Practice Problems 2

Question 1. Write down the equation for a pure seasonal S-AR(1) model. Describe how its AC and PAC functions look like.

Question 2. Write down the equation for a model with regular AR(1) and seasonal S-AR(1) components and describe how its AC and PAC functions look like.

Question 3. Explain the difference between estimation sample and prediction sample.

Question 4. Explain the difference between in-sample evaluation and out-of-sample evaluation.

Question 5. Explain how Mean Squared Error, Mean Absolute Error, and Mean Loss are used in the assessment of forecasts.

Question 6. Give an example of a deterministic trend $g(t)$ other than a linear trend and plot its graph. Write the model equation for this trend.

Question 7. Write the equation for pure random walk process and the equation for a random walk process with a drift. Explain the main difference between the two.

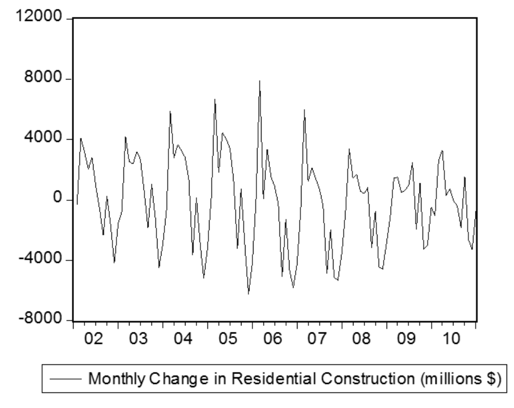
Question 8. Draw a typical correlogram for a random walk process.

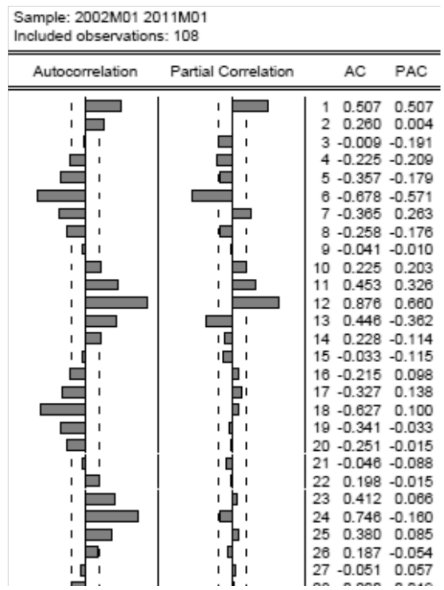
Question 9. Explain the difference between a trend stationary time series and a difference stationary time series.

Question 10. Explain the idea behind the Dickey Fuller unit root test.

Question 11. Explain what it means that a time series process is $I(1)$, and what it means that a process is $I(0)$.

Question 12. Consider the data for monthly changes in U.S. residential construction for the period January 2002-January 2011 shown below. Discuss what kind of model would you estimate for this time series and explain why.





	AC	PAC
1	0.507	0.507
2	0.260	0.004
3	-0.009	-0.191
4	-0.225	-0.209
5	-0.357	-0.179
6	-0.678	-0.571
7	-0.365	0.263
8	-0.258	-0.176
9	-0.041	-0.010
10	0.225	0.203
11	0.453	0.326
12	0.876	0.660
13	0.446	-0.362
14	0.228	-0.114
15	-0.033	-0.115
16	-0.215	0.098
17	-0.327	0.138
18	-0.627	0.100
19	-0.341	-0.033
20	-0.251	-0.015
21	-0.046	-0.088
22	0.198	-0.015
23	0.412	0.066
24	0.746	-0.160
25	0.380	0.085
26	0.187	-0.054
27	-0.051	0.057

Question 13. Consider two candidate models for change in private residential construction spending, AR(1)+SAR(1) and AR(2)+SAR(1), the results for which are below. Discuss which of these models would be preferred based on plots and correlograms of residuals, AIC and BIC, and statistical significance of coefficients.

