Written Examination in Time Series Analysis

2017-05-24

Solutions

Task 1

A1.

A stochastic process is a sequence of random variables.

A2.

The stochastic process is never observed, the realization is.

B.

In order to be covariance stationary, a stochastic process needs to fulfill the following:

1. $E(Y_t) = \mu$, $\forall t$. The mean is constant over time.
2. $Var(Y_t) = \gamma_0 < \infty$, $\forall t$. The variance is finite and constant over time.
3. $Cov(Y_t, Y_{t+k}) = Cov(Y_0, Y_k)$, $\forall t, k$. The covariance is independent of t and is only a function of the lag length, k.

The third condition actually implies the second.

C.

The model

$$Y_t = \alpha_0 + Y_{t-1} + e_t$$

where $e_t \sim NID(0, \sigma^2)$.

is difference-stationary i.e. the first difference is a covariance stationary process.

D.

The model

$$Y_t = \alpha_0 + \alpha_1 t + e_t$$

where $e_t \sim NID(0, \sigma^2)$.

is trend stationary. By removing the trend we get $Y_t - \alpha_1 t = e_t$, which is a stationary process.

E.

1. Identication: Identify the DGP, or find a DGP that is suitable for modelling the data, this can be done by looking at correlograms.
2. Estimation: Estimate the parameters of the model, this can be done by OLS or maximum likelihood estimation.
3. Diagnostics: Investigate to what extent the model captures the systematic variation in the data. The residuals should be white noise, i.e. uncorrelated. This can be done by looking at correlograms and plots and performing statistical tests on the estimated parameters and residuals.
4. Forecasting: Use the model to forecast the conditional mean of the process. Forecasts can be done based on conditional means.

F.

If we have two unit root processes that are integrated of order one, and a linear combination of them is stationary (integrated of order zero), then the two are said to be co-integrated.

G.

The ARMA models the conditional mean, whilst the GARCH models the conditional variance.

Task 2

The process of interest is

$$Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1} \tag{1}$$

where $e_t \sim NID(0, \sigma^2)$.

A.

It is an ARIMA(1,0,1).

B.

The ARMA process is stationary if the AR part is stationary. That is, whenever $|\phi| < 1$ and for all values of θ.

C.

The ARMA process is invertible if the MA part is invertible. That is, whenever $|\theta| < 1$ and for all values of ϕ.

D.

See Figure 1.

We have that

$$(1 - B\phi)Y_t = (1 - B\theta)e_t \tag{2}$$

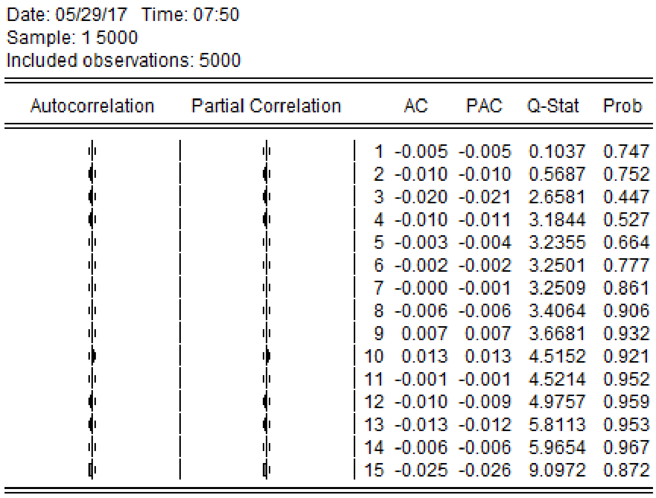
where $e_t \sim NID(0, \sigma^2)$. That is, if $\theta = \phi$

$$\begin{aligned} Y_t &= \frac{(1 - B\theta)}{(1 - B\phi)} e_t \\ &= e_t \end{aligned} \tag{3}$$

E.

