ECOM073 Midterm test

Wednesday, 27 March, 9:00am-10:00am (60 minutes)

Place: PL-301 (Peter Landin teaching rooms)

Midterm 2024

Wednesday, 27 March, 9:00-10:00 am

1) Basic definitions: stationarity, white noise, i.i.d and others
2) Summary statistics: mean, variance, skewness, kurtosis, Jarque Bera test
3) Testing for absence of correlation
4) AR(p), MA(q) models, selection of order p, q
5) Checking the fit of the model, and significance of parameter estimates
6) AIC and BIC information criterions

Test c**overs: Lecture 2-5, Problem Sets 2-5**

Examples for preparation:

Problem Set 2: 2.1

Problem Set 3: 3.1, 3.2,

Problem Set 4: 4.3, 4.5

Problem Set 5: 5.2

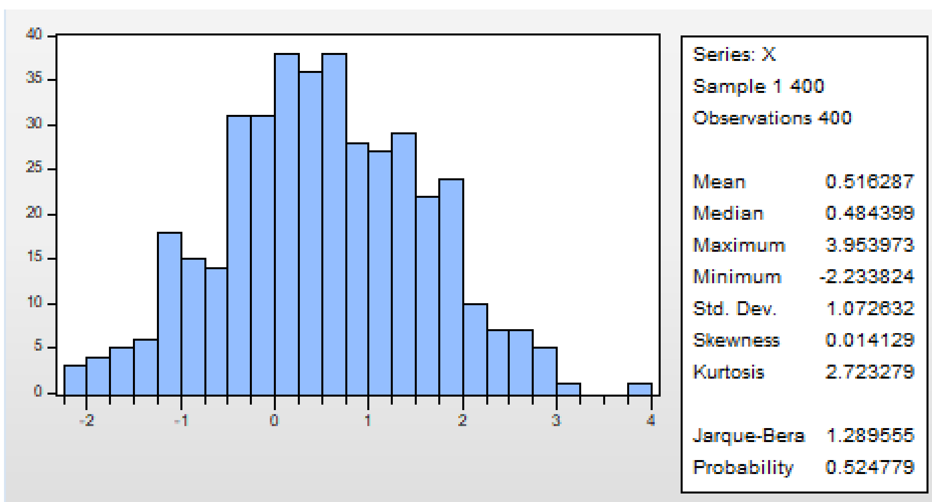
See solutions of Quizzes 2-4 (QMPLUS, Week 7)

Solution Quiz 3

Question 3.

The researcher computed summary statistics of the time series X using a sample containing $N = 100$ observations, see the output below.

1. Test whether skewness $S(X)$ equals 0.
2. Test whether kurtosis $K(X)$ equals 3
3. Test whether X has normal distribution



Solution of Question 3.

1. We test the hypothesis

H_0: $S(X) = 0$ against alternative H_1: $S(X) \neq 0$.

at 5% significance level.
We construct the test statistics:

$$t = \frac{\hat{S}(X)}{\sqrt{6/N}} = \frac{0.014129}{\sqrt{6/400}} = 0.11536$$

Under the null hypothesis, $t \sim N(0,1)$ is normally distributed.

Rule: we reject H_0 at 5% significance level, if

$$|t| \geq 2.$$

In our case $|t| = 0.11536 < 2$. Hence, the test shows that there no evidence in the data to reject the null hypothesis of zero skewness $S(X) = 0$.

2. Next we test at 5% significance level the hypothesis:

H_0: $K(X) - 3 = 0$ against alternative H_1: $K(X) \neq 3$.

We use the test statistics:

$$t = \frac{\hat{K}(X) - 3}{\sqrt{24/N}} = \frac{2.723 - 3}{\sqrt{24/400}} = -1.1308.$$

By theory, under null hypothesis, $t \sim N(0,1)$ is normally distributed. Therefore the testing rule is similar as for testing skewness:

Rule: reject H_0 at 5% significance level, if

$$|t| \geq 2.$$

In our case $|t| = 1.1308 < 2$. Hence, the test does not rejects the null hypothesis of zero kurtosis $K(X) = 3$.

3. Jargue-Bera test is used to test the hypothesis:

H_0: $S(X) = 0$ and $K(X) - 3 = 0$ ("normal distribution") against alternative

H_1: $S(X) \neq 0$ or $K(X) - 3 \neq 0$ ("distribution is not normal").