Dependent Variable: DCONST
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/08/17 Time: 16:14
Sample: 1993M02 2013M12
Included observations: 251
Convergence achieved after 5 iterations
Coefficient covariance computed using outer product of gradients

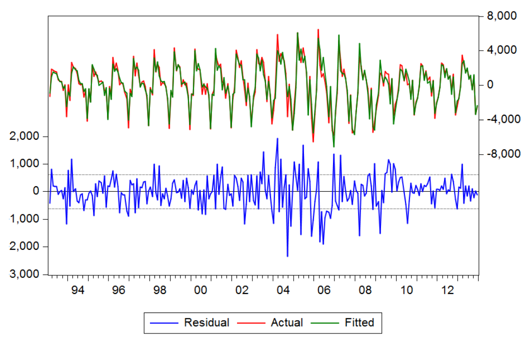
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	218.5690	1235.621	0.176890	0.8597
AR(1)	0.563509	0.044817	12.57349	0.0000
SAR(12)	0.944410	0.013547	69.71517	0.0000
SIGMASQ	379202.0	27048.41	14.01938	0.0000

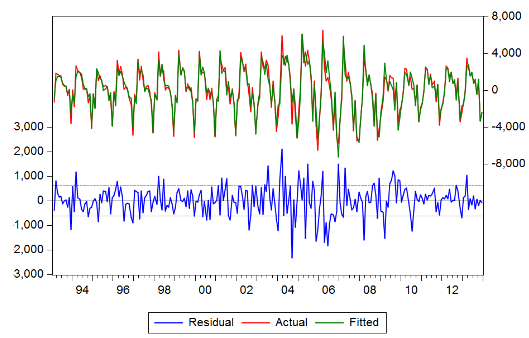
R-squared	0.933686	Mean dependent var	49.44223
Adjusted R-squared	0.932881	S.D. dependent var	2396.078
S.E. of regression	620.7600	Akaike info criterion	15.82347
Sum squared resid	95179704	Schwarz criterion	15.87965
Log likelihood	-1981.845	Hannan-Quinn criter.	15.84608
F-statistic	1159.242	Durbin-Watson stat	2.130377
Prob(F-statistic)	0.000000		

Dependent Variable: DCONST
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/08/17 Time: 16:14
Sample: 1993M02 2013M12
Included observations: 251
Convergence achieved after 7 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	214.4219	1374.109	0.156044	0.8761
AR(1)	0.497140	0.054595	9.105947	0.0000
AR(2)	0.116143	0.052147	2.227211	0.0268
SAR(12)	0.944592	0.013249	71.29646	0.0000
SIGMASQ	373960.6	26485.07	14.11968	0.0000

R-squared	0.934603	Mean dependent var	49.44223
Adjusted R-squared	0.933540	S.D. dependent var	2396.078
S.E. of regression	617.7066	Akaike info criterion	15.81783
Sum squared resid	93864109	Schwarz criterion	15.88805
Log likelihood	-1980.137	Hannan-Quinn criter.	15.84609
F-statistic	878.9103	Durbin-Watson stat	1.975678
Prob(F-statistic)	0.000000		





Date: 04/08/17 Time: 16:14
Sample: 1993M01 2013M12
Included observations: 251
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.066	-0.066	1.1109	
		2	0.136	0.132	5.8246	
		3	-0.022	-0.005	5.9460	0.015
		4	-0.010	-0.030	5.9704	0.051
		5	0.058	0.061	6.8484	0.077
		6	-0.075	-0.065	8.3212	0.080
		7	-0.081	-0.109	10.030	0.074
		8	0.116	0.132	13.549	0.035
		9	-0.088	-0.054	15.602	0.029
		10	-0.017	-0.073	15.675	0.047
		11	-0.044	-0.013	16.196	0.063
		12	-0.004	0.010	16.201	0.094
		13	-0.016	-0.046	16.267	0.132
		14	-0.008	0.010	16.284	0.179
		15	0.004	0.030	16.288	0.234
		16	0.009	-0.027	16.308	0.295
		17	0.050	0.054	16.997	0.319
		18	-0.093	-0.085	19.357	0.251
		19	0.047	0.020	19.956	0.276
		20	0.068	0.098	21.224	0.268
		21	-0.041	-0.048	21.684	0.300
		22	0.029	-0.011	21.918	0.345
		23	-0.020	0.022	22.033	0.398
		24	0.152	0.147	28.514	0.159
		25	-0.000	-0.029	28.514	0.197
		26	-0.059	-0.051	29.501	0.202
		27	-0.028	-0.033	29.724	0.235
		28	0.025	0.020	29.903	0.272
		29	-0.047	-0.041	30.532	0.291
		30	-0.033	-0.026	30.842	0.324
		31	-0.043	0.000	31.369	0.348
		32	0.058	0.025	32.353	0.351
		33	0.002	0.019	32.354	0.400
		34	0.033	0.043	32.675	0.434
		35	-0.022	-0.003	32.818	0.476
		36	-0.072	-0.129	34.329	0.452

