6.10 Example of SARIMA using Consumption Data

Construct SARIMA model using monthly and seasonally unadjusted consumption expenditure data and STATA12.

Estimation Period: Jan., 1970 — Dec., 2012 ($T = 516$)

```
. gen time=_n                                             → Generate time.

. tsset time                                              → Defined as time series data.
        time variable:  time, 1 to 516
                delta:  1 unit

. corrgram expend                                    → Variable name: expend
                                                     → corrgram: Compute autocorrelation
                                                                 and partial autocorrelation.

                                                  -1       0       1 -1       0       1
 LAG       AC       PAC      Q     Prob>Q  [Autocorrelation]  [Partial Autocor]
-------------------------------------------------------------------------------
1        0.8488   0.8499   373.88  0.0000          |------            |------
2        0.8231   0.3858   726.18  0.0000          |------            |---
3        0.8716   0.5266     1122  0.0000          |------            |----
4        0.8706   0.4025   1517.6  0.0000          |------            |---
5        0.8498   0.3447   1895.3  0.0000          |------            |--
6        0.8085   0.0074   2237.9  0.0000          |------            |
7        0.8378   0.1528   2606.5  0.0000          |------            |-
8        0.8460   0.1467     2983  0.0000          |------            |-
9        0.8342   0.3006   3349.9  0.0000          |------            |--
```

```
10        0.7735  -0.1518     3666  0.0000          |------         -|
11        0.7852  -0.1185   3992.3  0.0000          |------          |
12        0.9234   0.9442   4444.5  0.0000          |-------         |-------
13        0.7754  -0.5486   4764.1  0.0000          |------      ----|
14        0.7482  -0.3248   5062.1  0.0000          |-----         --|
15        0.7963  -0.2392   5400.5  0.0000          |------         -|
```

→ Autocorrelation does not approach zero for large lag.

→ Time series has unit root.

```
. gen dexp=expend-l.expend
(1 missing value generated)
```

→ Generate dexp=expend-expend(-1), excluding unit root.

```
. corrgram dexp

                                          -1       0       1 -1       0       1
 LAG       AC       PAC      Q     Prob>Q  [Autocorrelation]  [Partial Autocor]
-------------------------------------------------------------------------------
1        -0.4316  -0.4329   96.485  0.0000       ---|               ---|
2        -0.2546  -0.5441   130.13  0.0000        --|              ----|
3         0.1721  -0.4091   145.53  0.0000          |-              ---|
4         0.0667  -0.3459   147.85  0.0000          |                --|
5         0.0715  -0.0036   150.52  0.0000          |                  |
6        -0.2428  -0.1489   181.36  0.0000         -|                 -|
7         0.0711  -0.1400   184.01  0.0000          |                 -|
8         0.0668  -0.2900   186.36  0.0000          |                --|
9         0.1704   0.1681   201.64  0.0000          |-                 |-
10       -0.2485   0.1306   234.21  0.0000         -|                  |-
11       -0.4293  -0.9305   331.56  0.0000       ---|           -------|
12        0.9773   0.6768   837.12  0.0000          |-------           |-----
13       -0.4152   0.3778   928.56  0.0000       ---|                  |---
14       -0.2583   0.2688   964.03  0.0000        --|                  |--
```

```
15       0.1712   0.0406   979.63  0.0000          |-                 |
```

