Written Examination in Time Series Analysis (B3)

Spring 2016

2016-04-29 14.00-18.00

Bergsbrunnagatan 15, room2.

Lars Forsberg, Department of Statistics, Uppsala University

Allowed means of assistance:

1. Pen or **pencil** (recommended) and eraser
2. **Calculators**,
   (a) 'programmable' calculator, e.g. calculator with graphing functions is OK.
   (b) Calculators with blue-tooth are not allowed.
   (c) Calculators with access to internet are not allowed.
   (d) Calcuators with which it is possible to send and recieve messages of any kind are not allowed.
3. **Physical (paper) dictionary** (no electronic dictionary allowed).
   (a) Dictionary must contain *no notes* of any kind.
   (b) Each student must have his/her own dictionary. It is not allowed for students to pass a dictionary between them.
4. **Ruler.**
5. Collection of formulae and Statistical Tables named '*Collection of Formulae and Statistical Tables for the B2-Econometrics and B3-Time Series Analysis courses and exams*', that the student brings to the exam location.
6. Please note that a collection of critical values for the Student´s t, Normal, Chi-square and F-distributions is given in the Appendix of the '*Collection of Formulae and Statistical Tables for the B2-Econometrics and B3-Time Series Analysis courses and exams*'.
7. Also note that the '*Test template*', that should be used when performing tests, is given in the '*Collection of Formulae and Statistical Tables for the B2-Econometrics and B3-Time Series Analysis courses and exams'.*

That is:

1. NO BOOK (except paper-dictionary) is allowed.
2. NO (student-written) notes are allowed.
3. NO other document than the one 'Collection of Formulae and Statistical Tables for Time Series Exam' is allowed.

Instructions: **Please note the following:**

1. Start with reading through the instructions!
2. Make sure you **follow** the instructions!
3. Start with reading through the exam.
4. You may write your solutions in Swedish or English.
5. Total score is **100** points
   (a) If you want the ECTS grades, please indicate that on the cover page!
   (b) For each task the maximum number of points is given within parenthesis, e.g. (16p in total).
   (c) For each subtask the number of points is given within parenthesis, e.g. (2p)
6. All solutions must be on separate sheets. No solutions on the questionnaire! (If so, they will be disregarded.)
7. Make sure your solutions are: easy to read and easy to understand, that is:
   (a) For each task that you solve, please start with a new sheet: after Task 1, start with a blank sheet for Task 2, etc.
   (b) Write the *task number* at the top of each page, in the

   ....**MIDDLE OF THE PAGE!!!**..

   Like:

   ....**TASK 1**..

   - if you write it in the upper left corner, the staple will cover it, and there is no for way for the examinator to know if the text of that sheet belongs to the previous sub-task or what it is. The Examinators will not make any 'qualified guesses' of what is being displayed on any given page. It is the responsibility of the student to make sure that every task and sub-task is easily identifiable.

(c) If you continue a sub-task on the next sheet of paper - indicate that at the top of the page - IN THE MIDDLE OF THE PAGE, like, for example:

....'Task 1B (cont.)'..

(d) Please separate each subtask A, B etc with a horizontal line across the sheet

_ _
if they are on the same sheet of paper - that way it will be easy for the examinator to actually see where one subtask ends and next begins.

(e) For examinator readability, it is highly recommended that you use a pencil, (and not a pen), which will allow you to erase and rewrite if you make a mistake. Crossed-over text and corrections using 'tipp-ex' will just cause blurriness and confusion to the examinator.

(f) For examinator readability: Write clearly, that is, letters, mathematical/statistical symbols and numbers should be easy recognizable!! Do not underestimate the correlation between readability and points scored, that is, when readability goes to zero, points scored also goes to zero, no matter your intentions or wheather *you* can read it or not.

(g) Also note that everything that you write will be taken at 'face value'. That is, for example, if you write β_1 the examinator will take that as a β_1 even though you may claim that it is given from the context it should be clear that you meant something else, like β_3. Thus, given this example, writing β_1, and that is not correct in that specific formula or statement, this will lead to subtraction of points, even if you will claim that it is just a typo, and that in another task or subtask, it is clear that you understand the issue.

(h) Please put the sheets in **order**, that is first Task 1, and then Task 2 etc...

8. Please keep the questionaire.

9. Do well!

Task 1

(24 points in total)

A) (2p) What is a stochastic process from a theoretical point of view? Explain using words, no formulae needed.