Date: 04/08/17 Time: 16:14
Sample: 1993M01 2013M12
Included observations: 251
Q-statistic probabilities adjusted for 3 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.011	0.011	0.0322	
		2	0.054	0.054	0.7743	
		3	-0.054	-0.056	1.5282	
		4	-0.032	-0.034	1.7972	0.180
		5	0.042	0.049	2.2539	0.324
		6	-0.091	-0.092	4.3953	0.222
		7	-0.081	-0.089	6.1091	0.191
		8	0.112	0.132	9.3965	0.094
		9	-0.073	-0.078	10.799	0.095
		10	-0.031	-0.065	11.048	0.137
		11	-0.041	-0.009	11.501	0.175
		12	-0.005	-0.001	11.507	0.243
		13	-0.013	-0.053	11.553	0.316
		14	-0.007	0.014	11.565	0.397
		15	0.006	0.021	11.575	0.480
		16	0.019	-0.024	11.670	0.555
		17	0.040	0.045	12.109	0.598
		18	-0.095	-0.096	14.580	0.482
		19	0.042	0.037	15.068	0.520
		20	0.070	0.082	16.416	0.495
		21	-0.041	-0.061	16.875	0.532
		22	0.010	-0.006	16.904	0.596
		23	-0.005	0.036	16.912	0.659
		24	0.158	0.146	23.933	0.296
		25	0.008	-0.038	23.952	0.350
		26	-0.075	-0.038	25.527	0.324
		27	-0.033	-0.021	25.837	0.362
		28	0.024	0.018	25.999	0.408
		29	-0.044	-0.047	26.564	0.432
		30	-0.043	-0.024	27.097	0.459
		31	-0.037	-0.000	27.498	0.491
		32	0.064	0.024	28.673	0.482
		33	0.022	0.028	28.816	0.527
		34	0.044	0.058	29.388	0.549
		35	-0.016	-0.004	29.459	0.596
		36	-0.071	-0.119	30.955	0.569

Question 14. Consider a regression for the test of equal predictive ability for fixed scheme forecast vs simple four quarter moving average forecast

$$\Delta L_{t+j,1} = \beta_0 + u_{t+j} \qquad \text{with } j = 0, 1, 2, \ldots, T - t - 1$$

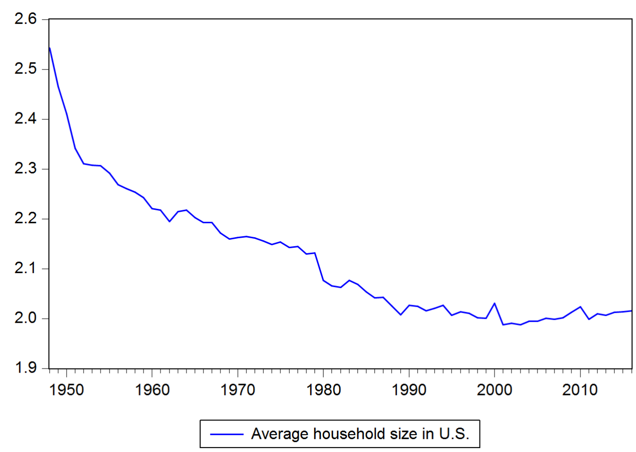
where $\Delta L_{t+j,1} = (e^{ma}_{t+j,1})^2 - (e^{fixed}_{t+j,1})^2$, the results of which are below. Explain the idea behind this test and interpret the results below.

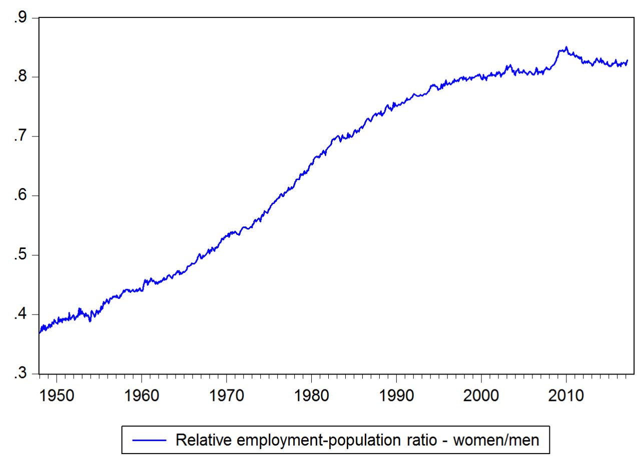
Dependent Variable: DL_MA
Method: Least Squares
Date: 04/08/17 Time: 19:21
Sample: 2009Q1 2016Q4
Included observations: 32
HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.81E-05	3.96E-05	2.224905	0.0335