Normal distribution has $S(X) = 0$ and $K(X) = 3$. Thus, in case of normal distribution, test will not reject H_0.

Test will reject H_0 if either skewness is not 0 or if kurtosis is not 3. That will indicate that distribution is not normal.

Since the $p = 0.5247$ value of Jargue-Bera test is larger than 0.05 we do not reject the asymptotic normality at significance level 5%.

Solution Quiz 3

Question 1. a) Suppose that

$$X_t = \varepsilon_t + t, \quad t = 1, 2, ...$$

where ε_t is a white noise sequence with zero mean and variance $E\varepsilon_t^2 = 1$. Investigate whether time series X_t is covariance stationary.

b) Suppose that

$$X_t = t\varepsilon_t, \quad t = 1, 2, ...$$

where ε_t is a white noise sequence with zero mean and variance $E\varepsilon_t^2 = 1$. Investigate whether time series X_t is covariance stationary.

Solution of Question 1.

Time series X_t is a covariance stationary time series if it satisfies three properties:

– $EX_t = \mu$ for all t (does not depend on t);

– $Var(X_t) = \sigma^2$ for all t (does not depend on t);

– $Cov(X_t, X_{t-k}) = \gamma_k$ - covariance function depends only on the lag k and does not depend on t.

a) We have

$$EX_t = E[\varepsilon_t + t] = E[\varepsilon_t] + t = 0$$

since by assumption $E[\varepsilon_t] = 0$. We see that mean EX_t changes with t. Therefore, this time series is not covariance stationary.

b) We have

$$EX_t = E[t\varepsilon_t] = tE[\varepsilon_t] = 0$$

since by assumption $E[\varepsilon_t] = 0$. Hence, mean EX_t does not change with t.

Next we compute the variance:

$$\text{Var}(X_t) = E(X_t - EX_t)^2 = EX_t^2 = E[(t\varepsilon_t)^2] = t^2 E[\varepsilon_t^2] = t^2.$$

since by assumption $E[\varepsilon_t^2] = 1$. We see that the variance $\text{Var}(x_t)$ changes with t. Therefore, this time series is not covariance stationary.

Question 2. Explain why the following sequence

$$\rho_1 = 0.8, \quad \rho_2 = 0.5, \quad \rho_3 = \rho_1 + \rho_2, \quad \rho_4 = \rho_1 + \rho_2 + \rho_3, ...$$

cannot be the auto-correlation function of a covariance stationary sequence.

Solution of Question 2.

Correlation function ρ_k at lag k of a covariance stationary time series X_t satisfies three properties:
- $\rho_0 = 1$
- $\rho_k = \rho_{-k}$ for any lag $k = 1, 2, 3, ...$
- $|\rho_k| \le 1$ for any k.

We have that

$$\rho_3 = \rho_1 + \rho_2 = 0.8 + 0.5 = 1.3 > 1.$$

Therefore ρ_k cannot ne correlation function of a covariance stationary time series.

Question 3. Using the following EVIEWS correlogram of time series Xt, determine whether x_t is a white noise time series (series of uncorrelated random variables).

Correlogram of Y

Date: 29/11/20 Time: 10:53
Sample: 1 400
Included observations: 400

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.023	-0.023	0.2196	0.639
		2	-0.032	-0.033	0.6458	0.724
		3	-0.036	-0.037	1.1600	0.763
		4	-0.057	-0.060	2.4792	0.648
		5	0.050	0.045	3.4877	0.625
		6	-0.076	-0.080	5.8704	0.438
		7	0.067	0.063	7.7087	0.359
		8	-0.009	-0.012	7.7392	0.459
		9	0.012	0.016	7.7940	0.555
		10	-0.041	-0.049	8.4897	0.581

Brief solution of question 3.

(i) (**Testing for correlation using ACF**.) Time series is a white noise if it is serially uncorrelated, that is $\rho_k = 0$ for $k \geq 1$. Hence, to test for white noise, we test the hypotheses

$H_0 : \rho_k = 0$ against alternative $H_1 : \rho_k \neq 0$
at each lag $k = 1, 2, ...$ at significance level 5%.

Rule: ACF ρ_k at lag k is significantly different from zero at 5% significance level if $|\hat{\rho}_k| > 2/\sqrt{N}$, where N is the number of observations.

If $|\hat{\rho}_k| \leq 2/\sqrt{N}$, then ACF at lag k is not significantly different from 0.

