# Chapter 6: Outline Solutions

The ARIMA models developed in these notes were obtained using Minitab. Other programs may give slightly different results. In each case, a complete solution should involve plotting the data and computing the ACF (and possibly the PACF). We include only the output and residual diagnostics for the final model selected. The simplest model that satisfies basic diagnostics is reported; other models may well be selected that perform adequately.

For comparative purposes in Exercises 6.1-6.4, the one-step-ahead forecasts are used for MA(3), MA(7) and the identified model. The comparative tables are provided in *Exercises\_6.1\_6.4.xlsx*.

### 6.1

An AR(1) model is adequate for this series (1963 – 2011).

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | 0.638 | 0.115 | 5.56 | 0.000 |
| Constant | 2.442 | 0.330 | 7.41 | 0.000 |
| Mean | 6.740 | 0.910 |  |  |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 47 | 249.054 | 5.29903 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 7.74 | 15.81 | 34.58 | 45.72 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.654 | 0.825 | 0.440 | 0.484 |



|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| GDP Change | |  |  |  |  |  |  |  |  |  |
|  | Actual | MA3 | MA7 | AR(1) | Absolute errors | | | Squared Errors | | |
| 2011 | 3.7 |  |  |  | MA3 | MA7 | AR(1) | MA3 | MA7 | AR(1) |
| 2012 | 4.1 | 1.83 | 3.46 | 4.80 | 2.27 | 0.64 | 0.70 | 5.14 | 0.41 | 0.49 |
| 2013 | 3.3 | 3.87 | 3.09 | 5.06 | 0.57 | 0.21 | 1.76 | 0.32 | 0.05 | 3.09 |
| 2014 | 4.2 | 3.70 | 2.73 | 4.55 | 0.50 | 1.47 | 0.35 | 0.25 | 2.17 | 0.12 |
| 2015 | 3.7 | 3.87 | 2.69 | 5.12 | 0.17 | 1.01 | 1.42 | 0.03 | 1.03 | 2.02 |
|  |  |  |  | MAE | 0.88 | 0.84 | 1.06 | 1.43 | 0.91 | 1.43 |
|  |  |  |  |  |  |  | RMSE | 1.20 | 0.96 | 1.20 |

MA(7) does best for the hold-out sample

### 6.2

An AR(2) model is adequate for this series (1963-2011), although there is a large outlier for 2009 (recession).

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | 1.056 | 0.140 | 7.55 | 0.000 |
| AR   2 | -0.322 | 0.141 | -2.29 | 0.027 |
| Constant | 1.098 | 0.236 | 4.65 | 0.000 |
| Mean | 4.129 | 0.888 |  |  |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 46 | 125.606 | 2.73056 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 10.63 | 15.34 | 29.37 | 34.94 |
| DF | 9 | 21 | 33 | 45 |
| P-Value | 0.302 | 0.805 | 0.648 | 0.860 |



|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| CPI Change | |  |  |  |  |  |  |  |  |  |
|  | Actual | MA3 | MA7 | AR(2) | Absolute errors | | | Squared Errors | | |
| 2010 | 1.64 |  |  |  | MA3 | MA7 | AR(2) | MA3 | MA7 | AR(2) |
| 2011 | 3.16 |  |  |  |  |  |  |  |  |  |
| 2012 | 2.07 | 1.48 | 2.54 | 3.90 | 0.59 | 0.47 | 1.83 | 0.35 | 0.22 | 3.36 |
| 2013 | 1.46 | 2.29 | 2.35 | 2.27 | 0.82 | 0.88 | 0.80 | 0.68 | 0.78 | 0.64 |
| 2014 | 1.62 | 2.23 | 2.10 | 1.98 | 0.61 | 0.47 | 0.36 | 0.37 | 0.22 | 0.13 |
| 2015 | 0.12 | 1.72 | 1.92 | 2.34 | 1.60 | 1.80 | 2.22 | 2.56 | 3.24 | 4.93 |
|  |  |  |  | MAE | 0.91 | 0.91 | 1.30 | 0.99 | 1.12 | 2.27 |
|  |  |  |  |  |  |  | RMSE | 0.99 | 1.06 | 1.51 |

MA(3) does best for the hold-out sample

### 6.3

The series appears to be a random walk, but this is as the result of the huge outlier in 2009.