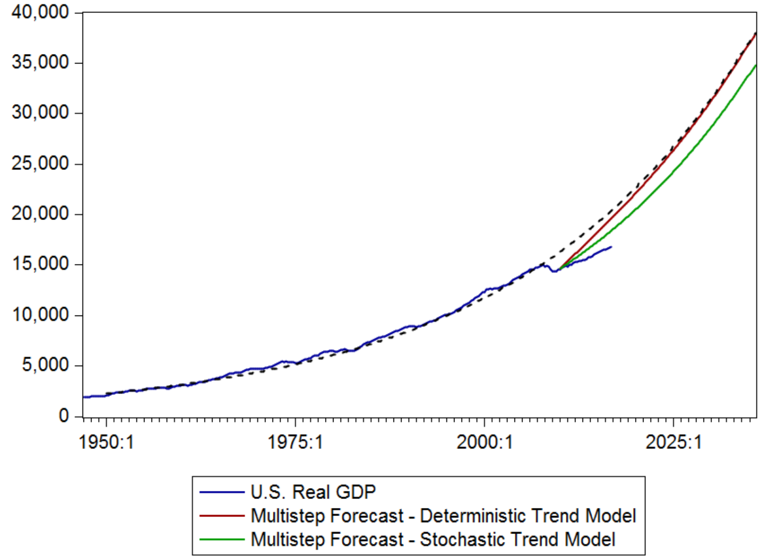
R-squared	0.000000	Mean dependent var	8.81E-05
Adjusted R-squared	0.000000	S.D. dependent var	0.000237
S.E. of regression	0.000237	Akaike info criterion	-13.82388
Sum squared resid	1.75E-06	Schwarz criterion	-13.77808
Log likelihood	222.1821	Hannan-Quinn criter.	-13.80870
Durbin-Watson stat	2.025517		

Question 15. Discuss the choice of a trend when developing models for the two series plotted below.





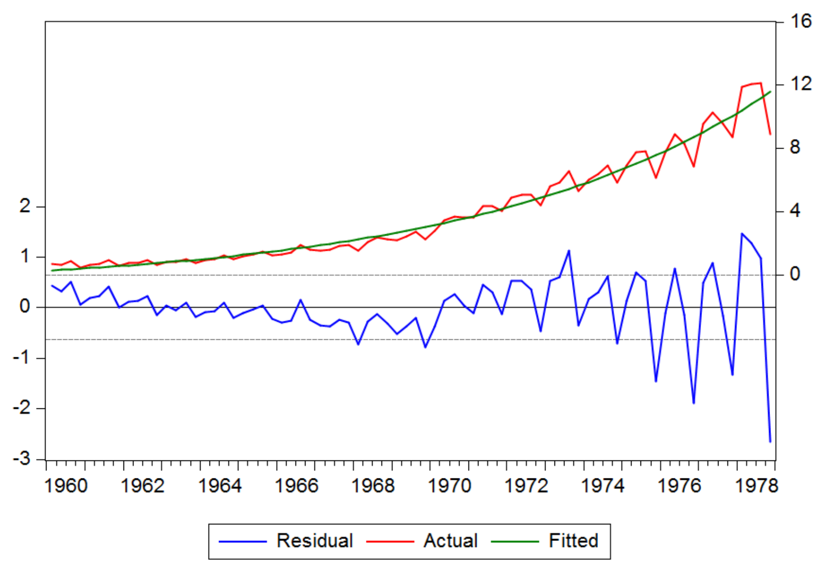
Question 16. The following figure shows the multistep forecasts for the U.S. real GDP, from the deterministic model and from the stochastic trend model, both for the period 2010Q1-2035Q4. Discuss the main difference in the behavior of the forecast and explain the reason for this difference.



