Autoregressive Integrated Moving Average (ARIMA) Models

Example

MA(1)

$$Y_{1t} = 0.1 + u_{1t} + 0.5\, u_{1t-1}$$

AR(1)

$$Y_{2t} = 0.1 + 0.5\, Y_{2t-1} + u_{2t}$$

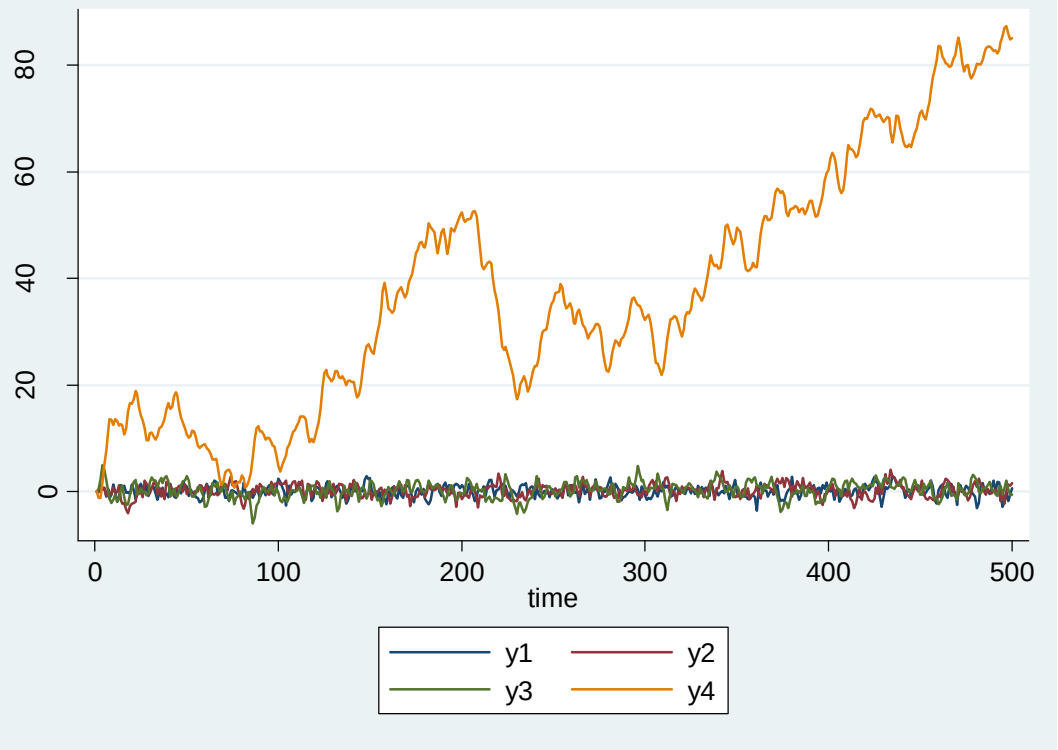
ARMA(1,1)

$$Y_{3t} = 0.1 + 0.5\, Y_{3t-1} + u_{3t} + 0.5\, u_{3t-1}$$

ARIMA(1,1,1)

$$\Delta Y_{4t} = 0.1 + 0.5\, \Delta Y_{4t-1} + u_{4t} + 0.5\, u_{4t-1}$$

```
. set obs 500
obs was 0, now 500

. gener time=_n

. tsset time
        time variable:  time, 1 to 500

. gener y1=0 in 1
(499 missing values generated)

. gener y2=0 in 1
(499 missing values generated)

. gener y3=0 in 1
(499 missing values generated)

. gener y4=0 in 1
(499 missing values generated)

. gener dy4 = 0 in 1
(499 missing values generated)

. gener u1= rnormal(0,1)

. gener u2= rnormal(0,1)

. gener u3= rnormal(0,1)

. gener u4= rnormal(0,1)

. replace y1=0.1+u1+0.5*l.u1 if time>1
(499 real changes made)

. replace y2=0.1+0.5*l.y2+u2 if time>1
(499 real changes made)

. replace y3=0.1+0.5*l.y3+u3+0.5*l.u3 if time>1
(499 real changes made)

. replace y4=0.1+l.y4+u4+0.5*l.u4 if time>1
(499 real changes made)

. replace dy4=d.y4 if time>1
(499 real changes made)

. replace y4=0.1+l.y4+0.5*l.dy4+u4+0.5*l.u4 if time>1
(499 real changes made)

. line y1 y2 y3 y4 time
```



Example: Finding ARIMA process of Price Index (CMI) ─ Box-Jenkins Methodology

Identification and Estimation: The first step is to identify the order of ARIMA(p,d,q). To find the value of p, d, and q, we should firstly identify value of d by performing unit root test to find which order the series is integrated.

