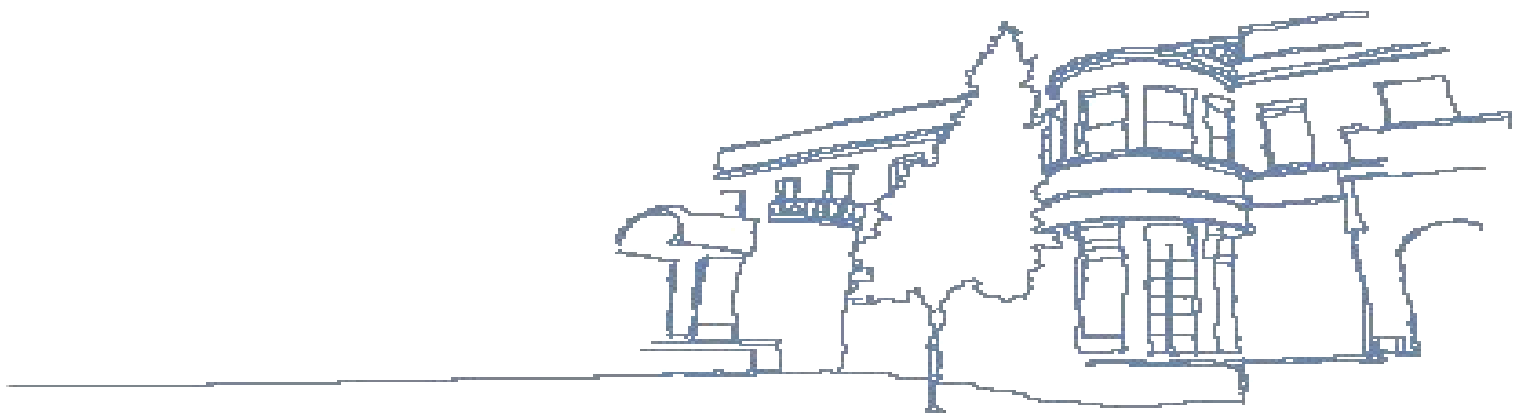
MBD SEPT 2022

BIG DATA & ARTIFICIAL INTELLIGENCE IN OPERATIONS MANAGEMENT

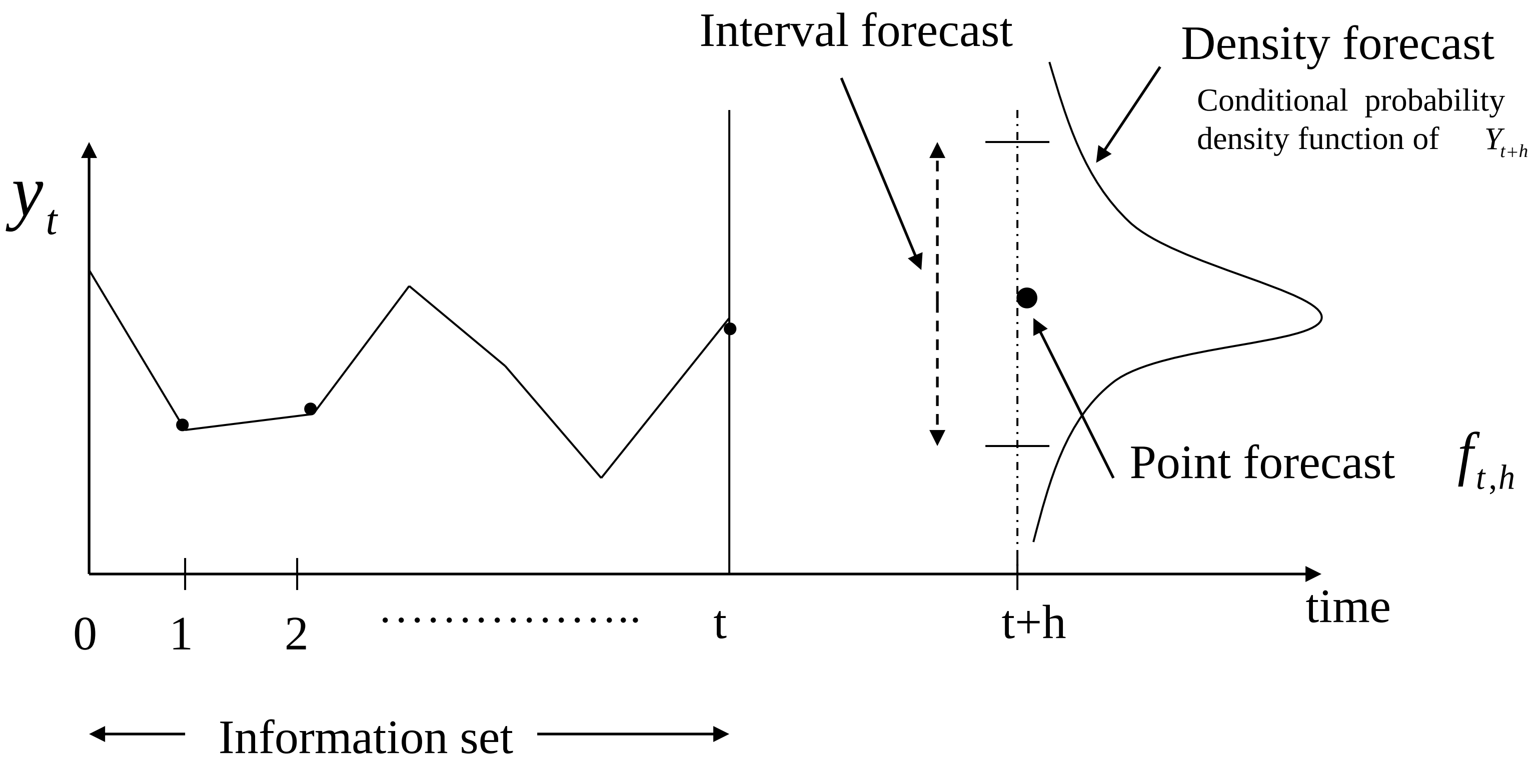
Session 5: The Box-Jenkins Methodology



Lorenzo Pascual

