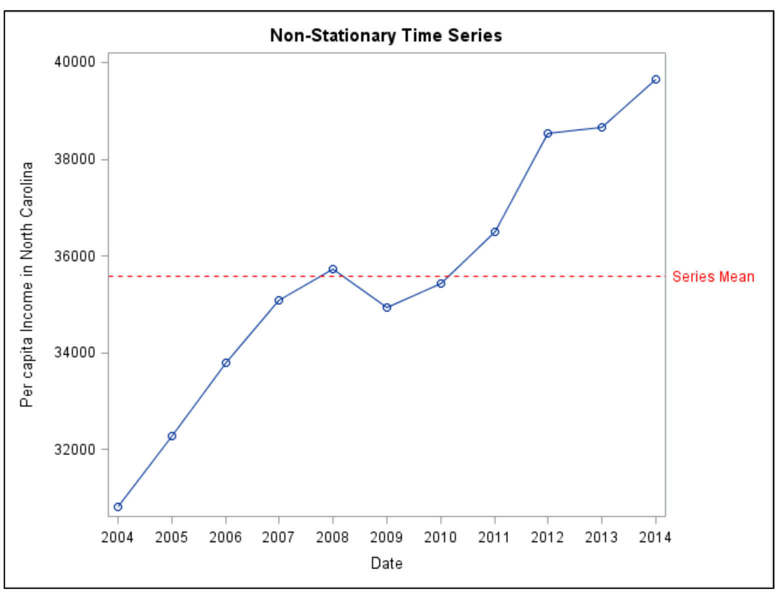
- Why we care so much about stationarity?
- Stationary time series : A stochastic process that generates data series with constant mean and constant variance
  - What if a series has trend?
  - What if a series show non-constant variance?

4

Visualizing Stationary Data



Visualizing Non-Stationary Data



History of ARMA models

- **AR** (Autoregressive) means a variable Y is modeled as a function of its own past values (lagged values of Y)
  - The idea is everything you want to know about Y is contained in its past values – so, why not just regress Y on its own past values?
    - Question is how many past values? 1, 2,…
  - The idea has been around but were not widely used until Box and Jenkins (1972) came up with their theory of combining AR with MA models in a way that can capture any type of series
  - In Box-Jenkins notation, ARIMA(p,d,q) are used to capture degree of autoregressive (p), differencing (d) and degree of MA (q)
  - ARIMA (1,0,2) means AR(1), no differencing and MA(2)
  - ARIMA (2,1,3) means AR(2), first-difference and MA(3)

7

What Is ARIMA?

AR
- AutoRegressive
  - Current values are related to past values.

I
- Integrated
  - Differenced values between successive time points can be modeled.

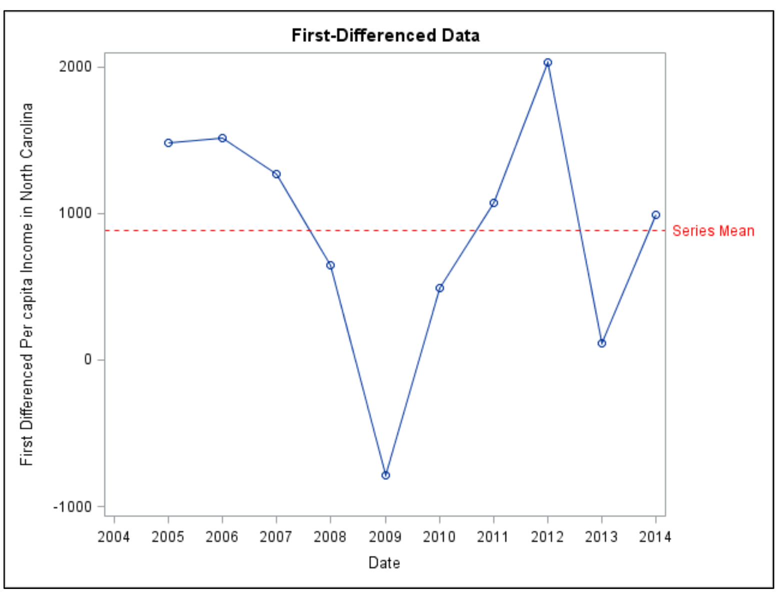
MA
- Moving Average
  - Current values are related to past estimation errors (that is, shocks).