B) (6p) State the three conditions for a stochastic process to be covariance stationary. For each condition, state that condition using formulae and also explain in words what it means.

C) (2p) Give an example of a process that is *trend-stationary.* That is, write down the process using a formula, make sure you define all the 'parts' of the process. Explain in words why it is not stationary (if you by any chance find this sentence confusing - yes, it says and should say *not* stationary - what condition of the stationary requirements is not fulfilled?).

D) (2p) Give an example of a process that is *difference-stationary.* That is, write down the process using a formula, make sure you define all the 'parts' of the process. Explain in words why it is not stationary (if you by any chance find this sentence confusing - yes, it says and should say *not* stationary - what condition of the stationary requirements is not fulfilled?).

E) (8p) State the four stages in the Box-Jenkins methodology for modeling and forecasting time series. For each stage, give at least one examle of a statistical method, procedur, statistic, estimator or test that can be used at that stage. Also, what property must the data", or the assumed true"underlying statistical process have, in order for it to make sense to even start with the first stage.

F) (4p) For a general (covariane stationary) process $\{Y_t\}$:

1. Interpret the fourth autocorrelation, (no need to state the defintion, just interpret it).
2. Interpret the fourth *partial* autocorrelation, (no need to state the defintion, just interpret it).

Task 2

(36 points in total)

Consider a SARMA process

$$\phi(B)\,\Phi(B)\,Y_t = \theta(B)\,\Theta(B)\,e_t$$

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

For each and every process specified by the lag polynomials in subtasks A through F, do the following:

1. Given the specification of the process as SARMA$(p,q) \times (P,Q)_s$, write down the specific values of p, q, P, Q and s, for this specific process, for example, white noise would be SARMA$(0,0) \times (0,0)_0$.
2. Write out the process in such a way that Y_t is alone on the Left Hand Side (LHS) of the equality sign, and the backshift operator does *not* occur on the Right Hand Side (RHS). Write out the process in terms of parameters using the greek letters, that is, do *not* replace the parameters with the numerical values below, these are for the ACF and PACF.
3. Sketch the ACF and PACF for the process using the parameter values in the right most column. Use as many lags as you see fit so that the pattern specific to the particular process becomes obvious.

			For ACF and PACF use
A) (6p)	$\phi(B) = (1 - \phi B)$ $\Phi(B) = 1$	$\theta(B) = 1$ $\Theta(B) = 1$	$\phi = 0.9$
B) (6p)	$\phi(B) = 1$ $\Phi(B) = 1$	$\theta(B) = (1 - \theta_1 B)$ $\Theta(B) = 1$	$\theta_1 = 0.9$
C) (6p)	$\phi(B) = (1 - \phi_1 B - \phi_2 B^2)$ $\Phi(B) = 1$	$\theta(B) = 1$ $\Theta(B) = 1$	$\phi_1 = 0.5$ $\phi_2 = 0.3$
D) (6p)	$\phi(B) = 1$ $\Phi(B) = 1$	$\theta(B) = (1 - \theta_1 B - \theta_2 B^2)$ $\Theta(B) = 1$	$\theta_1 = 0.5$ $\theta_2 = 0.3$
E) (6p)	$\phi(B) = 1$ $\Phi(B) = 1$	$\theta(B) = 1$ $\Theta(B) = (1 - \Theta_1 B^4)$	$\Theta_1 = 0.9$
F) (6p)	$\phi(B) = 1$ $\Phi(B) = (1 - \Phi_1 B^4)$	$\theta(B) = 1$ $\Theta(B) = 1$	$\Phi_1 = 0.9$

Task 3

(11 points in total)

Consider the following process

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

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

Let

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

and

$$\theta(B) = 1. \tag{3}$$

A) (3p) For the model defined by (1), (2) and (3), write down the characteristic equation and solve that. For what values of the root is the process covariance stationary?

B) (2p) For the model defined by (1), (2) and (3), derive the *expected value* of the process. Be explicit in what assumptions, if any, you make to be able to derive this result.

C) (2p) For the model defined by (1), (2) and (3), derive the *variance* of the process. Be explicit in what assumptions, if any, you make to be able to derive this result.