The objective of the forecaster - reminder

The Forecasting Problem



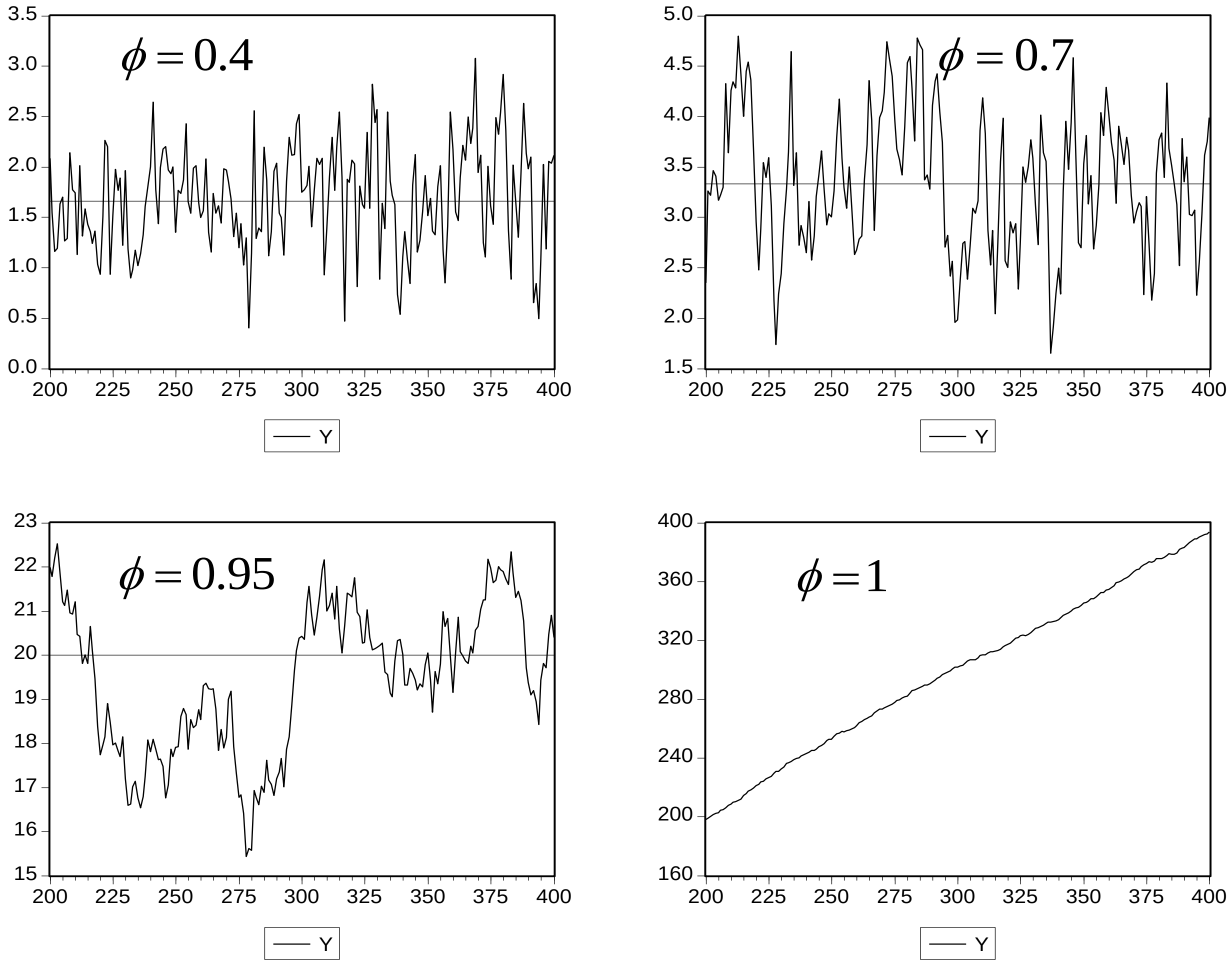
The Box-Jenkins Methodology - reminder

The Box-Jenkins methodology to fit an ARIMA model is as follows:

1. Decide about the stationary transformation
   a. Plot the series
   b. Correlogram (ACF y PACF)
   c. Dickey-Fuller test
2. Decide the stationary ARMA model that better fits the estimated autocorrelations
3. Estimate the parameters by Maximum Likelihood
4. Obtain the residuals and check whether they are white noise
5. Forecasting: obtain point and interval forecasts

Forecasting with AR models - reminder

Autoregressive Processes AR(1)



Forecasting with AR models - reminder

Q2. What do the Autocorrelation Functions look like?

Sample: 2 1000
Included observations: 999

Autocorrelation	Partial Correlation		AC	PAC
		1	0.334	0.334
		2	0.082	-0.034
		3	0.018	0.002
		4	0.043	0.044
		5	-0.022	-0.056
		6	0.006	0.033
		7	0.060	0.058
		8	0.016	-0.031
		9	0.041	0.053
		10	0.051	0.024
		11	0.028	-0.006
		12	0.001	-0.001
		13	0.036	0.037
		14	0.008	-0.023
		15	-0.026	-0.023

$\phi = 0.4$

Sample: 2 1000
Included observations: 999

Autocorrelation	Partial Correlation		AC	PAC
		1	0.732	0.732
		2	0.546	0.021
		3	0.400	-0.015
		4	0.279	-0.030
		5	0.168	-0.058
		6	0.085	-0.025
		7	0.032	-0.003
		8	-0.026	-0.057
		9	-0.071	-0.033
		10	-0.084	0.011
		11	-0.060	0.057
		12	-0.048	-0.010
		13	-0.060	-0.053
		14	-0.063	-0.013
		15	-0.054	0.006

$\phi = 0.7$

Sample: 2 1000
Included observations: 999

Autocorrelation	Partial Correlation		AC	PAC
		1	0.951	0.951
		2	0.904	-0.006
		3	0.861	0.009
		4	0.817	-0.019
		5	0.774	-0.016
		6	0.734	0.005
		7	0.696	-0.002
		8	0.658	-0.022
		9	0.622	0.013
		10	0.590	0.010
		11	0.561	0.011
		12	0.534	0.017
		13	0.508	-0.014
		14	0.481	-0.016
		15	0.457	0.007

$\phi = 0.95$

Forecasting with MA models

In this section we analyze the general MA(q) model answering the following three questions

1. What does a time series of a MA process look like?
2. What do the corresponding autocorrelation functions look like?
3. What is the optimal forecast?

Let us proceed to answer all these questions for the **MA(1) process**

$$\boldsymbol{Y_t = \mu + \theta\varepsilon_{t-1} + \varepsilon_t}$$

where $\boldsymbol{\varepsilon_t}$ is a White-Noise process

Forecasting with MA models

Q1. What does a time series of a MA(1) process look like?

We simulate time series of such a process for different values of θ

$$\boldsymbol{Y_t = 2 + \theta\varepsilon_{t-1} + \varepsilon_t}$$

$$\textbf{for } \boldsymbol{\theta = 0.05, 0.5, 0.95 \; and \; 2, and \; \varepsilon_t \sim N(0, 0.25)}$$

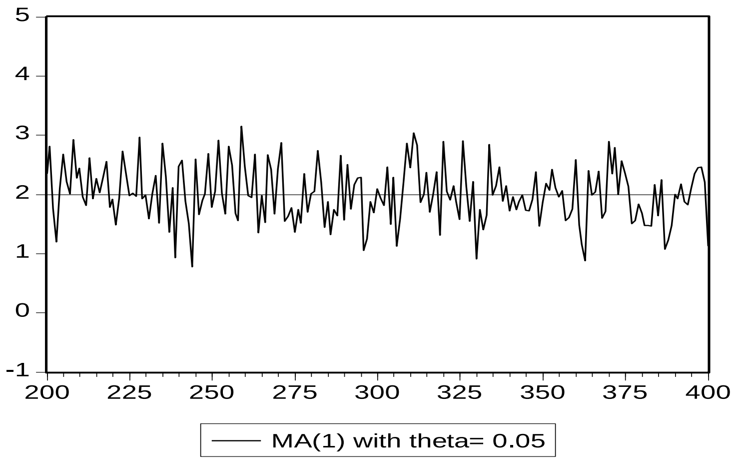
- We can see in the following graphs that
  - When $\theta \to 0$ the MA(1) model becomes a white noise process
  - The unconditional mean of the process is exactly the value of the constant of the model $\mu = 2$
  - The variance of the process is directly proportional to the magnitude of the parameter θ