Big autocorrelation at lag 12.

```
. gen sdex=dexp-l12.dexp          → Generate sdex=dexp-dexp(-12),
(13 missing values generated)       excluding seasonality.

. corrgram sdex

                                              -1       0       1 -1       0       1
 LAG       AC       PAC      Q     Prob>Q  [Autocorrelation]  [Partial Autocor]
-------------------------------------------------------------------------------
1       -0.4752  -0.4753   114.28  0.0000          ---|                ---|
2       -0.0244  -0.3235   114.58  0.0000             |                 --|
3        0.1163  -0.0759   121.46  0.0000             |                   |
4       -0.1246  -0.1365   129.37  0.0000             |                  -|
5        0.0341  -0.1016   129.96  0.0000             |                   |
6       -0.0151  -0.1136   130.08  0.0000             |                   |
7       -0.0395  -0.1413   130.88  0.0000             |                  -|
8        0.1123   0.0092   137.35  0.0000             |                   |
9       -0.0664  -0.0100   139.62  0.0000             |                   |
10       0.0168   0.0069   139.76  0.0000             |                   |
11       0.1642   0.2422   153.68  0.0000             |-                  |-
12      -0.3888  -0.2469    231.9  0.0000          ---|                  -|
13       0.2242  -0.1205   257.96  0.0000             |-                  |
14      -0.0147  -0.0941   258.07  0.0000             |                   |
15      -0.0708  -0.0591   260.68  0.0000             |                   |
```

Big autocorrelation at lag 1 (ignore lag 12).
→ MA(1)
Big partial autocorrelation at lags 1 and 2.
→ AR(2)

```
. arima sdex, ar(1,2) ma(1)          → Model specification is:
                                       sdex∼ARMA(2,1), i.e., expend∼SARIMA(2,1,1).

(setting optimization to BHHH)
Iteration 0:   log likelihood = -5107.4608
Iteration 1:   log likelihood =  -5102.391
Iteration 2:   log likelihood = -5099.9071
Iteration 3:   log likelihood = -5099.4216
Iteration 4:   log likelihood = -5099.2463
(switching optimization to BFGS)
Iteration 5:   log likelihood = -5099.2361
Iteration 6:   log likelihood = -5099.2346
Iteration 7:   log likelihood = -5099.2346
Iteration 8:   log likelihood = -5099.2346

ARIMA regression

Sample:  14 - 516                               Number of obs      =       503
                                                Wald chi2(3)       =    973.93
Log likelihood = -5099.235                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |                 OPG
        sdex |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sdex         |
       _cons |  -15.64573   59.17574    -0.26   0.791     -131.628    100.3366
-------------+----------------------------------------------------------------
```

```
ARMA         |
          ar |
         L1. |   .1271774   .0581883     2.19   0.029     .0131304    .2412244
         L2. |   .1009983    .053626     1.88   0.060    -.0041068    .2061034
             |
          ma |
         L1. |  -.8343264   .0419364   -19.90   0.000    -.9165202   -.7521326
-------------+----------------------------------------------------------------
      /sigma |   6111.128   139.0105    43.96   0.000     5838.673    6383.584
------------------------------------------------------------------------------
Note: The test of the variance against zero is one sided, and the two-sided
    confidence interval is truncated at zero.


. estat ic

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    503           .   -5099.235      5     10208.47    10229.57
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note

. predict resid, r
(13 missing values generated)
```

```
. corrgram resid                   → Make sure sdex∼ARMA(2,1), using the residuals.

                                                  -1       0       1 -1       0       1
 LAG       AC       PAC      Q     Prob>Q  [Autocorrelation]  [Partial Autocor]
-------------------------------------------------------------------------------
1        -0.0132  -0.0132  .08814  0.7666          |                    |
2        -0.0095  -0.0097   .1341  0.9351          |                    |
3         0.1248   0.1246  8.0433  0.0451          |                    |
4        -0.0644  -0.0624  10.154  0.0379          |                    |
5        -0.0001   0.0011  10.154  0.0710          |                    |
6        -0.0138  -0.0309  10.252  0.1144          |                    |
7        -0.0032   0.0126  10.257  0.1745          |                    |
8         0.0958   0.0938   14.97  0.0597          |                    |
9        -0.0317  -0.0255  15.487  0.0784          |                    |
10        0.0126   0.0112  15.569  0.1127          |                    |
11       -0.0053  -0.0305  15.583  0.1573          |                    |
12       -0.3773  -0.3837  89.235  0.0000       ---|                 ---|
13        0.0408   0.0258  90.098  0.0000          |                    |
14       -0.0233  -0.0307  90.381  0.0000          |                    |
15       -0.0911  -0.0059  94.703  0.0000          |                    |
```

Big autocorrelation and
big partial correlation at lag 12.
→ We need to re-specify the model.