Final Estimates of Parameters

Series differenced once

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| MA   1 | 0.031 | 0.203 | 0.15 | 0.878 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 32 | 3842.82 | 120.088 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 6.54 | 15.59 | \* | \* |
| DF | 11 | 23 | \* | \* |
| P-Value | 0.835 | 0.872 | \* | \* |



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | SA Oil |  |  |  |  |  |  |  |  |  |  | |  | Actual | MA3 | MA7 | RW | Absolute errors | | | Squared Errors | | | |  |  |  |  |  | MA3 | MA7 | AR(1) | MA3 | MA7 | RW | | 2011 | 109.43 |  |  |  |  |  |  |  |  |  | | 2012 | 108.93 | 84.06 | 75.94 | 109.43 | 24.87 | 32.99 | 0.50 | 618.52 | 1088.06 | 0.25 | | 2013 | 103.44 | 99.34 | 83.95 | 108.93 | 4.10 | 19.49 | 5.49 | 16.84 | 380.03 | 30.14 | | 2014 | 96.83 | 107.27 | 90.08 | 103.44 | 10.44 | 6.75 | 6.61 | 108.92 | 45.58 | 43.69 | | 2015 | 50.00 | 103.07 | 93.68 | 96.83 | 53.07 | 43.68 | 46.83 | 2816.07 | 1907.94 | 2193.05 | |  |  |  |  | MAE | 23.12 | 25.73 | 14.86 | 890.09 | 855.40 | 566.78 | |  |  |  |  |  |  |  | RMSE | 29.83 | 29.25 | 23.81 |   The random walk is superior to the MA schemes. 6.4 The original series clearly shows increasing variation over time (1963-2011) so a log transform is in order. Differencing is required to account for the trend: an ARIMA(1,1,0)+C model is reasonable. |  |  |  |  |

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | 0.358 | 0.138 | 2.59 | 0.013 |
| Constant | 0.0433 | 0.0214 | 2.02 | 0.049 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 46 | 1.01345 | 0.0220315 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 7.56 | 24.86 | 28.14 | \* |
| DF | 10 | 22 | 34 | \* |
| P-Value | 0.672 | 0.304 | 0.750 | \* |
|  |  |  |  |  |



|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Dulles |  |  |  |  |  |  |  |  |  |  |  |
|  | Actual | LN (Act) | MA3 | MA7 | ARIMA(1,1,0)+C: log transform | Absolute errors | | | Squared Errors | | |
| 2010 | 17214 | 9.75 |  |  |  | MA3 | MA7 | ARIMA | MA3 | MA7 | ARIMA |
| 2011 | 16663 | 9.72 |  |  |  |  |  |  |  |  |  |
| 2012 | 15883 | 9.67 | 16947 | 18170 | 17199 | 1064 | 2287 | 1316 | 1132096 | 5228409 | 1731534 |
| 2013 | 14958 | 9.61 | 16587 | 17277 | 16304 | 1629 | 2319 | 1346 | 2652555 | 5379086 | 1810658 |
| 2014 | 14393 | 9.57 | 15835 | 16873 | 15288 | 1442 | 2480 | 895 | 2078403 | 6151109 | 800942 |
| 2015 | 14463 | 9.58 | 15078 | 16245 | 14824 | 615 | 1782 | 361 | 378225 | 3174506 | 130428 |
|  |  |  |  |  | MAE | 1187.3 | 2216.9 | 979.4 | 1560320 | 4983277 | 1118390 |
|  |  |  |  |  |  |  |  | RMSE | 1249.1 | 2232.3 | 1057.5 |

The AR model with the log transform is best.

### 6.5

The parameter estimates are; the negative signs hint at over-differencing.

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | -0.7487 | 0.0807 | -9.28 | 0.000 |
| AR   2 | -0.3010 | 0.0807 | -3.73 | 0.000 |

The LBP results are:

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 22.75 | 31.64 | 42.02 | 55.78 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.012 | 0.084 | 0.162 | 0.153 |
|  |  |  |  |  |

The residual sum of squares is:

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 140 | 2551.17 | 18.2226 |

*Back forecasts excluded*

This model is clearly inferior to the simpler model in the next exercise.



### 6.6

The model ARIMA(0,1,1)+C works OK for the log version; again the MA term is not significant. The 10/2001 outlier appears to be even more extreme.