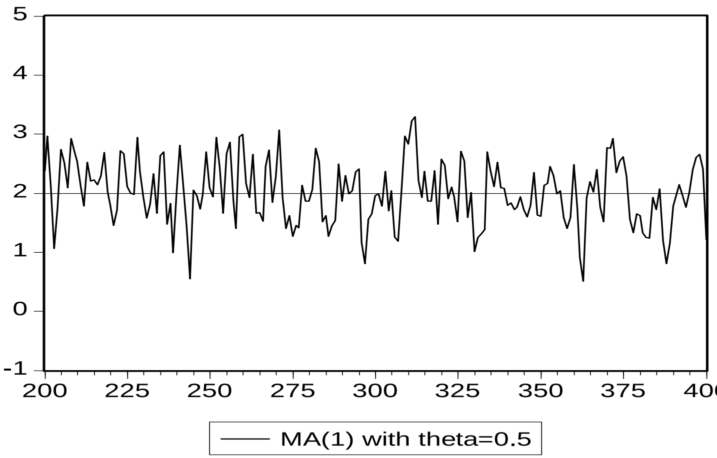
$$\boldsymbol{E(Y) = \mu}$$

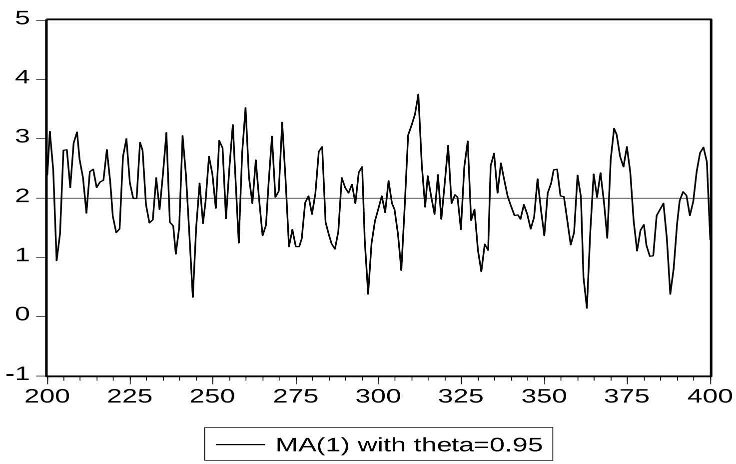
$$\boldsymbol{Var(Y) = \sigma_Y^2 = E(Y_t - \mu)^2 = \left(1 + \theta^2\right)\sigma_\varepsilon^2}$$

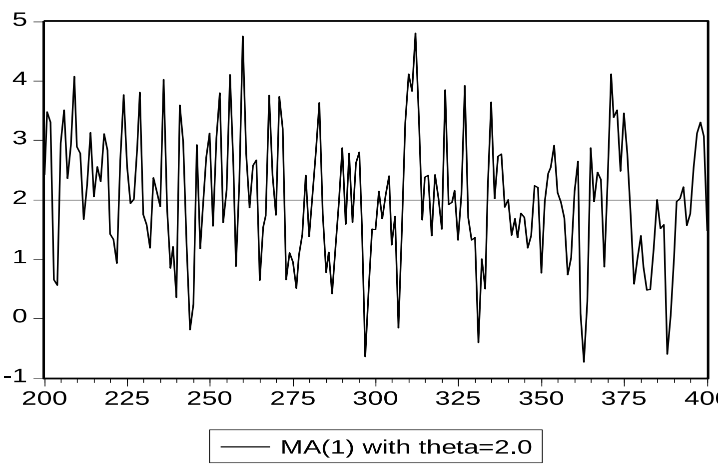
Forecasting with MA models

Q1. What does a time series of an MA(1) process look like?









Forecasting with MA models

Q2. What do the autocorrelation functions look like?