6.11 ARCH and GARCH Models

Autoregressive Conditional Heteroskedasticity (ARCH)

Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

1. ARCH (p) Model

$$\epsilon_t|\epsilon_{t-1}, \epsilon_{t-2}, \cdots, \epsilon_1 \sim N(0, h_t),$$

where,

$$h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \cdots + \alpha_p\epsilon_{t-p}^2.$$

The unconditional variance of ϵ_t is:

$$\sigma_\epsilon^2 = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \cdots - \alpha_p}$$

2. GARCH (p, q) Model

$$\epsilon_t|\epsilon_{t-1}, \epsilon_{t-2}, \cdots, \epsilon_1 \sim N(0, h_t),$$

where

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_p \epsilon_{t-p}^2 + \beta_1 h_{t-1} + \cdots + \beta_q h_{t-q}.$$

3. Application to OLS (Case of ARCH(1) Model):

$$y_t = x_t\beta + \epsilon_t, \qquad \epsilon_t|\epsilon_{t-1}, \epsilon_{t-2}, \cdots, \epsilon_1 \sim N(0, \alpha_0 + \alpha_1 \epsilon_{t-1}^2).$$

The joint density of $\epsilon_1, \epsilon_2, \cdots, \epsilon_T$ is:

$$\begin{aligned} f(\epsilon_1, \cdots, \epsilon_T) &= f(\epsilon_1) \prod_{t=2}^{T} f(\epsilon_t|\epsilon_{t-1}, \cdots, \epsilon_1) \\ &= (2\pi)^{-1/2} \left(\frac{\alpha_0}{1-\alpha_1} \right)^{-1/2} \exp\left(-\frac{1}{2\alpha_0/(1-\alpha_1)} \epsilon_1^2 \right) \\ &\times (2\pi)^{-(T-1)/2} \prod_{t=2}^{T} (\alpha_0 + \alpha_1 \epsilon_{t-1}^2)^{-1/2} \exp\left(-\frac{1}{2} \sum_{t=2}^{T} \frac{\epsilon_t^2}{\alpha_0 + \alpha_1 \epsilon_{t-1}^2} \right). \end{aligned}$$

The log-likelihood function is:

$$
\begin{aligned}
&\log L(\beta, \alpha_0, \alpha_1; y_1, \cdots, y_T) \\
&\quad = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\left(\frac{\alpha_0}{1-\alpha_1}\right) - \frac{1}{2\alpha_0/(1-\alpha_1)}(y_1 - x_1\beta)^2 \\
&\qquad -\frac{T-1}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{T}\log\left(\alpha_0 + \alpha_1(y_{t-1} - x_{t-1}\beta)^2\right) \\
&\qquad -\frac{1}{2}\sum_{t=2}^{T}\frac{(y_t - x_t\beta)^2}{\alpha_0 + \alpha_1(y_{t-1} - x_{t-1}\beta)^2}.
\end{aligned}
$$

Obtain α_0, α_1 and β such that the log-likelihood function is maximized.

$\alpha_0 > 0$ and $\alpha_1 > 0$ have to be satisfied.

These two conditions are explicitly included, when the model is modified to: $\mathrm{E}(\epsilon_t^2|\epsilon_{t-1}, \epsilon_{t-2}, \cdots$
$\alpha_0^2 + \alpha_1^2\epsilon_{t-1}^2$.

Testing the ARCH(1) Effect:

(a) Estimate $y_t = x_t\beta + u_t$ by OLS, and compute $\hat{\beta}$ and $\hat{u}_t = y_t - x_t\hat{\beta}$.

(b) Estimate $\hat{u}_t^2 = \alpha_0 + \alpha_1\hat{u}_{t-1}^2$ by OLS. If $\hat{\alpha}_1$ is significant, there is the ARCH(1) effect in the error term.

This test corresponds to LM test.