8

First Differencing Example

Year	Income	Lag(Income)	First Difference
2004	$ 30,818	⚠	⚠
2005	$ 32,296	$ 30,818	$ 1,478
2006	$ 33,808	$ 32,296	$ 1,512
2007	$ 35,076	$ 33,808	$ 1,268
2008	$ 35,725	$ 35,076	$ 649
2009	$ 34,942	$ 35,725	$ -783
2010	$ 35,435	$ 34,942	$ 493
2011	$ 36,508	$ 35,435	$ 1,073
2012	$ 38,538	$ 36,508	$ 2,030
2013	$ 38,653	$ 38,538	$ 115
2014	$ 39,646	$ 38,653	$ 993

Visualizing First-Differenced Data



Conceptual Issues (Continued)

- What if a series has seasonality?
  - If we know the seasonal order (such as retail sales spike in December), then differencing by the lag of seasonal order (?) will handle it
- Data =Trend + Season + Cycle + Irregular
- Cycle + Irregular = Data – Trend – Season
- Cycle + Irregular = (Approx.) Stationary Process

Stationary Process

- Series Y_t is stationary if:

$\mu_t = \mu$,	constant for all t
$\sigma_t = \sigma$,	constant for all t
$\rho(Y_t, Y_{t+h}) = \rho_h$	does not depend on t

- WN (White Noise) is a special example of a stationary process. It is a series that varies randomly around its mean

Models For a Stationary Process

- Autoregressive Process, AR(p)
- Moving Average Process, MA(q)
- Autoregressive Moving Average Process, ARMA(p, q)

13

Parameters of ARMA Models

Specification Parameters

ϕ_k Autoregressive Process Parameter

θ_k Moving Average Process Parameter

Characterization Parameters

ρ_k Autocorrelation Coefficient

ϕ_{kk} Partial Autocorrelation Coefficient

14

AR Process

- AR (1) : $(Y_t - \mu) = \phi_1 (Y_{(t-1)} - \mu) + \varepsilon_t$

$-1 < \phi_1 < 1$ (stationarity condition), ε_t is a WN (σ)

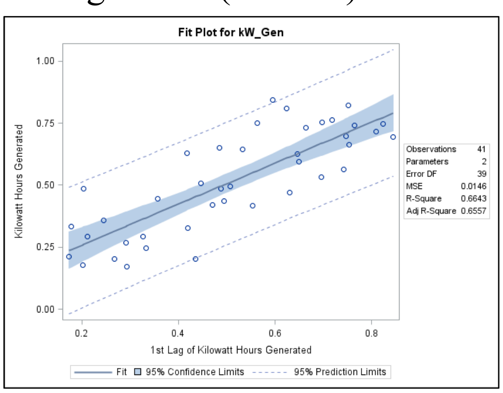
- $Y_t = \phi_0 + \phi_1 Y_{(t-1)} + \varepsilon_t$

this is the regression of Y on Lag(1) of Y: $Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$

- A high value away from mean in last week will imply a high value away from mean in this week.
- This will mean that the autocorrelation will be **high** for Lag 1!

Regression of Y on Past Y: *Autoregression*

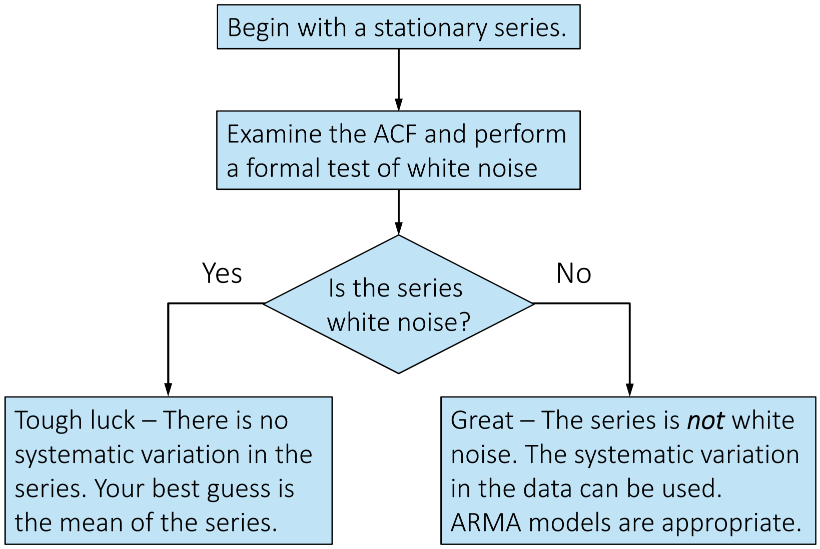
- Reminder (OLS regression): $$Y = \beta_0 + \beta_1 X + \varepsilon$$
- Autoregressive (Order 1) Model: $$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$



$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

$$\phi_0 = \mu(1 - \phi_1)$$

ARMA Models: Initial Process Flow



17

AR Process (Contd.)