Sample: 1 2000
Included observations: 1999

$\theta = 0.05$

Lag	AC	PAC
1	0.072	0.072
2	-0.027	-0.032
3	-0.061	-0.057
4	-0.024	-0.017
5	0.016	0.016
6	-0.004	-0.011
7	0.022	0.022
8	0.003	0.001
9	0.000	0.000
10	-0.011	-0.009

Sample: 1 2000
Included observations: 1999

$\theta = 0.5$

Lag	AC	PAC
1	0.406	0.406
2	-0.038	-0.243
3	-0.076	0.053
4	-0.038	-0.039
5	0.006	0.029
6	0.009	-0.015
7	0.020	0.030
8	0.011	-0.012
9	-0.003	0.002
10	-0.010	-0.009

Sample: 1 2000
Included observations: 1999

$\theta = 0.95$

Lag	AC	PAC
1	0.498	0.498
2	-0.042	-0.386
3	-0.081	0.220
4	-0.042	-0.183
5	0.003	0.157
6	0.013	-0.127
7	0.019	0.131
8	0.013	-0.112
9	-0.004	0.092
10	-0.010	-0.088

Sample: 1 2000
Included observations: 1999

$\theta = 2$

Lag	AC	PAC
1	0.406	0.406
2	-0.039	-0.243
3	-0.076	0.053
4	-0.038	-0.039
5	0.006	0.030
6	0.009	-0.016
7	0.020	0.030
8	0.010	-0.013
9	-0.004	0.002
10	-0.010	-0.008

Forecasting with MA models

Q2. What do the autocorrelation functions look like?

The most prominent features of the autocorrelation functions of a MA(1) are

✓ There is only one spike in the ACF that is different from zero $\rho_1 \neq 0$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta\sigma_\varepsilon^2}{(1+\theta^2)\sigma_\varepsilon^2} = \frac{\theta}{(1+\theta^2)}$$

$$\rho_k = 0, for\ k > 1$$

✓ The magnitude of the spike is directly proportional to the magnitude of θ and has the same sign than θ

THE MA(1) PROCESS IS COVARIANCE-STATIONARY

✓ The PACF decreases toward zero in an alternating and a very smooth fashion

Forecasting with MA models

Q3. What is the optimal forecast?

Forecasting horizon h=1

Optimal point forecast

$$f_{t,1} = E(Y_{t+1}|I_t) = E_t(Y_{t+1}) = \mu + \theta\varepsilon_t$$

One-perior ahead forecast error

$$e_{t,1} = Y_{t+1} - f_{t,1} = \varepsilon_{t+1}$$

Uncertainty associated with the forecast

$$\sigma^2_{t+1/t} = Var(Y_{t+1}|I_t) = E\left(Y_{t+1} - f_{t,1}|I_t\right)^2 = E\left(e^2_{t,1}\right) = \sigma^2_\varepsilon$$

Density forecast (if ε_t is normally distributed)

$$f(Y_{t+1}|I_t) \to N\left(f_{t,1}, \sigma^2_{t+1/t}\right) = N(\mu + \theta\varepsilon_t, \sigma^2_\varepsilon)$$

and a 95% confidence interval is $f_{t,1} \pm 1.96\sigma_{t+1/t}$

Forecasting with MA models

Q3. What is the optimal forecast?

Forecasting horizon h=2

Optimal point forecast

$$f_{t,2} = E(Y_{t+2}|I_t) = E_t(Y_{t+2}) = \mu$$

One-perior ahead forecast error

$$e_{t,2} = Y_{t+2} \text{-} f_{t,2} = \theta\varepsilon_{t+1} + \varepsilon_{t+2}$$

Uncertainty associated with the forecast

$$\sigma^2_{t+2/t} = Var(Y_{t+2}|I_t) = E\left(e^2_{t,2}\right) = \sigma^2_\varepsilon(1+\theta^2)$$

Density forecast (if ε_t is normally distributed)

$$f(Y_{t+2}|I_t) \to N\left(f_{t,2}, \sigma^2_{t+2/t}\right)$$

and a 95% confidence interval is $f_{t,2} \pm 1.96\sigma_{t+2/t}$

Forecasting with MA models

The analysis of an MA(q) process proceeds in the same fashion as that of an MA(1)

Let us focus on an **MA(2) process** to illustrate that the generalization of MA processes is straightforward