If time series is white noise, then we do not reject H_0 for any $k = 1, 2, ...$

(ii) **Ljung-Box test**. This test can be also used to test for zero correlation. We select $m = 1, 2, ...$ and test the hypothesis

$H_0 : \rho_1 = ...\rho_m = 0$ against alternative

$H_1 : \rho_j \neq 0$ for some $j = 1, ..., m$.

We reject the H_0 at 5% significance level, if $p-$ value satisfies $p < 0.05$. If time series is white noise, then we do not reject H_0 for any $m = 1, 2, ...$

(iii) We have $2/\sqrt{N} = 2/\sqrt{400} = 0.1$. Since $|\rho_1| = 0.023 < 0.1$, $|\rho_2| = 0.032 < 0.1$, ..., and so on. We find that ρ_k is not significant at any lag $k = 1, 2, ...$ at 5% significance level, because $|\rho_k| \leq 2/\sqrt{N} = 2/\sqrt{400} = 0.1$. So, this time series is a white noise.

From the correlogram we see that p-values of Ljung Box test satisfy $p > 0.05$ for all $m = 1, ...$ So H_0 is not rejected at any m and this time series is a white noise.

Solution Quiz 4

Question 1.

Using the following Eviews output determine order p of AR(p) model and order q of MA(q) model you would fit to the data.

Correlogram of Y

Date: 04/10/20 Time: 09:39
Sample: 1 625
Included observations: 625

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.487	-0.487	149.22	0.000
		2	0.289	0.067	201.67	0.000
		3	0.012	0.232	201.77	0.000
		4	-0.016	0.074	201.94	0.000
		5	0.062	0.015	204.35	0.000
		6	-0.013	0.000	204.47	0.000
		7	-0.019	-0.051	204.70	0.000
		8	0.039	-0.001	205.68	0.000
		9	0.007	0.059	205.72	0.000
		10	-0.047	-0.033	207.13	0.000

Write down the model you would fit to the data.

Comment on how would you check the fit of your model to the data.

Brief solution of Question 1

(i) To select the order p for AR(p) model, we use the sample PACF function. We test the hypothesis

$H_0: \rho_k = 0$ against alternative $H_1: \rho_k \neq 0$

at lags $k = 1, 2, ...$ at significance level 5%, where ρ_k is the PACF function.

PACF $\hat{\rho}_k$ at lag k is significantly different from 0 at 5% significance level, if $|\hat{\rho}_k| > 2/\sqrt{N}$, where N is the number of observations.

If $|\hat{\rho}_k| \leq 2/\sqrt{N}$, then PACF at lag k is not significantly different from 0.

Rule: we select for p the largest lag k at which the PACF is significant.

This rule can be used because PACF of the AR(p) model becomes 0 for $k > p$.

(ii) To select the order q of MA(q) model, we use the sample ACF function. We test the hypothesis

$H_0: \rho_k = 0$ against alternative $H_1: \rho_k \neq 0$

at lags $k = 1, 2, ...$ at significance level 5%, where ρ_k is ACF function.

Rule: ACF ρ_k is significant at lag k at 5% significance level, if $|\hat{\rho}_k| > 2/\sqrt{N}$, where N is the number of observations.

If $|\hat{\rho}_k| \leq 2/\sqrt{N}$, then ACF at lag k is not significantly different from 0.

We select for q the largest lag k at which the ACF is significant.

(iii) We have $2/\sqrt{N} = 2/\sqrt{625} = 0.08$. The PACF is significant only at lag 1 and 3. Hence we would fit AR(3) model.

The ACF shows significant correlation at the lags 1 and 2. Hence we would fit MA(2) model.

From the two models AR(3) and MA(2) we select a simpler model MA(2) with smaller number of parameters which should to be fitted to the data.

(iv) According to above, we can fit MA(2) model $Y_t = \phi_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$ where ε_t is white noise.

(v) This MA(2) model fits the data if residuals $\hat{\varepsilon}_t$ are uncorrelated. We could use the correlogram of residuals to test whether residuals are uncorrelated.

Question 2.

Using AIC information criterion values obtained fitting an AR(p) model, select the order p of an AR model you would fit to the data:

p	0	1	2	3	4	5	6
AIC	3	-2.3	-2	-1.1	0.6	1.7	1.8

Write down equation of your AR(p) model.

Solution of Question 2. Using AIC information criterion we select the model which minimizes the AIC value. In this case the minimum value -2.3 corresponds to $p = 1$, so AR(1) model should be fitted to the data.