D) (4p) For the model defined by (1), (2) and (3), derive the first *two* auto*covariances* for the process, that is, derive $Cov(Y_t, Y_{t-1})$ and $Cov(Y_t, Y_{t-2})$. You may write the autocovariances as functions of other autocovariances if you so see fit.

Task 4

(11 points in total)

Consider the following stochastic process

$$Y_t = e_t - \theta e_{t-1}, \tag{4}$$

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

A) (3p) For the model above, write down the characteristic equation and solve that. For what values of the root is the process covariance stationary? For what values of the root is it invertible?

B) (2p) Derive the expected value of the process. Be explicit in what assumptions, if any, you make to be able to derive this result.

C) (2p) Derive the variance of the process in (4). Be explicit in what assumptions, if any, you make to be able to derive this result.

D) (4p) Derive the first two autocovariances for the process in (4), that is derive $Cov(Y_t, Y_{t-1})$ and $Cov(Y_t, Y_{t-2})$. Be explicit in what assumptions, if any, you make to be able to derive these results.

Task 5

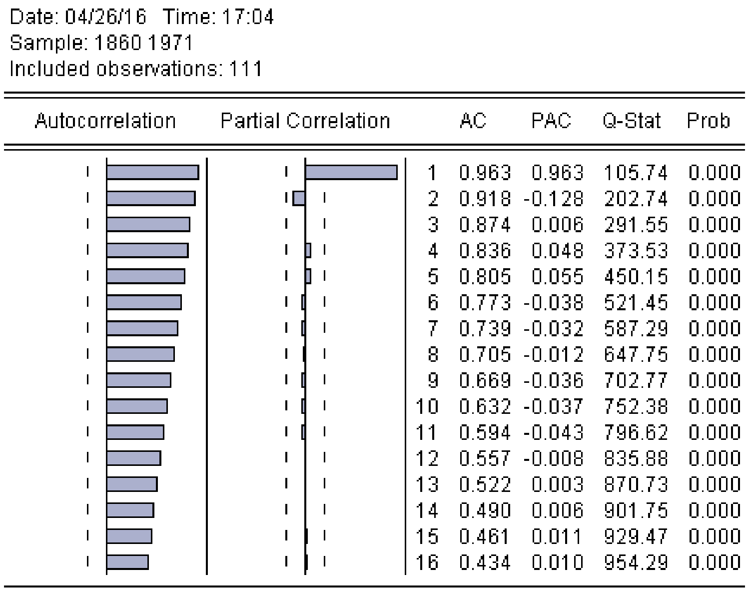
(18 points in total)

A time series of US CPI yearly data from 1860 to 1970 is being analyzed. Your task is to perform some tests on this data, you are given Eviews outputs to this end. Use them to perform the tests in the following subtasks.

A) (6p) Perform an unit root test, testing if $\ln(CPI)$ has a unit root, use significance level 5%. Document the test procedure as outlined in the test-template.

B) (6p) For the first difference of $\ln(CPI)$, test if the second (and only the second) autocorrelation is equal to zero, against the alternative that it is not zero. Use significance level 5 %. Document the test procedure as outlined in the test-template.

C) (6p) An analyst has fitted an ARMA(1,1) model to the first difference of the $\ln(CPI)$. Perform a test to test the null hypothesis that the residuals are uncorrelated up to and including lag 10. Use significance level 5%. Document the test procedure as outlined in the test-template.


Date: 04/26/16 Time: 17:04
Sample: 1860 1971
Included observations: 111

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.963	0.963	105.74	0.000
		2	0.918	-0.128	202.74	0.000
		3	0.874	0.006	291.55	0.000
		4	0.836	0.048	373.53	0.000
		5	0.805	0.055	450.15	0.000
		6	0.773	-0.038	521.45	0.000
		7	0.739	-0.032	587.29	0.000
		8	0.705	-0.012	647.75	0.000
		9	0.669	-0.036	702.77	0.000
		10	0.632	-0.037	752.38	0.000
		11	0.594	-0.043	796.62	0.000
		12	0.557	-0.008	835.88	0.000
		13	0.522	0.003	870.73	0.000
		14	0.490	0.006	901.75	0.000
		15	0.461	0.011	929.47	0.000
		16	0.434	0.010	954.29	0.000

Fgure 5.1: Correlogram of ln(CPI)

Null Hypothesis: LOG_CPI has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-0.482385	0.8894
Test critical values:	1% level	-3.491345	
	5% level	-2.888157	
	10% level	-2.581041	