- AR (2) : $(Y_t - \mu) = \phi_1 (Y_{(t-1)} - \mu) + \phi_2 (Y_{(t-2)} - \mu) + \varepsilon_t$

  $\phi_2 + \phi_1 < 1$, $\phi_2 - \phi_1 < 1$, $-1 < \phi_2 < 1$ (stationarity condition), ε_t is a WN (σ)

- $Y_t = \phi_0 + \phi_1 Y_{(t-1)} + \phi_2 Y_{(t-2)} + \varepsilon_t$

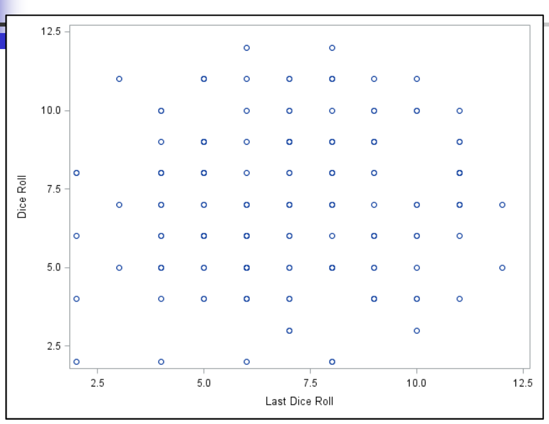
This is the regression of **Y on Lag(1) and Lag(2) of Y**

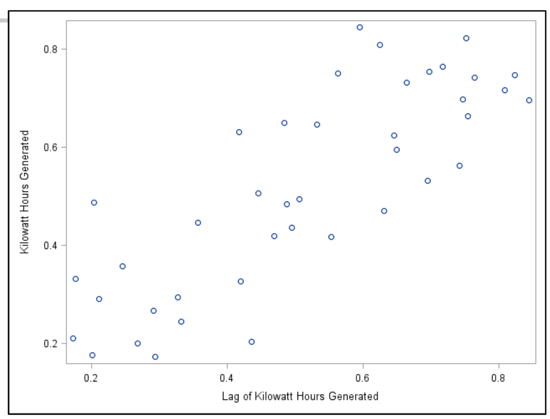
18

Determining Autoregressive Order

- Confirm that the series is stationary and contains autocorrelation.
- Determine which lagged values (Y_{t-1}, Y_{t-2}, Y_{t-3}, and so on) are correlated with the current value (Y_t), and adjust for the autocorrelation of all lower order lags.
- The partial autocorrelation function plot (PACF) helps determine this by answering the following questions:
  - Is there significant autocorrelation between Y_t and Y_{t-1}?
  - Is there significant autocorrelation between Y_t and Y_{t-2}, holding constant the autocorrelation between Y_t and Y_{t-1}?

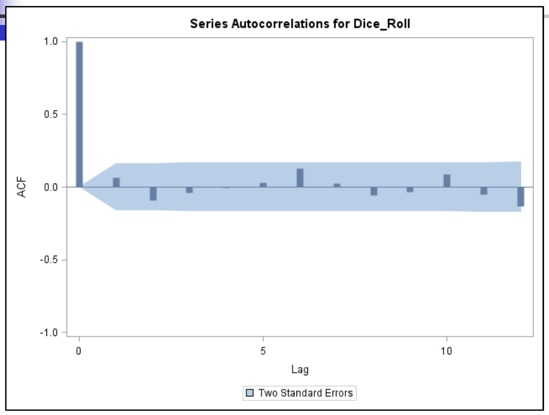
Autocorrelation Scatter Plots

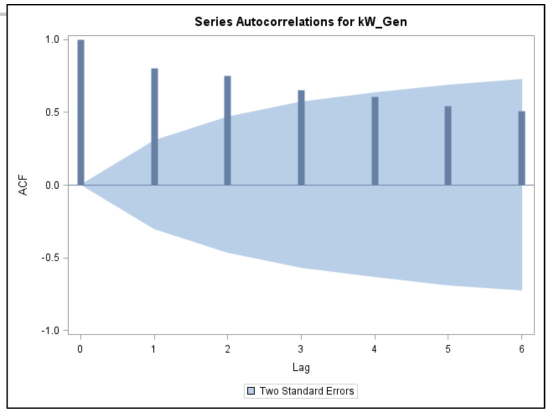




- Autocorrelation is the correlation of present values versus lagged values.
- Autocorrelation between the present value and the first lagged value is called *first order* autocorrelation.