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| MA   1 | 0.0777 | 0.0840 | 0.92 | 0.357 |
| Constant | 0.002583 | 0.000993 | 2.60 | 0.010 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 141 | 0.0233748 | 0.0001658 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 8.76 | 15.75 | 25.93 | 35.25 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.555 | 0.828 | 0.838 | 0.875 |



Adjusting the outlier (e.g. replacing it by the average of the values before and after it) serves to reduce the RMSE and makes the MA coefficient even smaller. The Ljung-Box tests are not materially affected.

6.7

The results are very similar. When the outlier is adjusted the MA term is very close to zero and the diagnostic tests produce very similar results.

### 6.8

As noted earlier the RMSE is reduced and the MA(1) coefficient becomes insignificant.

**ARIMA Model: LN\_Sales, with outlier (Oct 2001) replaced by average of Sept and Nov.**

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| MA   1 | -0.0214 | 0.0842 | -0.25 | 0.800 |
| Constant | 0.002583 | 0.000980 | 2.63 | 0.009 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 141 | 0.0185805 | 0.0001318 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 10.41 | 20.28 | 28.51 | 38.88 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.405 | 0.566 | 0.733 | 0.763 |



### 6.9

Note that the series is monthly and not quarterly as stated in the exercise.

The forecasts for the 36 months from 12/2012 as origin are:

Forecasts from period 144

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | 95% Limits | |  |
| Period | Forecast | Lower | Upper | Actual |
| 145 | 366.424 | 359.440 | 373.407 | 367.676 |
| 146 | 367.224 | 357.498 | 376.949 | 372.573 |
| 147 | 368.024 | 356.175 | 379.873 | 369.575 |
| 148 | 368.824 | 355.178 | 382.471 | 368.614 |
| 149 | 369.625 | 354.392 | 384.857 | 370.671 |
| 150 | 370.425 | 353.756 | 387.094 | 372.894 |
| 151 | 371.225 | 353.234 | 389.216 | 374.280 |
| 152 | 372.025 | 352.803 | 391.248 | 372.768 |
| 153 | 372.826 | 352.446 | 393.205 | 373.610 |
| 154 | 373.626 | 352.152 | 395.100 | 374.196 |
| 155 | 374.426 | 351.910 | 396.942 | 374.691 |
| 156 | 375.226 | 351.715 | 398.738 | 377.425 |
| 157 | 376.027 | 351.560 | 400.493 | 373.036 |
| 158 | 376.827 | 351.441 | 402.212 | 377.666 |
| 159 | 377.627 | 351.355 | 403.899 | 382.356 |
| 160 | 378.427 | 351.297 | 405.558 | 385.855 |
| 161 | 379.228 | 351.265 | 407.190 | 385.572 |
| 162 | 380.028 | 351.258 | 408.797 | 388.226 |
| 163 | 380.828 | 351.273 | 410.383 | 387.541 |
| 164 | 381.628 | 351.308 | 411.949 | 389.924 |
| 165 | 382.429 | 351.362 | 413.495 | 388.478 |
| 166 | 383.229 | 351.433 | 415.025 | 390.183 |
| 167 | 384.029 | 351.521 | 416.537 | 391.799 |
| 168 | 384.829 | 351.624 | 418.035 | 388.116 |
| 169 | 385.630 | 351.741 | 419.518 | 384.417 |
| 170 | 386.430 | 351.872 | 420.988 | 383.423 |
| 171 | 387.230 | 352.015 | 422.445 | 387.665 |
| 172 | 388.030 | 352.171 | 423.890 | 387.235 |
| 173 | 388.831 | 352.338 | 425.323 | 392.268 |
| 174 | 389.631 | 352.516 | 426.746 | 391.955 |
| 175 | 390.431 | 352.704 | 428.158 | 395.100 |
| 176 | 391.231 | 352.901 | 429.561 | 394.944 |
| 177 | 392.032 | 353.109 | 430.954 | 394.429 |
| 178 | 392.832 | 353.325 | 432.339 | 394.200 |
| 179 | 393.632 | 353.549 | 433.715 | 395.261 |
| 180 | 394.432 | 353.782 | 435.083 | 395.950 |

Using the log transform the results are:

Forecasts from period 144

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | 95% Limits | |  |
| Period | Forecast | Lower | Upper | Actual |
| 145 | 5.90394 | 5.88071 | 5.92716 | 5.90720 |
| 146 | 5.90645 | 5.87487 | 5.93803 | 5.92043 |
| 147 | 5.90896 | 5.87081 | 5.94710 | 5.91235 |
| 148 | 5.91147 | 5.86773 | 5.95520 | 5.90975 |
| 149 | 5.91398 | 5.86529 | 5.96267 | 5.91531 |
| 150 | 5.91649 | 5.86331 | 5.96967 | 5.92129 |
| 151 | 5.91900 | 5.86167 | 5.97632 | 5.92500 |
| 152 | 5.92151 | 5.86032 | 5.98269 | 5.92096 |
| 153 | 5.92402 | 5.85920 | 5.98884 | 5.92321 |
| 154 | 5.92653 | 5.85827 | 5.99479 | 5.92478 |
| 155 | 5.92904 | 5.85750 | 6.00057 | 5.92610 |
| 156 | 5.93155 | 5.85688 | 6.00621 | 5.93337 |
| 157 | 5.93406 | 5.85639 | 6.01173 | 5.92167 |
| 158 | 5.93657 | 5.85600 | 6.01713 | 5.93401 |
| 159 | 5.93908 | 5.85572 | 6.02243 | 5.94635 |
| 160 | 5.94159 | 5.85553 | 6.02764 | 5.95546 |
| 161 | 5.94410 | 5.85542 | 6.03277 | 5.95473 |
| 162 | 5.94661 | 5.85539 | 6.03783 | 5.96159 |
| 163 | 5.94912 | 5.85542 | 6.04281 | 5.95982 |
| 164 | 5.95163 | 5.85552 | 6.04774 | 5.96595 |
| 165 | 5.95414 | 5.85568 | 6.05260 | 5.96224 |
| 166 | 5.95665 | 5.85589 | 6.05741 | 5.96662 |
| 167 | 5.95916 | 5.85615 | 6.06216 | 5.97075 |
| 168 | 5.96167 | 5.85646 | 6.06687 | 5.96130 |
| 169 | 5.96418 | 5.85682 | 6.07153 | 5.95173 |
| 170 | 5.96669 | 5.85722 | 6.07616 | 5.94914 |
| 171 | 5.96920 | 5.85766 | 6.08074 | 5.96014 |
| 172 | 5.97171 | 5.85813 | 6.08528 | 5.95903 |
| 173 | 5.97422 | 5.85864 | 6.08979 | 5.97195 |
| 174 | 5.97673 | 5.85919 | 6.09426 | 5.97115 |
| 175 | 5.97924 | 5.85977 | 6.09870 | 5.97914 |
| 176 | 5.98175 | 5.86038 | 6.10311 | 5.97874 |
| 177 | 5.98426 | 5.86102 | 6.10750 | 5.97744 |
| 178 | 5.98677 | 5.86168 | 6.11185 | 5.97686 |
| 179 | 5.98928 | 5.86238 | 6.11618 | 5.97955 |
| 180 | 5.99179 | 5.86310 | 6.12048 | 5.98129 |

The summary statistics, after transformation back to original data, are:

|  |  |  |
| --- | --- | --- |
|  | Original | Log Transform |
| MAE | 3.21 | 2.77 |
| RMSE | 4.07 | 3.38 |
| MAPE | 0.83 | 0.72 |

All forecasts lie within the 95% limits, which are rather wide. Adjusting for the outlier narrows the PIs, but has very little effect on the point forecasts.

### 6.10

We start with the model



For the time period (*t* + 2), we can substitute the expression for to obtain



Since is known at the time we make the forecast and δ is a constant, the variance of the one-step-ahead forecast is:



Similarly, the variance of the two-step-ahead forecast is:



Successive errors are assumed to be uncorrelated so that:



The MA(1) model is:



At time t, the previous error term has been recorded and µ is constant so that the variance of the one-step-ahead forecast is:



The two-step-ahead version may be written as:



Both error terms are unknown, but they are uncorrelated so the variance is:



### 6.11

Original results from Table 6.9:

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | 0.7044 | 0.0946 | 7.45 | 0.000 |
| Constant | 0.03037 | 0.00813 | 3.74 | 0.000 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 57 | 0.221881 | 0.0038927 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 20.91 | 42.32 | 52.17 | 58.75 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.022 | 0.006 | 0.024 | 0.098 |

Revised results after dropping first two years

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | 0.533 | 0.121 | 4.42 | 0.000 |
| Constant | 0.03629 | 0.00662 | 5.48 | 0.000 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 49 | 0.109614 | 0.0022370 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 18.93 | 34.82 | 52.41 | 58.28 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.041 | 0.040 | 0.023 | 0.106 |



The AR coefficient is lower and the MSE drops by a third. The diagnostic plots look better as well.