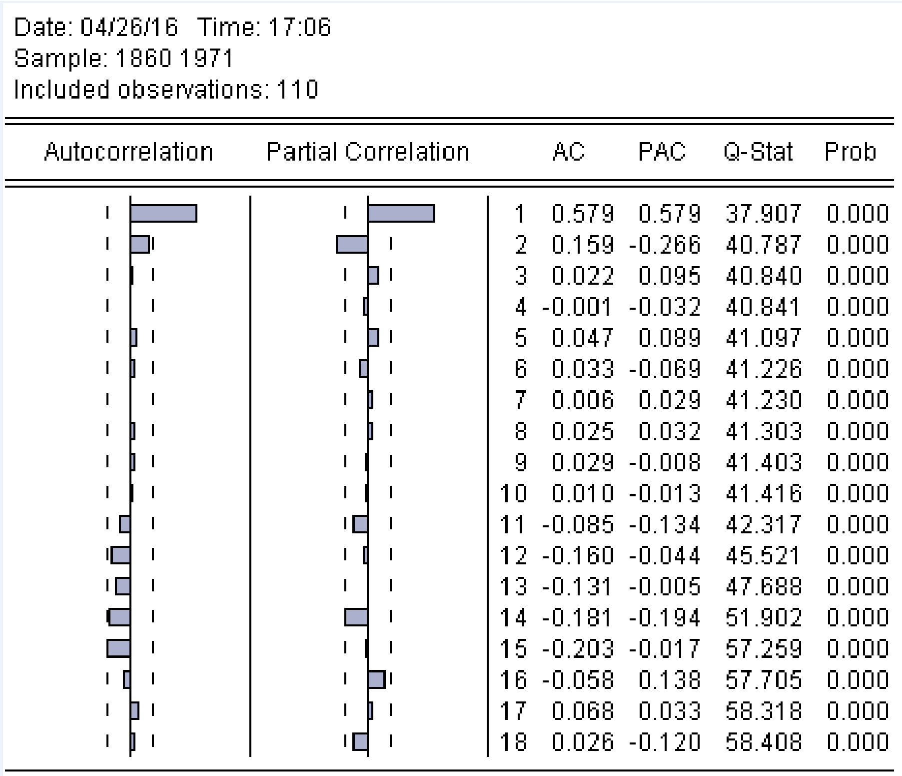
*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG_CPI)
Method: Least Squares
Date: 04/26/16 Time: 17:05
Sample (adjusted): 1862 1970
Included observations: 109 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG_CPI(-1)	-0.005172	0.010722	-0.482385	0.6305
D(LOG_CPI(-1))	0.589982	0.080740	7.307192	0.0000
C	0.025231	0.040335	0.625526	0.5330

R-squared	0.338878	Mean dependent var	0.013398
Adjusted R-squared	0.326404	S.D. dependent var	0.057578
S.E. of regression	0.047256	Akaike info criterion	-3.239357
Sum squared resid	0.236707	Schwarz criterion	-3.165283
Log likelihood	179.5449	Hannan-Quinn criter.	-3.209317
F-statistic	27.16674	Durbin-Watson stat	1.640865
Prob(F-statistic)	0.000000		

Figure 5.2: ADF output



Date: 04/26/16 Time: 17:06
Sample: 1860 1971
Included observations: 110

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.579	0.579	37.907	0.000
		2	0.159	-0.266	40.787	0.000
		3	0.022	0.095	40.840	0.000
		4	-0.001	-0.032	40.841	0.000
		5	0.047	0.089	41.097	0.000
		6	0.033	-0.069	41.226	0.000
		7	0.006	0.029	41.230	0.000
		8	0.025	0.032	41.303	0.000
		9	0.029	-0.008	41.403	0.000
		10	0.010	-0.013	41.416	0.000
		11	-0.085	-0.134	42.317	0.000
		12	-0.160	-0.044	45.521	0.000
		13	-0.131	-0.005	47.688	0.000
		14	-0.181	-0.194	51.902	0.000
		15	-0.203	-0.017	57.259	0.000
		16	-0.058	0.138	57.705	0.000
		17	0.068	0.033	58.318	0.000
		18	0.026	-0.120	58.408	0.000

Figure 5.3: Correlogram of the first difference of ln(CPI).

