ECO 5375-701 Prof. Tom Fomby
Eco and Bus Forecasting Fall 2016

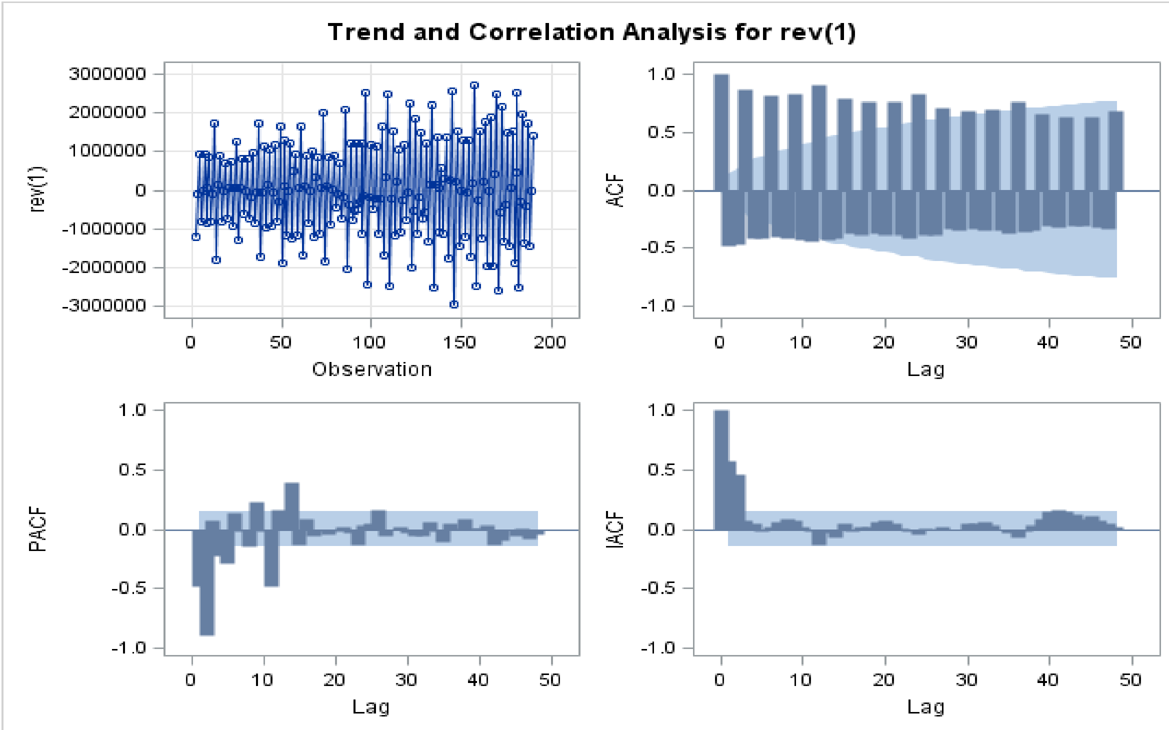
EXERCISE 8
KEY

Purpose: To learn how to use the **Hasza-Fuller (1982)** and **Dickey-Hasza-Fuller (1984) Seasonal Unit Root tests** to determine the appropriate differencing of the Plano Sales Tax Revenue Data set. For more on these tests see the pdf file **Seasonal Differencing.pdf** on the class website. **You are to turn in this homework on Tuesday, November 9.**

Go to the course website and download the SAS program **Plano_Unit_2.sas** program and use it to complete the following tasks:

(i) Examine the ACF of the first difference of the Plano Sales Tax Revenue. Does it indicate the presence of seasonality in the data? Explain your answer. (Note we differenced the data because the data has trend in it. If it had not had trend in it, we would have simply looked at the ACF of the original data.) Hand in the ACF with this exercise.

Answer: The ACF of the first differences of the data has very slowly damping autocorrelations at the seasonal lags of 12, 24, 36, and 48, thus hinting that there is substantial seasonal variation in the data.



(ii) Use the SAS program **Plano_Unit_2.sas** to conduct the Hasza-Fuller (1982) test of the appropriateness of using the first and seasonal span difference to achieve stationarity in the Plano Sales Tax Revenue data. What is the null hypothesis of this test? What is the alternative hypothesis of this test? What conclusion do you draw from conducting this test?

Answer:

The relevant documents are Seasonal Differencing.pdf and Seasonal Unit Root Test Tables.pdf. You can find these on the class website.