### 6.12

**GDP Change**

The two simple models ARIMA(1,0,0)+C and ARIMA(0,1,1) both fit well. The second one has a lower MSE (4.87 against 5.30).

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| MA   1 | 0.654 | 0.110 | 5.94 | 0.000 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 47 | 229.122 | 4.87493 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 5.25 | 12.93 | 28.78 | \* |
| DF | 11 | 23 | 35 | \* |
| P-Value | 0.919 | 0.953 | 0.762 | \* |

**Forecasts**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Actual | Forecast | Lower | Upper |
| 2012 | 4.1 | 2.99 | -1.34 | 7.32 |
| 2013 | 3.3 | 2.99 | -1.59 | 7.57 |
| 2014 | 4.2 | 2.99 | -1.83 | 7.81 |
| 2015 | 3.7 | 2.99 | -2.06 | 8.03 |

**CPI Change**

The two simple models ARIMA(1,0,0)+C and ARIMA(0,1,1) both fit well. The first one has a lower MSE (2.97 against 3.05).

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | 0.8065 | 0.0873 | 9.24 | 0.000 |
| Constant | 0.759 | 0.247 | 3.07 | 0.004 |
| Mean | 3.92 | 1.28 |  |  |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 47 | 139.433 | 2.96666 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 19.82 | 23.44 | 37.05 | 41.50 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.031 | 0.377 | 0.330 | 0.661 |

**Forecasts**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Actual | Forecast | Lower | Upper |
| 2012 | 2.07 | 3.30 | -0.07 | 6.68 |
| 2013 | 1.46 | 3.42 | -0.91 | 7.76 |
| 2014 | 1.62 | 3.52 | -1.34 | 8.38 |
| 2015 | 0.12 | 3.60 | -1.58 | 8.77 |

**SA Oil Prices**

The series appears to be a random walk. The standard deviation of the series is 10.56. The prediction interval for a random walk h steps ahead is “*forecast +/- 1.96\*SD\*sqrt(h)*”.

Even the very wide PI fails to capture the huge drop in 2015.

**Forecasts**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Actual | Forecast | Lower | Upper |
| 2012 | 108.93 | 109.43 | 88.73 | 130.13 |
| 2013 | 103.44 | 109.43 | 80.16 | 138.70 |
| 2014 | 96.83 | 109.43 | 73.58 | 145.28 |
| 2015 | 50 | 109.43 | 68.03 | 150.83 |

**Dulles**

The series has a major step in 1985 (discussed earlier) so only observations from 1986 on are used. The ARIMA(0,1,1) model provides an adequate fit. The prediction intervals are very wide; the series was more volatile in previous years.



**Forecasts**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Actual | Forecast | Lower | Upper |
| 2012 | 15883 | 16616 | 12774 | 20458 |
| 2013 | 14958 | 16616 | 10955 | 22276 |
| 2014 | 14393 | 16616 | 9593 | 23639 |
| 2015 | 14463 | 16616 | 8455 | 24777 |

### 6.13

The two fitted models (to Dec 2014) are shown below; they perform similarly. On closer examination, the ARIMA(0,2,2) model is over-differenced – check the coefficients. An ARIMA(0,1,1)+C does just as well. The forecasts and prediction intervals are very similar.

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | -0.0416 | 0.0781 | -0.53 | 0.596 |
| Constant | 0.846 | 0.281 | 3.01 | 0.003 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 165 | 2179.44 | 13.2088 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 9.61 | 17.18 | 27.50 | 40.91 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.475 | 0.754 | 0.777 | 0.685 |

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| MA   1 | 1.07964 | 0.00009 | 11596.05 | 0.000 |
| MA   2 | -0.0870 | 0.0133 | -6.52 | 0.000 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 164 | 2205.34 | 13.4472 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 10.02 | 17.59 | 27.90 | 41.69 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.439 | 0.730 | 0.760 | 0.653 |