Unit Root Test

Find integrated order of the series using unit root test.

```
. tsset time
        time variable:  time, 1997m1 to 2005m6

. dfuller cmi

Dickey-Fuller test for unit root                   Number of obs   =       101

                              ---------- Interpolated Dickey-Fuller ---------
                 Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value            Value
------------------------------------------------------------------------------
 Z(t)             -1.391            -3.510            -2.890           -2.580
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.5868

. dfuller d.cmi

Dickey-Fuller test for unit root                   Number of obs   =       100

                              ---------- Interpolated Dickey-Fuller ---------
                 Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value            Value
------------------------------------------------------------------------------
 Z(t)             -5.967            -3.510            -2.890           -2.580
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000
```

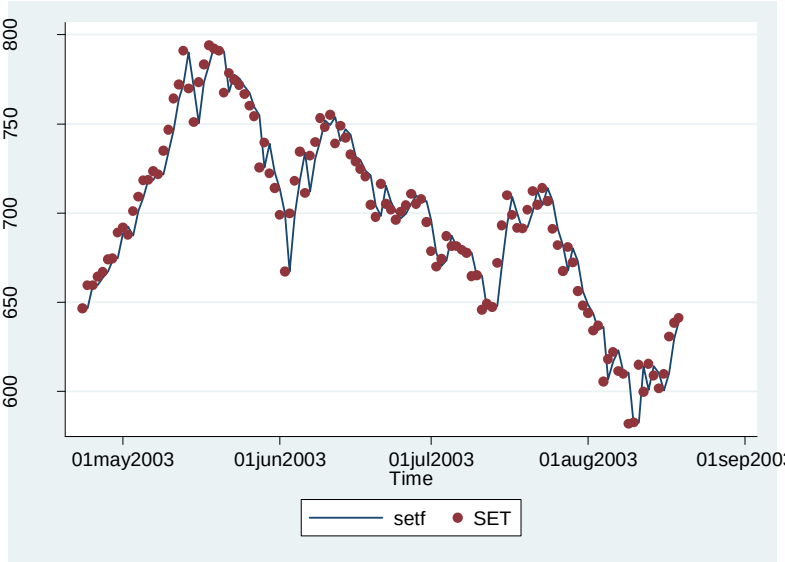
```
. arima cmi, arima(1,1,1) nolog

ARIMA regression

Sample:  1997m2 to 2005m6                       Number of obs      =       101
                                                Wald chi2(2)       =     43.07
Log likelihood = -169.6031                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |                 OPG
       D.cmi |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
cmi          |
       _cons |   .3614335   .2202549     1.64   0.101    -.0702582    .7931251
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |   .1668652   .2874552     0.58   0.562    -.3965367    .7302672
          ma |
         L1. |   .3087396    .238521     1.29   0.196     -.158753    .7762322
-------------+----------------------------------------------------------------
      /sigma |   1.295842   .0659596    19.65   0.000     1.166564    1.425121
------------------------------------------------------------------------------

. estat ic

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC          BIC
-------------+---------------------------------------------------------------
           . |    101           .   -169.6031      4     347.2061    357.6666
-----------------------------------------------------------------------------


. do "E:\TU\MIF\MF651\Manual\Documents\STATA\2009\ARIMA.do"

. *Define Variable y
.         g y = set
(1 missing value generated)

. *Specify order p d q
. forvalue d = 1(1)1 {
  2. forvalue p = 1(1)2 {
  3. forvalue q = 1(1)2 {
  4.         display "estimate arima`p'`d'`q'"
  5.         capture: quietly arima y, arima(`p',`d',`q') nolog
  6.         if _rc~=0 {
  7.         display "flatlog when pdq =" `p'`d'`q'
  8.         continue
  9.         }
 10.         estimates store arima`p'`d'`q'
 11.         display "arima`p'`d'`q' already estimated"
 12. }
 13. }
 14. estimates table arima1`d'1 arima1`d'2, star(0.1 0.05 0.01) stat(aic bic ll)
 15. estimates table arima2`d'1 arima2`d'2, star(0.1 0.05 0.01) stat(aic bic ll)
 16. }
estimate arima111
arima111 already estimated
estimate arima112
arima112 already estimated
estimate arima211
arima211 already estimated
estimate arima212
arima212 already estimated


-----------------------------------------------
    Variable |   arima111        arima112
-------------+---------------------------------
y            |
       _cons | -.05254828      -.04186597
-------------+---------------------------------
ARMA         |
        L.ar | -.91746548***   -.44764234
        L.ma |  .89391226**     .45501476
       L2.ma |                   .0894573
```

```
-------------+----------------------------------
sigma        |
       _cons |  12.970114***     12.94033***
-------------+----------------------------------
Statistics   |
         aic |  947.67414       949.13046
         bic |  958.75688       962.98388
          ll | -469.83707      -469.56523
-------------------------------------------------
           legend: * p<.1; ** p<.05; *** p<.01

-------------------------------------------------
    Variable |   arima211        arima212
-------------+----------------------------------
y            |
       _cons |  -.0417522       .02663207
-------------+----------------------------------
ARMA         |
        L.ar | -.45694176       .08426726
       L2.ar |  .07334288       .57125955
        L.ma |  .45785522      -.09578595
       L2.ma |                 -.50183303
-------------+----------------------------------
sigma        |
       _cons |  12.946917***    12.947449***
-------------+----------------------------------
Statistics   |
         aic |  949.24708       951.25664
         bic |  963.10051       967.88075
          ll | -469.62354      -469.62832
-------------------------------------------------
           legend: * p<.1; ** p<.05; *** p<.01

. drop y

.
.
end of do-file
```