The test equation for the Hasza-Fuller (1982) test is:

$$y_t = \beta_1 y_{t-1} + \beta_2 (y_{t-1} - y_{t-S-1}) + \beta_3 (y_{t-S} - y_{t-S-1}) + \gamma_1 \Delta_S \Delta_1 y_{t-1} + \cdots + \gamma_p \Delta_S \Delta_1 y_{t-p} + a_t$$

where for monthly data S = 12.

The null and alternative hypotheses are

$H_0 : \Delta_S \Delta_1 y_t$ is the appropriate transformation to stationarity.

(i.e. $\beta_1 = 1, \beta_2 = 0, \beta_3 = 1$)

$H_1 : \Delta_S \Delta_1 y_t$ is not the appropriate transformation to stationarity.

```
Hasza-Fuller Test for (1,12) Differencing
```

The REG Procedure
Model: MODEL1

Test 1 Results for Dependent Variable rev

Source	DF	Mean Square	F Value	Pr > F
Numerator	3	0.12152	1.91	0.1300
Denominator	166	0.06363		

The appropriate critical value at the 5% level is 14.78. (See the document Seasonal Unit Root Test Tables.pdf.) The observed F-statistic of 1.91 is less than this critical value, thus we accept the null hypothesis that the (1,12) differencing is appropriate for the data. (Notice that you cannot use the p-value reported for the F-Statistic here because in this test situation the usual F-distribution is not appropriate.)

(iii) Separately, using **Plano_unit_2.sas**, conduct the Dickey-Hasza-Fuller (1984) test of the appropriateness of using the seasonal span difference alone to achieve stationarity in the Plano Sales Tax Revenue data. What is the null hypothesis of this test? What is the alternative hypothesis of this test? What arethe 1%, 5%, and 10% critical values for this test? What conclusion do you draw from this test?

Answer:

The test equation for the Dickey-Hasza-Fuller (1984) test is:

$$y_t - y_{t-S} = \beta_S y_{t-S} + \gamma_1 \Delta_S y_{t-1} + \cdots + \gamma_p \Delta_S y_{t-p} + a_t$$

where S = 12 in this case (monthly data).

The null and alternative hypotheses are

$H_0 : \Delta_S y_t$ **is the appropriate transformation to stationarity (i.e.** $\beta_S = 0$**)**

$H_1 : \Delta_S y_t$ **is not the appropriate transformation to stationarity**

The estimated test equation is

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|

Intercept	1	118313	55099	2.15	0.0332
rev_12	1	-16922	16178	-1.05	0.2971
rev12_1	1	0.29926	0.07692	3.89	0.0001
rev12_2	1	0.00179	0.07567	0.02	0.9812
rev12_3	1	0.36066	0.07521	4.80	<.0001
rev12_4	1	0.00488	0.07636	0.06	0.9491

The observed t-statistic associated with the variable rev_12 is -1.05. The appropriate critical value at the 5% level is -2.02. (Again see the document Seasonal Unit Root Test Tables.pdf.) Therefore, the null hypothesis that the seasonal span difference is needed is supported by the data. (Notice that you cannot use the p-value reported for the t-statistic of rev_12 here because in this test situation the usual t-distribution is not appropriate.)

(iv) Finally, given the results you derived from the above tests, what is your choice of seasonal filter for the Plano Sales Tax Revenue data? $\Delta_1\Delta_{12}$ or Δ_{12} ?

Answer:

The result of the Hasza-Fuller test indicated that the (1,12) differencing is supported by the data. Therefore, we don't need to conduct the Dickey-Hasza-Fuller. We did it here just as an illustration of what one might do if the Hasza-Fuller test was rejected and we needed to go fishing for a different transformation, i.e. seasonal span differencing. For the Plano data, apparently $\Delta_S\Delta_1$ is an appropriate choice of seasonal filter to render the Plano data stationary.

(v) Given the results produced by the **Plano_unit._2.sas** program, write out the "best" Multiplicative Seasonal Box-Jenkins $ARIMA(P,D,Q)_S x(p,d,q)$, with coefficient estimates and standard errors, etc. when using the $\Delta_1\Delta_{12}$ transformation? Using the backshift polynomial form to write out your estimated model would probably be the best way to go. Fill in the following blanks: P = **__(24)_**, D = **__1_**, Q = **_1__**, p = **_0__**, d = **_1__**, q = **_3__** .