Example: GARCH(1,1) Model

```
. arch sdex l.sdex l2.sdex, arch(1) garch(1)        → 誤差項の MA 項を GRARCH に

(setting optimization to BHHH)
Iteration 0:   log likelihood = -5089.3558
Iteration 1:   log likelihood = -5086.7468
.........
.........

Iteration 22:  log likelihood = -5064.9328  (backed up)
Iteration 23:  log likelihood = -5064.9328

ARCH family regression
```

```
Sample: 16 - 516                                Number of obs   =        501
Distribution: Gaussian                          Wald chi2(2)    =     225.19
Log likelihood = -5064.933                      Prob > chi2     =     0.0000

------------------------------------------------------------------------------
             |                 OPG
        sdex |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
sdex         |
        sdex |
         L1. |  -.6357273   .0426939   -14.89   0.000    -.7194059   -.5520488
         L2. |   -.370862   .0466222    -7.95   0.000    -.4622398   -.2794842
             |
       _cons |  -55.28043   261.2057    -0.21   0.832    -567.2341    456.6733
-------------+----------------------------------------------------------------
ARCH         |
        arch |
         L1. |    .041632   .0123474     3.37   0.001     .0174317    .0658324
             |
       garch |
         L1. |   .9526041   .0148639    64.09   0.000     .9234715    .9817367
             |
       _cons |   312143.8   227564.3     1.37   0.170    -133873.9    758161.6
------------------------------------------------------------------------------
```

7 Vector Autoregressive (VAR) Model – Causality, Impulse Response Function and etc

We can consider VARMA (vector autoregressive moving average) model.

However, it is very difficult to estimate MA terms in the case of vector.

Usually, we consider VAR (Vector Autoregressive) process:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

where

$$y_t : k \times 1, \qquad \mu : k \times 1, \qquad \epsilon_t : k \times 1, \qquad \phi_i : k \times k.$$

Rewriting the above equation,

$$\phi(L) y_t = \mu + \epsilon_t,$$

where $\phi(L) = I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$.

VAR(1) Model:

$$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \text{i.e.,} \qquad (I_k - \phi_1 L) y_t = \epsilon_t.$$

When y_t is stationary, we obtain:

$$\begin{aligned} y_t &= (I_k - \phi_1 L)^{-1} \epsilon_t \\ &= (I_k + \phi_1 L + \phi_1^2 L^2 + \phi_1^3 L^3 + \cdots) \epsilon_t \\ &= \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_1^2 \epsilon_{t-2} + \phi_1^3 \epsilon_{t-3} + \cdots \end{aligned}$$

VAR(1)=VMA(∞)

VAR(2) Model:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t, \qquad \text{i.e.,} \qquad (I_k - \phi_1 L - \phi_2 L^2) y_{t-1} = \epsilon_t.$$

When y_t is stationary, we obtain:

$$\begin{aligned} y_{t-1} &= (I_k - \phi_1 L - \phi_2 L^2)^{-1} \epsilon_t \\ &= \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots \end{aligned}$$

VAR(2)=VMA(∞)

VAR(p) Model:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

i.e.,

$$(I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) y_t = \epsilon_t.$$

When y_t is stationary, we obtain:

$$y_t = (I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)^{-1} \epsilon_t$$

$$= \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots$$

VAR(p)=VMA(∞)

7.1 Autocovariance Matrix and Autocorrelation Matrix

Let y_t be a $k \times 1$ vector.

Autocovariance Function Matrix:

$$\Gamma(\tau) = \mathrm{E}((y_t - \mu)(y_{t-\tau} - \mu)'), \qquad \tau = 0, 1, 2, \cdots,$$

where $\mathrm{E}(y_t) = \mu$. $\Gamma(\tau)$ is a $k \times k$ matrix.

$$\Gamma(\tau) = \Gamma(-\tau)'$$

Autocorrelation Function Matrix:

$$\rho(\tau) = D^{-1/2}\Gamma(\tau)D^{-1/2},$$

where the (i, j)th element of D is given by $\gamma_{ii}(0) = \mathrm{V}(y_{it})$ for $i = j$ and zero otherwise.

$$\rho(\tau) = \rho(-\tau)'$$