The characteristic equation for the MA part in (1) is $1 - \theta x = 0$. Solving for x we get

Figure 1: SACF and SPACF of an ARMA where $\theta = 0.9$ and $\phi = 0.9$



Date: 05/29/17 Time: 07:50
Sample: 1 5000
Included observations: 5000

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.005	-0.005	0.1037	0.747
		2	-0.010	-0.010	0.5687	0.752
		3	-0.020	-0.021	2.6581	0.447
		4	-0.010	-0.011	3.1844	0.527
		5	-0.003	-0.004	3.2355	0.664
		6	-0.002	-0.002	3.2501	0.777
		7	-0.000	-0.001	3.2509	0.861
		8	-0.006	-0.006	3.4064	0.906
		9	0.007	0.007	3.6681	0.932
		10	0.013	0.013	4.5152	0.921
		11	-0.001	-0.001	4.5214	0.952
		12	-0.010	-0.009	4.9757	0.959
		13	-0.013	-0.012	5.8113	0.953
		14	-0.006	-0.006	5.9654	0.967
		15	-0.025	-0.026	9.0972	0.872

$$x = \frac{1}{\theta} \tag{4}$$

The process is invertible whenever $|x| > 1$

F.

We can rewrite (1) as

$$\begin{aligned} Y_t &= e_t - \theta e_{t-1} \\ Y_t &= (1 - B\theta) e_t \end{aligned} \tag{5}$$

We make use of the assumption that $|\theta| < 1$, which is equivalent of $|B\theta| < 1$. Under this assumption, we use geometric series and the fact that

$$\frac{1}{(1 - B\theta)} = \sum_{j=0}^{\infty} (B\theta)^j \tag{6}$$

so (3) can be rewritten as

$$\sum_{j=0}^{\infty} (B\theta)^j Y_t = e_t \tag{7}$$

and Y_t is then

$$Y_t = -\sum_{j=1}^{\infty} \theta^j Y_{t-j} + e_t \tag{8}$$

which is a $AR(\infty)$ representation with $\phi = -\theta^j$. If then $|\theta| < 1$, then $|\phi| < 1$ and the process is stationary.

Task 3

The process of interest is

$$Y_t = e_t - \theta_1 e_{t-1} \tag{9}$$

where $e_t \sim NID(0, \sigma^2)$. That is, a MA(1) process.

A.

The expected value of the process is

$$\begin{aligned} E(Y_t) &= E(e_t) - \theta_1 E(e_{t-1}) \\ &= 0 \end{aligned} \tag{10}$$

B.

The variance, γ_0, of the process is

$$\begin{aligned} \gamma_0 &= Var(e_t - \theta_1 e_{t-1}) \\ &= Var(e_t) + \theta_1^2 Var(e_{t-1}) - 2\theta_1 Cov(e_t, e_{t-1}) \\ &\text{since the error terms are assumed to be independent, the cov. term is zero and} \\ \gamma_0 &= \sigma^2 + \theta_1^2 \sigma^2 \\ &= \sigma^2(1 + \theta_1^2) \end{aligned} \tag{11}$$

C.

The first autocovariance is

$$\begin{aligned} \gamma_1 &= Cov(Y_t, Y_{t-1}) \\ &= Cov(e_t - \theta_1 e_{t-1}, e_{t-1} - \theta_1 e_{t-2}) \\ &= Cov(e_t, e_{t-1}) - \theta_1 Cov(e_t, e_{t-2}) - \theta_1 Var(e_{t-1}) + \theta_1^2 Cov(e_{t-1}, e_{t-2}) \\ &\text{since the error terms are assumed to be independent, the cov. terms are all zero and} \\ \gamma_1 &= -\theta_1 \sigma^2 \end{aligned} \tag{12}$$

The second autocovariance is

$$
\begin{aligned}
\gamma_2 &= Cov(Y_t, Y_{t-2}) \\
&= Cov(e_t - \theta_1 e_{t-1}, e_{t-2} - \theta_1 e_{t-3}) \\
&= Cov(e_t, e_{t-2}) - \theta_1 Cov(e_t, e_{t-3}) - \theta_1 Cov(e_{t-1}, e_{t-2}) + \theta_1^2 Cov(e_{t-1}, e_{t-3})
\end{aligned}
$$

since the error terms are assumed to be independent, the cov. terms are all zero and

$$\gamma_2 = 0 \tag{13}$$

D.

We are to calculate the conditional expected value of the process at time $t+1$, given all the information up to and including time t. That is,

$$
\begin{aligned}
\hat{Y}_{t+1} &= E[Y_{t+1}|I_t] \\
&= E[e_{t+1}|e_t, e_{t-1}, ...] - \theta_1 E[e_t|e_t, e_{t-1}, ...] \\
&= -\theta_1 e_t
\end{aligned} \tag{14}
$$

since we can use the fact that $E[e_{t+1}|e_t, e_{t-1}, ...] = 0$ and that the expected value of a constant is the constant itself.