Forecast

Command: predict, xb

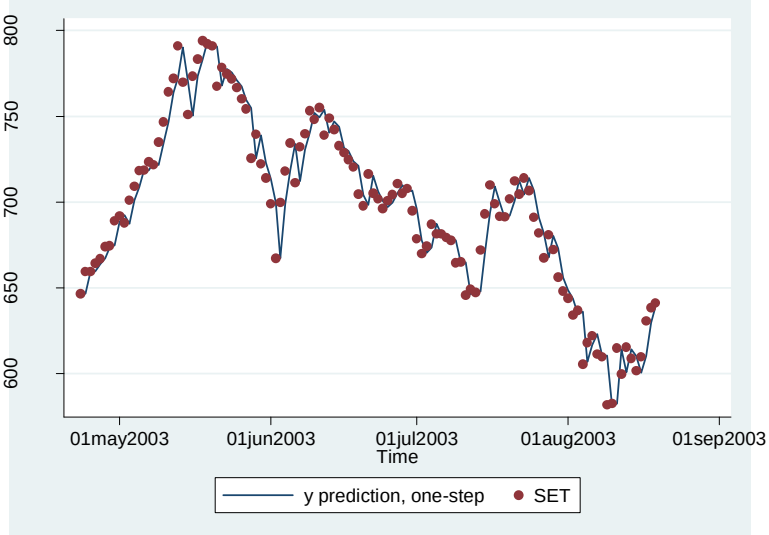
```
. arima set, arima(1,1,1) nolog

ARIMA regression

Sample:  24apr2003 - 19aug2003                  Number of obs      =       118
                                                Wald chi2(2)       =     10.97
Log likelihood = -469.8371                      Prob > chi2        =    0.0042

------------------------------------------------------------------------------
             |                 OPG
       D.set |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
set          |
       _cons |  -.0525483   1.212681    -0.04   0.965    -2.429359    2.324263
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |  -.9174655   .3483242    -2.63   0.008    -1.600168   -.2347626
          ma |
         L1. |   .8939123   .3766723     2.37   0.018     .1556482    1.632176
-------------+----------------------------------------------------------------
      /sigma |   12.97011   .8838035    14.68   0.000     11.23789    14.70234
------------------------------------------------------------------------------

. predict dsetf, xb

. g setf=l.set+dsetf
(2 missing values generated)

. twoway (line setf time, sort) (scatter set time, sort)
```



Command: predict, y

Static Forecast (One step ahead forecast)

```
. predict setff, y dynamic(.) t0(120)
(2 missing values generated)

. twoway (line setff time, sort) (scatter set time, sort)
```



Forecasting Error

```
. g e1=setf-set
(2 missing values generated)
```

SARIMA

```
. arima set, arima(1,1,1) sarima(1,1,1,5) nolog

ARIMA regression

Sample:  29apr2003 - 19aug2003                  Number of obs      =       113
                                                Wald chi2(4)       =     42.63
Log likelihood = -455.9555                      Prob > chi2        =    0.0000
------------------------------------------------------------------------------
             |                 OPG
      DS5.set |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
set            |
       _cons |  -.2077929   .2532667    -0.82   0.412    -.7041866    .2886007
-------------+----------------------------------------------------------------
ARMA           |
          ar |
         L1. |  -.9189895   .1894008    -4.85   0.000    -1.290208   -.5477708
          ma |
         L1. |   .8645402   .2277527     3.80   0.000      .418153    1.310927
-------------+----------------------------------------------------------------
ARMA5          |
          ar |
         L1. |   .2141882   .1462771     1.46   0.143    -.0725097     .500886
          ma |
         L1. |  -1.000018   734.7429    -0.00   0.999     -1441.07     1439.07
-------------+----------------------------------------------------------------
      /sigma |   12.86195   4725.669     0.00   0.998    -9249.279    9275.003
------------------------------------------------------------------------------
```

ARMAX

```
. arima set ibr, arima(1,1,1) nolog

ARIMA regression

Sample:  24apr2003 - 19aug2003                  Number of obs      =       118
                                                Wald chi2(3)       =     10.70
Log likelihood = -469.7221                      Prob > chi2        =    0.0134
------------------------------------------------------------------------------
             |                 OPG
       D.set |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
set          |
         ibr |
         D1. |  -20.27917   41.04736    -0.49   0.621    -100.7305    60.17217
       _cons |  -.0544715   1.210328    -0.05   0.964    -2.426671    2.317728
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |  -.9136402   .3653158    -2.50   0.012    -1.629646   -.1976343
          ma |
         L1. |   .8893427   .3968214     2.24   0.025      .111587    1.667098
-------------+----------------------------------------------------------------
      /sigma |   12.95795   .8950096    14.48   0.000     11.20377    14.71214
------------------------------------------------------------------------------
```