Answer:

$$\left(1+\underset{(0.08369)}{0.20597}B^{24}\right)\Delta_1\Delta_{12}rev_t = -0.00155 + \left(1-\underset{(0.07122)}{0.62574}B^1-\underset{(0.08344)}{0.27826}B^2+\underset{(0.07148)}{0.36875}B^3\right)\left(1-\underset{(0.07402)}{0.49117}B^{12}\right)\hat{a}_t$$

AIC = -17.4905, SBC = 1.566441, $Q_{24} = 16.95\ (0.5932)$

All coefficients are statistically significant and the residuals of the model are white noise.

(vi) Given the results produced by the **Plano_unit._2.sas** program, write out the "best" Multiplicative Seasonal Box-Jenkins $ARIMA(P,D,Q)_s x(p,d,q)$, with coefficient estimates and standard errors, etc. when using the Δ_{12} transformation? Using the backshift polynomial form to write out your estimated model would probably be the best way to go. Fill in the following blanks: P = _**0**_, D = _**1**_, Q = _**1**_, p = _**(1,3,5)**_, d = _**0**__, q = _**0**_ .

Answer:

$$\left(\mathbf{1} - \mathbf{0.21645}\boldsymbol{B}^{\mathbf{1}} - \mathbf{0.43541}\boldsymbol{B}^{\mathbf{3}} - \mathbf{0.18435}\boldsymbol{B}^{\mathbf{5}}\right)\Delta_{\mathbf{12}}\boldsymbol{rev_t} = \mathbf{0.035956} +$$

(0.06475) (0.06644) (0.06499)

$$(\mathbf{1} - \mathbf{0.61971}\boldsymbol{B}^{\mathbf{12}})\boldsymbol{\hat{a}_t}$$

(0.06495)

AIC = -32.4861, SBC = -16.5771, $\boldsymbol{Q_{24}} = \mathbf{20.49}\ (\mathbf{0.4274})$

(vii) As the Hasza-Fuller and Dickey-Hasza-Fuller tests provided contradictory results, one has to rely on other means to determine which model to forecast the 2006 Sales Tax Revenue for the City of Plano. The **Plano_unit._2.sas** program provides such a method. Describe to me this method and what conclusion you draw from the method that is used. That is, which model should you use and why?

Answer:

Since the models are non-nested we cannot compare the AIC and SBC goodness-of-fit measures across the two estimated models. (They have different dependent variables, i.e. $\Delta_{\mathbf{1}}\Delta_{\mathbf{12}}\boldsymbol{rev_t}$ **versus** $\Delta_{\mathbf{12}}\boldsymbol{rev_t}$.) **But instead we can compare the correlations between the actual and fitted values offered by the competing models. See the output below:**

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	rev	**forecast_112**	**forecast_12**
rev	1.00000	0.98356	0.98483
		<.0001	<.0001
	177	177	177
forecast_112	0.98356	1.00000	0.99759
	<.0001		<.0001

	177	190	190
forecast_12	0.98483	0.99759	1.00000
	<.0001	<.0001	
	177	190	190

As we can see corr(rev, forecast_112) = 0.98356 < corr(rev, forecast_12) = 0.98483. Therefore, the $\Delta_{12} rev_t$ model is slightly preferred. Hereafter, we use this model to do our projection of the growth in Plano Sales Tax Revenue data from 2005 to 2006.

(viii) Given the forecasts from your preferred model of the Plano Sales Tax Revenue data, I want you to fill in the following blanks:

Total Tax Revenue for Plano in 2005 (including the December 2005 forecast) = ________**52,577,324**________________.

Total Forecasted Tax Revenue for Plano in 2006 = ___**55,495,100**__________________.

The percentage increase in Tax Revenue that is forecasted for 2006 as compared to 2005 = __**5.55%**______%.

Answer: The forecasts from the $\Delta_{12} rev_t$ model are (in millions):

Forecasts for variable rev

Obs	Forecast	Std Error	95% Confidence Limits	
191	3.6392	0.2178	3.2122	4.0661
192	4.0425	0.2229	3.6056	4.4793
193	6.5158	0.2231	6.0785	6.9530
194	3.9830	0.2433	3.5062	4.4599
195	3.7049	0.2468	3.2212	4.1887
196	5.5338	0.2526	5.0387	6.0288
197	4.2536	0.2602	3.7436	4.7636
198	4.0472	0.2624	3.5330	4.5614
199	5.5738	0.2668	5.0509	6.0967
200	4.1580	0.2706	3.6276	4.6884
201	4.1589	0.2728	3.6243	4.6936
202	5.6377	0.2757	5.0973	6.1781

Forecasts for variable rev

Obs	Forecast	Std Error	95% Confidence Limits	
203	3.8859	0.2997	3.2985	4.4733