Autocorrelation Plots

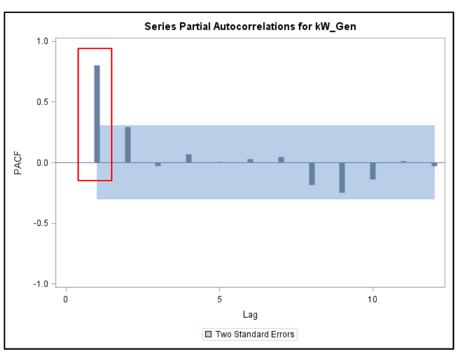


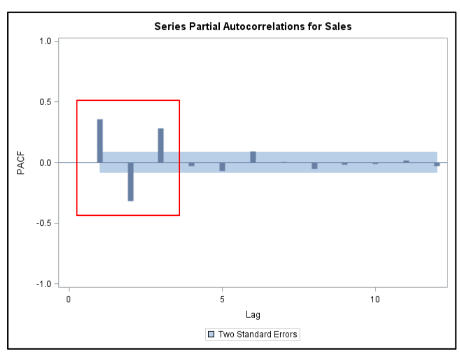


- The autocorrelation plot enables you to see the autocorrelation at multiple lags.
- The blue range indicates 95% confidence intervals for each lag.

Partial Autocorrelation Function Plot (PACF)

- Significant spikes in the PACF are the most important source of information for identifying **an autoregressive** series.





Autoregressive versus Moving Average Models

- First Order Autoregressive Model AR(1)

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

- Y_t is a function of the previous value plus some error.

- First Order Moving Average MA(1)

$$Y_t = \theta_0 - \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

  - Y_t is a function of its immediately previous shock plus error (significant autocorrelation between Y_t and ε_{t-1}).

MA(1) Model

$$Y_t = \theta_0 - \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

- A moving average model is a function of past shocks going back (q) periods.
- Past shocks are uncorrelated with **other past** shocks. $\varepsilon_{t-n} \sim iid\ N(0, \sigma^2)$
- Unlike AR(p) models, MA(q) models are used to model short-lived, abrupt patterns in the data.
- Forecasts after (q) periods immediately revert to the mean of the series.

Moving Average of Order *q*

$$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

- Moving average models are a sum of stationary processes.
- As a result, *pure* moving average models are always stationary.
- Time series containing autoregressive *and* moving average variations are called ARMA models.

MA Process

- MA (1) : $Y_t - \mu = \varepsilon_t + \theta_1 \varepsilon_{(t-1)}$

  $-1 < \theta_1 < 1$

  (invertibility condition)

- MA (2) : $Y_t - \mu = \varepsilon_t + \theta_1 \varepsilon_{(t-1)} + \theta_2 \varepsilon_{(t-2)}$

  $\theta_2 + \theta_1 > -1, \quad \theta_2 - \theta_1 > -1, \quad -1 < \theta_2 < 1$

  (invertibility condition)

ε_t is a WN (σ)

ARIMA Ordering - ARIMA(p,d,q)

AR
- Autoregressive order = *p*

I
- Differencing order = *d*

MA
- Moving average order = q

Note: ARMA models are ARIMA models with *d*=0 and are denoted ARMA(*p*,*q*).

27

sas

ARMA (p, q) Models

- ARMA(1, 1):

$$(Y_t - \mu) = \phi_1 (Y_{(t-1)} - \mu) + \varepsilon_t + \theta_1 \varepsilon_{(t-1)}$$

- ARMA(2, 1):

$$(Y_t - \mu) = \phi_1 (Y_{(t-1)} - \mu) + \phi_2 (Y_{(t-2)} - \mu) + \varepsilon_t + \theta_1 \varepsilon_{(t-1)}$$

- ARMA(1, 2):

$$(Y_t - \mu) = \phi_1 (Y_{(t-1)} - \mu) + \varepsilon_t + \theta_1 \varepsilon_{(t-1)} + \theta_2 \varepsilon_{(t-2)}$$

28

How to Figure out Which is the Right Process?

- Challenges include:
  - Too many terms mean model is unnecessary complex
  - Too few terms mean model is not complete
- One way:
  - Test all possible AR process and select the one that works the best using criteria such as MAPE, AIC, SBC,..
- Another way:
  - Use diagnostics from models to figure out if complex models are worth it
    - ACF plots (Autocorrelation)
    - PACF plots (Partial autocorrelation)