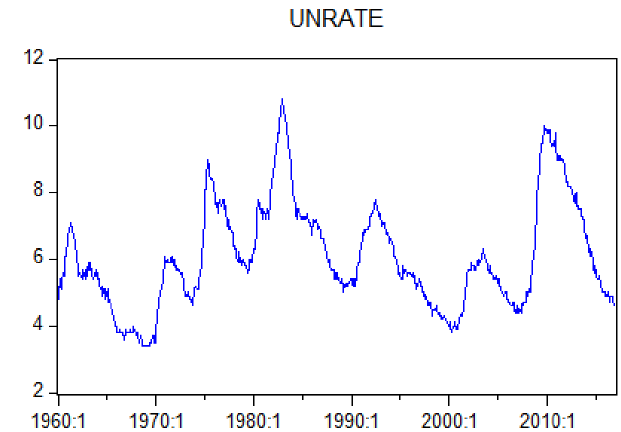
Question 17. Consider a model for quarterly earnings per share of the Johnson and Johnson company

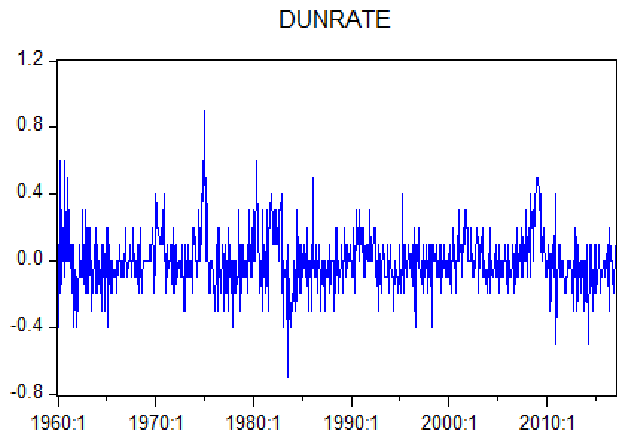
$$JNJ_t = \beta_0 + \beta_1 e^{\beta_2 t} + \varepsilon_t$$

Given the plot with actual values, fitted values, and residuals below, explain how you would proceed with modifying/developing the model further.



Question 18. Interpret the below results of the Augmented Dickey-Fuller test for unemployment rate `UNRATE` and its first difference `DUNRATE`, determine whether unemployment is $I(0)$ or $I(1)$. Explain why only constant was used in both tests.





Null Hypothesis: UNRATE has a unit root
Exogenous: Constant
Lag Length: 4 (Automatic - based on SIC, maxlag=19)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.057041	0.0304
Test critical values:	1% level	-3.439682	
	5% level	-2.865549	
	10% level	-2.568961	

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

Null Hypothesis: D(UNRATE) has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=19)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.891697	0.0000
Test critical values:	1% level	-3.439682	
	5% level	-2.865549	
	10% level	-2.568961	

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

Question 19. Below are the results for the Augmented Dickey-Fuller unit root test for log transformed earnings per share $\log JNJ_t$, and for the first difference of the log transformed earnings per share $\Delta \log JNJ_t$. Interpret the results, and determine whether $\log JNJ_t$ is $I(0)$ or $I(1)$. Explain why trend and constant were used in the test for $\log JNJ_t$ but only constant was used in the test for $\Delta \log JNJ_t$.

Null Hypothesis: LJNJ has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 3 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.696535	0.7428
Test critical values:	1% level	-4.090602	
	5% level	-3.473447	
	10% level	-3.163967	

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

Null Hypothesis: D(LJNJ) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-19.93554	0.0001
Test critical values:	1% level	-4.090602	
	5% level	-3.473447	
	10% level	-3.163967	

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

Question 20. Consider two models for U.S. real GDP, used to construct forecast for the period 2010Q1-2016Q4:

(A) deterministic trend model

$$\log rGDP_t = \beta_0 + \beta_1 t + u_t$$
$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t$$

for which the sequence of 1-step ahead forecasts has RMSE=103.459 and the multistep forecast has RMSE=1649.069

(B) stochastic trend model

$$\Delta \log rGDP_t = \beta_0 + u_t$$
$$u_t = \phi_1 u_{t-1} + \varepsilon_t$$

for which the sequence of 1-step ahead forecasts has RMSE=77.3231 and the multistep forecast has RMSE=905.1898.

Discuss how we would choose which model is preferred based on this information. How would we conduct a formal test that one of the models produces more precise forecasts?