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

where ε_t is a White-Noise process

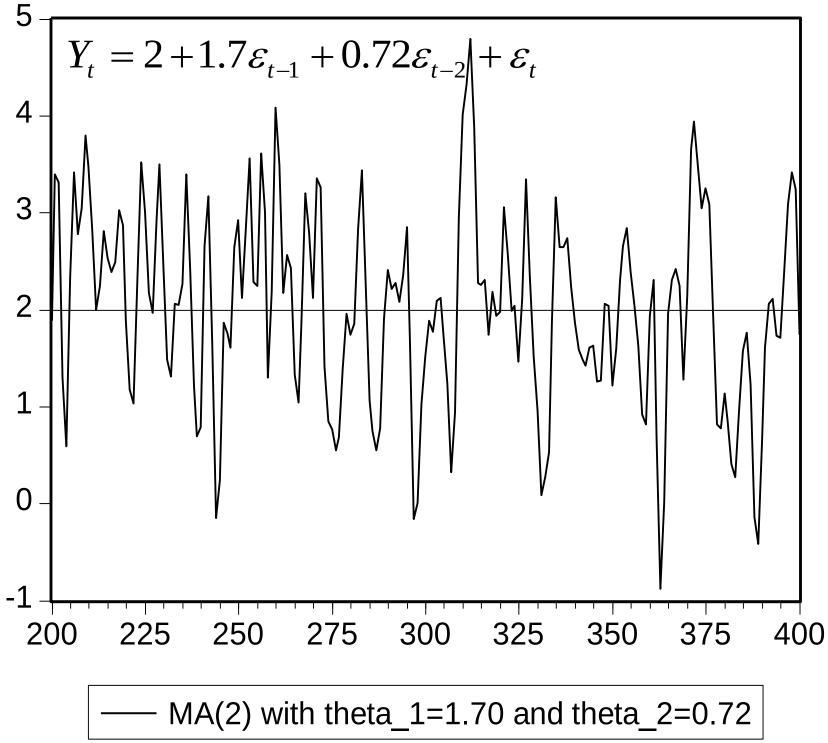
Let us proceed to answer the three questions for different values of the parameters and $\varepsilon_t \sim N(0, 0.25)$

The marginal moments are:

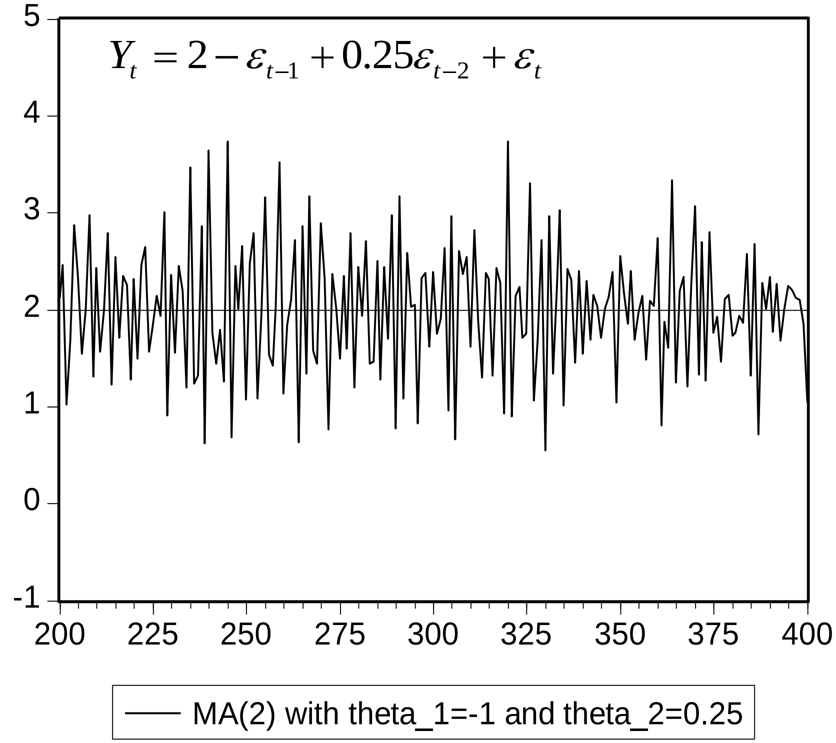
$$E(Y) = \mu$$
$$Var(Y) = \sigma_Y^2 = E(Y_t - \mu)^2 = \left(1 + \theta_1^2 + \theta_2^2\right)\sigma_\varepsilon^2$$

Forecasting with MA models

Q1. What does a time series of a MA(2) process look like?



(a)

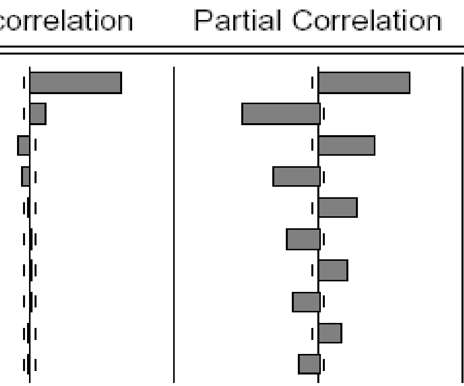


(b)

Forecasting with MA models

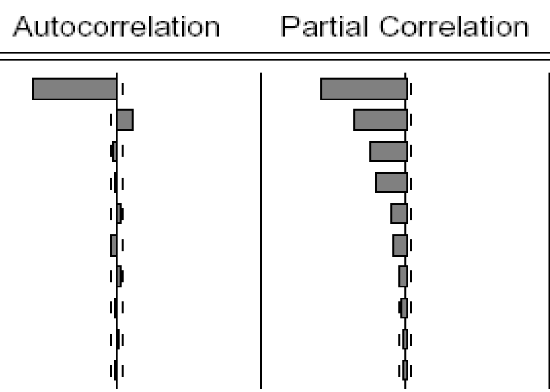
Q2. What do the autocorrelation functions look like?