Index 1 = ARIMA(1,1,0)+C

Index 2 = ARIMA(0,2,2)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ACTUAL | FORE\_1 | LOWER\_1 | UPPER\_1 | FORE\_2 | LOWER\_2 | UPPER\_2 |
| 384.4 | 389.1 | 382.0 | 396.2 | 389.4 | 382.2 | 396.6 |
| 383.4 | 389.9 | 380.1 | 399.8 | 390.3 | 380.5 | 400.1 |
| 387.7 | 390.7 | 378.7 | 402.7 | 391.2 | 379.4 | 403.0 |
| 387.2 | 391.5 | 377.7 | 405.4 | 392.1 | 378.5 | 405.7 |
| 392.3 | 392.4 | 376.9 | 407.8 | 393.0 | 377.8 | 408.2 |
| 392.0 | 393.2 | 376.3 | 410.0 | 393.9 | 377.2 | 410.6 |
| 395.1 | 394.0 | 375.8 | 412.2 | 394.8 | 376.8 | 412.8 |
| 394.9 | 394.8 | 375.3 | 414.2 | 395.7 | 376.4 | 415.0 |
| 394.4 | 395.6 | 375.0 | 416.2 | 396.6 | 376.1 | 417.1 |
| 394.2 | 396.4 | 374.7 | 418.1 | 397.5 | 375.8 | 419.2 |
| 395.3 | 397.2 | 374.5 | 420.0 | 398.4 | 375.6 | 421.2 |
| 396.0 | 398.0 | 374.3 | 421.8 | 399.3 | 375.4 | 423.2 |

### 6.14

The MS is slightly lower and the LBP results are better. The forecast results are very similar even though the models look different.

Final Estimates of Parameters

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Coef | SE Coef | T-Value | P-Value |
| AR   1 | -0.348 | 0.121 | -2.88 | 0.005 |
| SAR  4 | -0.347 | 0.125 | -2.78 | 0.007 |

Residual Sums of Squares

|  |  |  |
| --- | --- | --- |
| DF | SS | MS |
| 61 | 58.7947 | 0.963848 |

*Back forecasts excluded*

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lag | 12 | 24 | 36 | 48 |
| Chi-Square | 10.29 | 24.99 | 33.35 | 40.19 |
| DF | 10 | 22 | 34 | 46 |
| P-Value | 0.415 | 0.298 | 0.499 | 0.713 |

Forecasts from period 68

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | 95% Limits | |  |
| Period | Forecast | Lower | Upper | Actual |
| 69 | 100.247 | 98.322 | 102.171 |  |
| 70 | 106.435 | 104.137 | 108.733 |  |
| 71 | 106.444 | 103.706 | 109.182 |  |
| 72 | 124.341 | 121.263 | 127.419 |  |
| 73 | 103.267 | 99.178 | 107.356 |  |
| 74 | 109.352 | 104.687 | 114.016 |  |
| 75 | 109.453 | 104.208 | 114.699 |  |
| 76 | 127.270 | 121.524 | 133.017 |  |



### 6.15

In Exercise 6.10 we considered the *conditional* variance for forecasts one and two steps ahead. Now we examine the *unconditional* values. The model is:



Thus the mean is:



The first step applies because the “mean of the sum is the sum of the means”. Then, since µ is a constant and the errors have zero means, the result follows. Similarly, the variance is given by:



This result follows because the mean is a constant (hence zero variance) and the errors are uncorrelated.

Next, we must consider the lag 1 autocovariance, given by:



The error terms are uncorrelated so



Multiplying out the terms, we find that only one has a non-zero expectation:



Finally, the first order (or lag 1) autocorrelation is:



Following the same approach for higher order autocovariances we find there are no terms with common coefficients, so that all higher terms are zero.

### 6.16 – 6.21

The solutions to Exercises 6.16 – 6.21 were generated using Minitab. That program does not provide one-step-ahead forecasts out of sample. Thus, in each case, the one-step-ahead forecasts were generated after fitting the model to the entire series. The tutorial in R for chapter 6 shows how this can be done.

### 6.16

The ACF shows a clear need for differencing. After differencing, the ACF shows spikes at lags 1 and 3 suggesting ARIMA(3,1,0) or ARIMA(0,1,3). ARIMA(1,1,2) was also considered.