E.

The expected value of the forecast error is

$$
\begin{aligned}
E[Y_{t+1} - \hat{Y}_{t+1}] &= E[e_{t+1} - \theta_1 e_t - (-\theta_1 e_t)] \\
&= E[e_{t+1}] \\
&= 0
\end{aligned} \tag{15}
$$

F.

The forecast error variance is

$$
\begin{aligned}
Var[Y_{t+1} - \hat{Y}_{t+1}] &= Var[e_{t+1} - \theta_1 e_t - (-\theta_1 e_t)] \\
&= Var[e_{t+1}] \\
&= \sigma^2
\end{aligned} \tag{16}
$$

G.

See Figure 2.

Figure 2: SACF and SPACF of a $SMA(0) \times (1)_4$ with $\Theta = 0.9$

Date: 05/29/17 Time: 08:00
Sample: 1 1000
Included observations: 998

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.027	0.027	0.7037	0.402
		2	0.004	0.004	0.7222	0.697
		3	0.012	0.012	0.8748	0.832
		4	-0.513	-0.514	264.90	0.000
		5	-0.053	-0.031	267.70	0.000
		6	0.007	0.019	267.76	0.000
		7	0.003	0.030	267.76	0.000
		8	0.021	-0.328	268.23	0.000
		9	0.044	0.010	270.23	0.000
		10	-0.015	-0.000	270.45	0.000
		11	-0.033	-0.014	271.59	0.000
		12	-0.014	-0.253	271.79	0.000
		13	-0.039	-0.024	273.31	0.000
		14	0.023	0.036	273.84	0.000
		15	0.015	-0.018	274.06	0.000

H.

See Figure 3.

Figure 3: SACF and SPACF of a $SAR(0) \times (1)_4$ with $\Phi = 0.9$

Date: 05/29/17 Time: 08:06
Sample: 1 1000
Included observations: 1000

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.113	0.113	12.771	0.000
		2	-0.007	-0.020	12.826	0.002
		3	0.114	0.119	25.948	0.000
		4	0.879	0.876	803.09	0.000
		5	0.113	0.014	816.00	0.000
		6	-0.020	-0.057	816.39	0.000
		7	0.117	0.015	830.18	0.000
		8	0.772	-0.003	1431.9	0.000
		9	0.106	-0.025	1443.2	0.000
		10	-0.035	-0.030	1444.5	0.000
		11	0.110	-0.026	1456.8	0.000
		12	0.681	0.013	1927.0	0.000
		13	0.100	0.010	1937.1	0.000
		14	-0.041	0.023	1938.8	0.000
		15	0.104	0.007	1949.9	0.000

Task 4

We consider the dataset of US GDP 1947Q1 - 2016Q4, quarterly data.

A.

We follow the test template for the test of whether the original data contains a unit root or not.

1. Hypotheses

$H_0 : \{Y_t\}$ has a unit root $\Leftrightarrow a = 0$
$H_1 : \{Y_t\}$ does not have a unit root $\Leftrightarrow a < 0$

2. Significance level

$\alpha = 0.05$

3. Estimator(s)/Statistics

OLS estimator of a in $\bigtriangledown Y_t = \phi_1 \bigtriangledown Y_{t-1} + ... + \phi_k \bigtriangledown Y_{t-k} + e_t$

4. Assumptions

T is large.

5. Test statistic

$$ADF_{obs} = \frac{\hat{a} - 0}{\hat{\sigma}_{\hat{a}}}$$

6. Rejection rule and figure

The null hypothesis is rejected if the p-value $< \alpha = 0.05$. We cannot draw a figure since the test statistic does not follow a standard distribution.

7. Calculations and results

From Figure 4.3, we can see that the p-value $= 0.8105 > 0.05$. We cannot reject the null hypothesis.

8. Conclusion

We cannot reject the null hypothesis that the original data contains a unit root at a 5 percent significance level.

B.

The second sample autocorrelation is the correlation between the realization of a process (in this case the residuals) at a given time point and the same process two time points back.

C.

