ECON 423 - AR regression and Forecast Evaluation Lab

Introduction

This lab focuses on generating AR forecasts and the evaluation of forecasts in the context of forecasting demand for tickets at Major League Baseball games. In this lab you will estimate a AR regression model for attendance OLS and use the estimated parameters from this to generate *ex post* and *ex ante* forecasts of attendance. These forecasts be evaluated and compared to alternative naive and multiple regression forecasts.

Goals

- 1. Estimate an AR forecasting model
- 2. Understand how to calculate and interpret forecast evaluation measures
- 3. Compare naive forecasts, multivariate regression forecasts, and AR regression forecasts

Data

The Excel file forecast_eval_1_f03_data.xls contains data on attendance, ticket prices and season wins for the Kansas City Royals professional baseball team. The file contains the following variables:

Variable	Description
year	Calendar year, 1990-2001
teamname	Name of baseball team
avg_attend	Average attendance per game for season
price	Average ticket price
wins	Total number of games won in season

This file contains the following worksheets:

regression data	Average attendance, average price and total wins 1990-2001
regression output	Regression output from previous lab
AR data	Average attendance 1969-2001
forecast evaluation	Actual and regression forecast attendance

Methods

Recall that the general form of an Autoregressive (AR) forecast regression model is

$$X_{t} = \beta_{0} + \beta_{1}X_{t-1} + \beta_{2}X_{t-2} + \ldots + \beta_{k}X_{t-k} + u_{t}$$

Also, recall that the four primary forecast diagnostic statistics discussed in class are

1. Mean Absolute Error (MAE)

$$MAE = \sum_{1}^{n} \frac{|FE_t|}{n} = \sum_{1}^{n} \frac{|F_t - A_t|}{n}$$

2. Mean Absolute Percentage Error (MAPE)

$$MAPE = \sum_{1}^{n} \frac{|FE_{t}|}{n} = \sum_{1}^{n} \left| \frac{F_{t} - A_{t}}{A_{t-1}} \right|$$

3. Mean Square Error (MSE)

$$MSE = \sum_{1}^{n} \frac{(FE_t)^2}{n} = \sum_{1}^{n} \frac{(F_t - A_t)^2}{n}$$

4. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\sum_{1}^{n} \frac{(FE_t)^2}{n}} = \sqrt{\sum_{1}^{n} \frac{(F_t - A_t)^2}{n}}$$

Part 1 - AR Forecasts

- 1. Plot the average annual attendance for the Royals over the period 1969-2000.
- 2. Using attendance data for the period 1969-2000 estimate the parameters of an AR(4) autoregressive regression model

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + e_t$$

- Start with worksheet AR data
- Label to four columns to the right of average attendance L1, L2, L3,L4. These are the "lagged" variables that make up the explanatory variables in the regression model.
- What is the first year you can estimate this model for?
- Should you use the entire series? Why or why not?
- Estimate the regression model. Put the results on a new worksheet and check the **Residuals** option.
- The results for the entire series are

SUMMARY OUT	ſPUT					
Regression	Sta	tistics				
Multiple R		0.858				
R Square		0.736				
Adjusted R Square 0.692						
Standard En	ror	2731				
Observation	າຣ	29				
ANOVA						
	df	SS	MS	F		
Regression	4	499851585	124962896	16.8		
Residual	24	179008072	7458670			
Total	28	678859657				
	Co	efficients	Standard	Error	t Stat	P-value
Intercept		6419		2237	2.869	0.008
X Variable	1	0.752		0.189	3.974	0.001
X Variable	2	0.103		0.240	0.431	0.671
X Variable	3	0.075		0.235	0.320	0.752
X Variable	4	-0.200		0.173	-1.154	0.260

- 3. Discuss the significance of the estimated regression parameters.
- 4. Plot the actual and predicted values for the AR(4) model over the period 1990-2000. How well does the forecast track the data?

Part 2 - Forecast Evaluation

- 1. Begin in the worksheet forecast evaluation; column A is the year, column B contains actual average attendance and column C contains the *ex post* regression forecast.
- 2. In column D calculate the forecast error for the regression forecast
- 3. In column E calculate the absolute value of the forecast error for the regression forecast
- 4. In column F calculate the square of the forecast error for the regression forecast
- 5. Copy the AR(4) forecast into column G
- 6. In column H calculate the forecast error for the AR(4) forecast
- 7. In column I calculate the absolute value of the forecast error for the AR(4) forecast
- 8. In column J calculate the square of the forecast error for the AR(4) forecast
- 9. Calculate the Mean Absolute Error, Mean Square Error and Root Mean Square Error for each forecast
- 10. Which forecast performed better within sample?