Sample: 1 2000
Included observations: 1998



	AC	PAC
1	0.649	0.649
2	0.109	-0.542
3	-0.082	0.394
4	-0.054	-0.315
5	-0.007	0.271
6	0.016	-0.215
7	0.022	0.199
8	0.014	-0.184
9	-0.002	0.157
10	-0.010	-0.143

Sample: 1 2000
Included observations: 1998



	AC	PAC
1	-0.592	-0.592
2	0.119	-0.356
3	-0.026	-0.250
4	-0.001	-0.199
5	0.031	-0.106
6	-0.034	-0.088
7	0.028	-0.043
8	-0.014	-0.022
9	0.008	-0.001
10	-0.012	-0.008

Forecasting with MA models

Q2. What do the autocorrelation functions look like?

The most prominent features of the autocorrelation functions of a MA(2) are

✓ Only two spikes in the ACF are different from zero $\rho_1 and\ \rho_2 \neq 0$

$$\boldsymbol{\rho_1 = \frac{(\theta_1 + \theta_1\theta_2)}{(1 + \theta_1^2 + \theta_2^2)}}$$

$$\boldsymbol{\rho_2 = \frac{\theta_2}{(1 + \theta_1^2 + \theta_2^2)}\ and\ \rho_k = 0\ for\ k > 2}$$

THE MA(1) PROCESS IS COVARIANCE-STATIONARY

✓ The PACF decreases toward zero in an alternating fashion for time series (a) and a smooth fashion for time series (b)

Forecasting with MA models

Q3. What is the optimal forecast?

Forecasting horizon h=1

Optimal point forecast

$$f_{t,1} = E(Y_{t+1}|I_t) = E_t(Y_{t+1}) = \mu + \theta_1\varepsilon_t + \theta_2\varepsilon_{t-1}$$

One-perior ahead forecast error

$$e_{t,1} = Y_{t+1} - f_{t,1} = \varepsilon_{t+1}$$

Uncertainty associated with the forecast

$$\sigma^2_{t+1/t} = Var(Y_{t+1}|I_t) = E\left(Y_{t+1} - f_{t,1}\middle|I_t\right)^2 = E\left(e^2_{t,1}\right) = \sigma^2_\varepsilon$$

Density forecast (if ε_t is normally distributed)

$$f(Y_{t+1}|I_t) \to N\left(f_{t,1}, \sigma^2_{t+1/t}\right) = N(\mu + \theta_1\varepsilon_t + \theta_2\varepsilon_{t-1}, \sigma^2_\varepsilon)$$

and a 95% confidence interval is $f_{t,1} \pm 1.96\sigma_{t+1/t}$

Forecasting with MA models

Q3. What is the optimal forecast?

Forecasting horizon h=3

Optimal point forecast

$$f_{t,3} = E(Y_{t+3}|I_t) = E_t(Y_{t+3}) = \mu$$

One-perior ahead forecast error

$$e_{t,3} = Y_{t+3} \text{-} f_{t,3} = \theta_1 \varepsilon_{t+2} + \theta_2 \varepsilon_{t+1} + \varepsilon_{t+3}$$

Uncertainty associated with the forecast

$$\sigma^2_{t+3/t} = Var(Y_{t+3}|I_t) = E\left(e^2_{t,3}\right) = \sigma^2_\varepsilon\left(1 + \theta_1^2 + \theta_2^2\right)$$

Density forecast (if ε_t is normally distributed)

$$f(Y_{t+3}|I_t) \to N\left(f_{t,3}, \sigma^2_{t+3/t}\right)$$

and a 95% confidence interval is $f_{t,3} \pm 1.96\sigma_{t+3/t}$

Forecasting with MA models

A real example: ***forecasting the percentage changes in the 5-year Treasure Note Yield***

U.S. Treasure securities are considered the least risky assets in the U. S. economy and constitute an asset of reference to monitor de level of risk of other fixed-income securities

