

Practical 4: Further Time Series and Forecasting

What you need to do

Before starting the Minitab practical, it would be a good idea (if you've not already done so) to bookmark the webpage containing all of the information for Semester 2 of ACE2013:

www.mas.ncl.ac.uk/~nlf8/teaching/ace2013

To add this page to your favourites, select Favourites → Add to Favourites from the menu bar at the top. You will now be able to get to this page easily by selecting Favourites and selecting the page from the list.

In this practical you will be using Minitab to perform time series analyses similar to those shown in Chapter 2. You should work through *all* the questions given below – *all* parts of this practical sheet are assessed! **It will really help if you have the worksheet for Practical 3 at hand – very little help is given with the first few questions, and I will assume you have already worked through Practical 3.**

Along with assessed work from the other practical sessions, solutions to *all* questions in this practical should be typed up and submitted (with a cover sheet – see the course website) to the homework submission postbox outside the Maths & Statistics General Office, on the third floor of the Herschel Building, by no later than **4pm, Monday 12th May 2014.**

Log into a PC and open Minitab by clicking on Start → All programs → Statistical Software → Minitab 16 Statistical Software English.

Note: you should open any data files from the course webpage from within the FireFox web browser; or right-click the data file link, save the file somewhere convenient as filename.MTW and then open from within Minitab itself.

Go to the ACE2013 website and click on the **Jungle Village** link in the material for Practical 4. Once opened, you will see that there are data in the first four columns of this worksheet.

- Column 1 shows Y , the revenue (in thousands of Australian dollars), for *PK's Jungle Village*, a hostel in Cape Tribulation in Northern Queensland, Australia.
- Column 2 indicates the “season” – 1 being the first quarter (January–March), 2 being the second quarter (April–June), 3 being the 3rd quarter (July–September) and 4 being the last quarter (October–December)
- Column 3 indicates the year
- Column 4 is a time index from 1 to 39

Look at the worksheet, and then answer the following questions.

- *1. Produce an appropriately labelled time series plot for these data. Copy–and–paste this into your solutions, and comment.
- *2. Use **Minitab** to find the series of centred moving averages Y^* . Make sure **Minitab** stores the resulting moving averages in a column of the worksheet, and re–label this column “MA”. There is nothing to write up here.
- *3. Now you should use **Minitab** to fit a simple linear regression model to the series of moving averages, again making sure the fitted values are stored in a column of the worksheet (re–label this column “Trend”). In your solutions, you should include the regression table produced by **Minitab**, and you should also write down the regression equation in the form:

$$Y^* = \beta_0 + \beta_1 T + \epsilon, \quad \epsilon \sim N(0, \sigma^2).$$

Is the slope term β_1 significant here? Justify your answer.

- *4. Now re–produce the time series plot from question 1 but with the moving averages and the regression line superimposed. You should copy–and–paste this plot into your solutions.
- *5. Use **Minitab** to obtain the seasonal effects, and copy–and–paste the relevant output into your solutions. Make sure you get **Minitab** to store these in a column of the worksheet, and re–label that column “Season”. *[Notice that the first observation is in the third quarter (“season 3”). Thus, in Stat→Time Series→Decomposition, you should select Options and tell Minitab that the First obs. is in seasonal period 3]*
- *6. You should now obtain the residual series for the revenue data. Click on **Editor** and select **Enable Commands**. Now go to the Session window (the white window above the worksheet) and at the **Minitab** command prompt (MTB >) type

MTB > LET C10 = C1-Trend-Season

Label column C10 as “Residuals”. Produce a time series plot of the residual series, and copy–and–paste this into your solutions.

- *7. Use **Minitab** to obtain a plot of the partial autocorrelation function for the residuals by clicking on Stat→Time Series→Partial Autocorrelation and entering column C10 in Series. Click OK, and copy–and–paste this plot into your solutions.
- *8. Using your plot in question 7, what autoregressive model would you suggest for the residual series? (i.e. $AR(1)$? $AR(2)$? $AR(3)$?)
- *9 Use the options Stat→Time Series→ARIMA to fit the model you proposed in question 8. You should copy–and–paste the “Final Estimates of Parameters” table into your solutions (nothing else).
- *10 Using the output you obtained in question 9, write down your estimated model in the form

$$Y_t = \mu + \alpha_1(Y_{t-1} - \mu) + \dots + \epsilon_t.$$

- *11. Repeat question 9, but before clicking OK within the ARIMA option, click **Forecasts**. In the box that says **Lead** enter the number 5, and then click OK and OK again. You will get exactly the same output as in question 9, but now you will also get the forecasted values for the model you wrote down in question 10 for the next 5 time–points. Copy–and–paste these forecasted values, with their lower and upper 95% confidence limits, into your solutions.

- *12.** You should now forecast revenue at *PK's Jungle Village* in April–June 2010 (time–point 40).
- (i) *By hand*, use the trend equation you obtained in question 3 to obtain an estimate of the trend at time–point 40.
 - (ii) Add the estimated trend from part (i), and the seasonal effect for April–June from question 5, to the forecasted value for time–point 40 in question 11. This is your *full forecast* of revenue, in Australian dollars, for April–June 2010.
 - (iii) Add the estimated trend from part (i), and the seasonal effect for April–June, to the lower and upper 95% confidence limits to obtain a 95% confidence interval for forecasted revenue in April–June 2010. You should write this down in the usual way for confidence intervals, i.e. (lower bound, upper bound) – and don't forget your units! (Australian dollars)
- *13.** Repeat question 12 to forecast revenue at *PK's Jungle Village* in July–September 2010, and obtain a 95% confidence interval for this forecast.
- *14.** Repeat question 9, but before clicking OK select **Graphs** and tick the boxes that say **Histogram of residuals** and **Residuals versus fits**. Click OK and OK again. You will get the same output as before, but you will now also get two graphs. Copy–and–paste these graphs into your solutions, and comment (see pages 28–29 and page 82 of the lecture notes for help with this).