The second sample partial autocorrelation is the correlation between the the realization of a process (in this case the residuals) at a given time point and the same process second time points back, when taking the first autocorrelation into account

D.

A test if the second autocorrelation is different from zero.

1. Hypotheses

$H_0 : \rho_2 = 0$
$H_1 : \rho_2 \neq 0$

2. Significance level

$\alpha = 0.05$

3. Estimator(s)/Statistics

$\hat{\rho}_2$, the sample autocorrelation for lag 2.

4. Assumptions

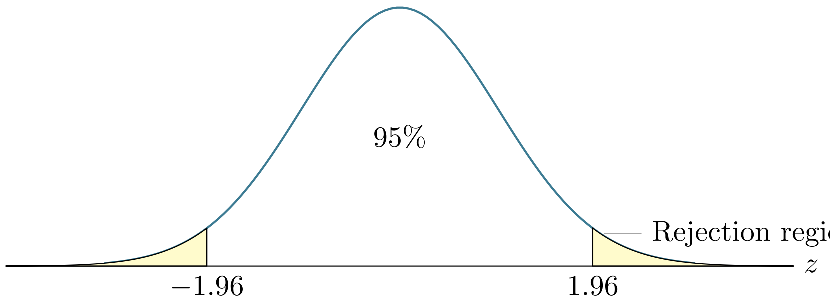
T is large

5. Test statistic

$$z_{obs} = \frac{\hat{\rho}_k - \rho_k^{H_0}}{\sqrt{\frac{1}{T}}} \sim N(0,1)$$

6. Rejection rule and figure

Reject the null hypothesis if $z_{obs} < -1.96$ or if $z_{obs} > 1.96$.



7. Calculations and results

$$z_{obs} = \frac{0.144 - 0}{\sqrt{\frac{1}{279}}} = 2.40... > 1.96$$

We reject the null hypothesis.

8. Conclusion

We reject the null hypothesis that the second autocorrelation is zero at the 5 percent significance level.

E.

A test whether the first three autocorrelations are simultaneously zero.

1. Hypotheses

$H_0 : \rho_1 = \rho_2 = \rho_3 = 0$
$H_1 :$ At least one $\rho_j \neq 0, \ j = 1, 2, 3$

2. Significance level

$\alpha = 0.05$

3. Estimator(s)/Statistics

$\hat{\rho}_j,\ j = 1, 2, 3$

4. Assumptions

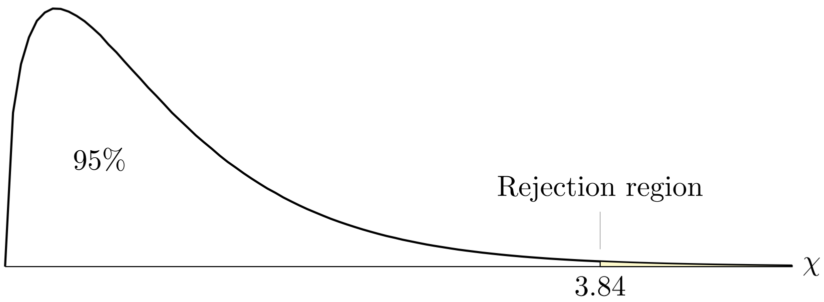
T is large

5. Test statistic

$Q_{LB} = T(T+2)\sum_{j=1}^{K} \frac{\hat{\rho}_j^2}{T-j} \sim \chi_{K-p-q-P-Q}$

6. Rejection rule and figure

Reject the null hypothesis if $\chi_{obs} > \chi_{1,0.05} = 3.84$



7. Calculations and results

$$\begin{aligned} Q_{LB} &= 279 * 281 * \left(\frac{(-0.001)^2}{279-1} + \frac{(-0.011)^2}{279-2} + \frac{(-0.001)^2}{279-3} \right) \\ &= 0.0348... \\ &< 3.84 \end{aligned} \tag{17}$$

We cannot reject the null hypothesis.

8. Conclusion

We cannot reject the null hypothesis that the first three autocorrelations are zero at the 5 percent significance level.

F

Looking at the correlogram in Figure 4.4, an AR(2) seems to be appropriate. The first 24 autocorrelations of the residuals are simultaneously zero and the adjusted R-squared is high (compared with the other estimated models). In addition, the estimators are significantly different from zero.