The available data goes from April 1953 to April 2008

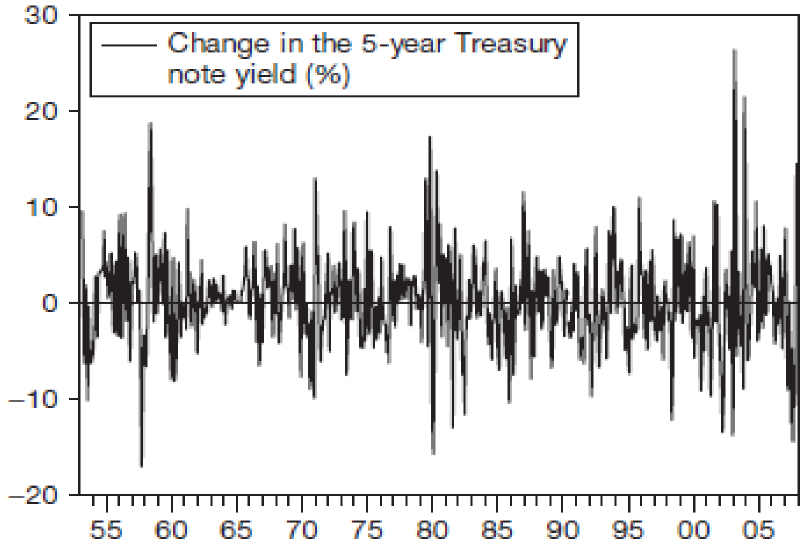
The profile of the time series is ragged and fluctuates around zero

The volatility of this series is very important to take into account. There are abrupt (positive and negative) changes in this series of approximaely 20%

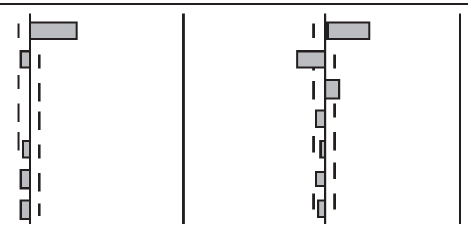
The ACF has only one positive spike different from zero with the remaining autocorrelations practically equal to zero

Forecasting with MA models

Percentage Changes in the 5-Year Treasury Note Yield



Sample: 1953M04 2008M04
Included observations: 660



Autocorrelation	Partial Correlation		AC	PAC
		1	0.339	0.339
		2	-0.073	-0.213
		3	0.007	0.129
		4	0.014	-0.063
		5	-0.043	-0.017
		6	-0.073	-0.060
		7	-0.069	-0.035

Forecasting with MA models

Estimation Output: 5-Year Treasury Yield (Monthly Percentage Changes)

Estimated model: $Y_t = 0,160 + 0,485\varepsilon_{t-1} + \varepsilon_t$

Dependent Variable: DY
Method: Least Squares
Sample (adjusted): 1953M05 2007M11
Included observations: 655 after adjustments
Convergence achieved after 7 iterations
Backcast: 1953M04

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.160159	0.258095	0.620544	0.5351
MA(1)	0.485011	0.034468	14.07130	0.0000

R-squared	0.165370	Mean dependent var	0.168613
Adjusted R-squared	0.164092	S.D. dependent var	4.866609
S.E. of regression	4.449443	Akaike info criterion	5.826484
Sum squared resid	12927.79	Schwarz criterion	5.840177
Log likelihood	-1906.173	F-statistic	129.3829
Durbin-Watson stat	2.055799	Prob(F-statistic)	0.000000

Inverted MA Roots -.49

Forecasting with MA models

December 2007-April 2008 Forecasts of 5-year Treasure Yield Changes

$h=1$ 12/2007	$f_{t,1}=\hat{\mu}+\hat{\theta}\varepsilon_t = 0.160+0.485\hat{\varepsilon}_t = -6.276\%$	$\sigma^2_{t+1\|t}=\hat{\sigma}^2_\varepsilon=4.449^2$	$f(Y_{t+1}\|I_t)\to N(\mu_{t+1\|t},\sigma^2_{t+1\|t}) = N(-6.276, 4.449^2)$
$h=2$ 1/2008	$f_{t,2}=\hat{\mu}=0.160\%$	$\sigma^2_{t+2\|t}=\hat{\sigma}^2_\varepsilon(1+\hat{\theta}^2) = 4.449^2(1+0.485^2) = 23.683=\hat{\sigma}^2_Y$	$f(Y_{t+2}\|I_t)\to N(0.16, 23.683)$
$h=3$ 2/2008	$f_{t,3}=\hat{\mu}=0.160\%$	$\sigma^2_{t+3\|t}=23.683=\hat{\sigma}^2_Y$	$f(Y_{t+3}\|I_t)\to N(0.16, 23.683)$
$h=4$ 3/2008	$f_{t,4}=\hat{\mu}=0.160\%$	$\sigma^2_{t+4\|t}=23.683=\hat{\sigma}^2_Y$	$f(Y_{t+4}\|I_t)\to N(0.16, 23.683)$
$h=5$ 4/2008	$f_{t,5}=\hat{\mu}=0.160\%$	$\sigma^2_{t+5\|t}=23.683=\hat{\sigma}^2_Y$	$f(Y_{t+5}\|I_t)\to N(0.16, 23.683)$

Forecasting with MA models

Multistep Forecast of Monthly Changes of 5-year Treasury Yield