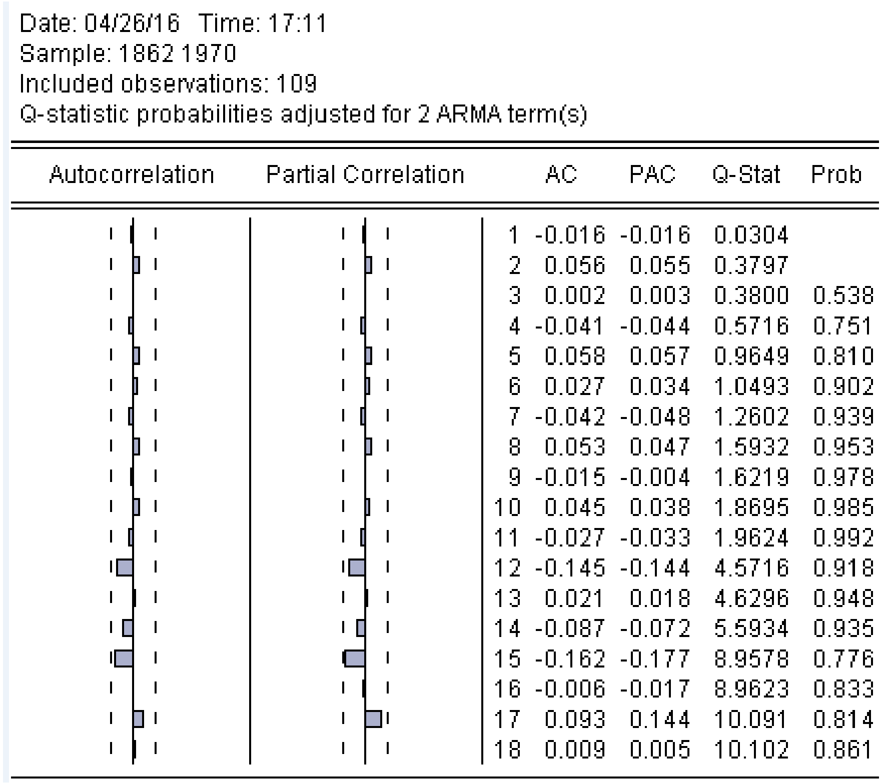
Dependent Variable: D(LOG_CPI)
Method: Least Squares
Date: 04/26/16 Time: 17:11
Sample (adjusted): 1862 1970
Included observations: 109 after adjustments
Convergence achieved after 7 iterations
MA Backcast: 1861

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.013710	0.008571	1.599516	0.1127
AR(1)	0.198767	0.128106	1.551581	0.1237
MA(1)	0.604180	0.101571	5.948342	0.0000

R-squared	0.406909	Mean dependent var	0.013398
Adjusted R-squared	0.395718	S.D. dependent var	0.057578
S.E. of regression	0.044758	Akaike info criterion	-3.347947
Sum squared resid	0.212349	Schwarz criterion	-3.273873
Log likelihood	185.4631	Hannan-Quinn criter.	-3.317907
F-statistic	36.36227	Durbin-Watson stat	2.007867
Prob(F-statistic)	0.000000		

Inverted AR Roots	.20
Inverted MA Roots	-.60

Figure 5.3



Date: 04/26/16 Time: 17:11
Sample: 1862 1970
Included observations: 109
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.016	-0.016	0.0304	
		2	0.056	0.055	0.3797	
		3	0.002	0.003	0.3800	0.538
		4	-0.041	-0.044	0.5716	0.751
		5	0.058	0.057	0.9649	0.810
		6	0.027	0.034	1.0493	0.902
		7	-0.042	-0.048	1.2602	0.939
		8	0.053	0.047	1.5932	0.953
		9	-0.015	-0.004	1.6219	0.978
		10	0.045	0.038	1.8695	0.985
		11	-0.027	-0.033	1.9624	0.992
		12	-0.145	-0.144	4.5716	0.918
		13	0.021	0.018	4.6296	0.948
		14	-0.087	-0.072	5.5934	0.935
		15	-0.162	-0.177	8.9578	0.776
		16	-0.006	-0.017	8.9623	0.833
		17	0.093	0.144	10.091	0.814
		18	0.009	0.005	10.102	0.861

Figure 5.4: Correlogram of the residuals from an ARMA(1,1) modell fitted to the first difference of ln(CPI).