The summary statistics indicate that the two original choices perform better. Both AIC and the hold-out point to ARIMA(3,1,0). The diagnostic plots for this model are shown.

|  |  |  |  |
| --- | --- | --- | --- |
|  | ARIMA | | |
|  | (3,1,0) | (0,1,3) | (1,1,2) |
| Parameters | 3 | 3 | 3 |
| MSE | 41108.0 | 42325.4 | 46798.4 |
| AIC | 10.66 | 10.69 | 10.79 |
| Hold-out |  |  |  |
| MAE | 155.7 | 163.6 | 159.8 |
| RMSE | 194.6 | 202.1 | 205.3 |
| MAPE | 3.36 | 3.53 | 3.41 |



### 6.17

The ACF shows a clear need for differencing, possibly both seasonal and regular. After differencing, the ACF shows strong seasonality and action at lag 3. Many models could be tried, but users should avoid being over-elaborate!

The summary statistics indicate that the ARIMA(3,0.0)(0,1,1)+C model performs best. Other options exist. The diagnostic plots for this model are shown.

|  |  |  |  |
| --- | --- | --- | --- |
|  | ARIMA | | |
|  | (011)(011) | (200)(011)+C | (300)(011)+C |
| Parameters | 2 | 4 | 5 |
| MSE | 165876.0 | 165064.0 | 115732.0 |
| AIC | 12.04 | 12.06 | 11.71 |
| Hold-out |  |  |  |
| MAE | 385.5 | 370.2 | 288.8 |
| RMSE | 439.1 | 433.6 | 352.8 |
| MAPE | 3.53 | 3.37 | 2.66 |

 

### 6.18

The ACF shows a clear need for both regular and seasonal differencing. After differencing, the ACF shows spikes at lags 1 and 3 suggesting ARIMA(3,1,0)(0,1,1)12..

The summary statistics indicate that this choice performs quite well but the simpler ARIMA(0,1,1)(0,1,1)12 model is selected on an out-of-sample basis. The diagnostic plots for the original choice are shown.

|  |  |  |  |
| --- | --- | --- | --- |
|  | ARIMA | | |
|  | (013)(011) | (011)(011) | (310)(011) |
| Parameters | 4 | 2 | 4 |
| MSE | 165876.0 | 165064.0 | 115732.0 |
| AIC | 12.06 | 12.04 | 11.70 |
| Hold-out |  |  |  |
| MAE | 5.6 | 5.3 | 5.7 |
| RMSE | 6.8 | 6.5 | 7.0 |
| MAPE | 1.47 | 1.40 | 1.49 |

 

### 6.19

The plot shows the need to transform the series – a log transform was used. The plot then shows strong seasonality and a downward trend. The benchmark ARIMA(0,1,1)(0,1,1)7 model performs well so no further models need to be considered. The diagnostic plots for this model are shown.



 

### 6.20

The ACF of the differenced series shows no structure, as expected.



The GARCH analysis via EViews produces the following output. That is we find a small positive values for the mean level, indicating a positive return over time. The GARCH model has α=0.149 and β =0.778.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: D(SP500,1) | | | |  |
| Method: ML ARCH - Normal distribution (BFGS / Marquardt steps) | | | | |
| Date: 08/31/17 Time: 15:54 | | |  |  |
| Sample (adjusted): 1/04/2012 12/31/2015 | | | |  |
| Included observations: 1005 after adjustments | | | |  |
| Convergence achieved after 20 iterations | | | |  |
| Coefficient covariance computed using outer product of gradients | | | | |
| Presample variance: backcast (parameter = 0.7) | | | | |
| GARCH = C(2) + C(3)\*RESID(-1)^2 + C(4)\*GARCH(-1) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 1.246265 | 0.398977 | 3.123654 | 0.0018 |
|  |  |  |  |  |
|  |  |  |  |  |
|  | Variance Equation | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 15.03354 | 3.535659 | 4.251976 | 0.0000 |
| RESID(-1)^2 | 0.148939 | 0.023776 | 6.264322 | 0.0000 |
| GARCH(-1) | 0.778271 | 0.031372 | 24.80802 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | -0.001133 | Mean dependent var | | 0.763065 |
| Adjusted R-squared | -0.001133 | S.D. dependent var | | 14.35931 |
| S.E. of regression | 14.36745 | Akaike info criterion | | 7.998453 |
| Sum squared resid | 207249.3 | Schwarz criterion | | 8.018006 |
| Log likelihood | -4015.223 | Hannan-Quinn criter. | | 8.005883 |
| Durbin-Watson stat | 1.957532 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

### 6.21

Conduct the analysis in similar fashion to Exercise 6.20 above. The results should not be too different.

6.22

[Question should cite the IMA(1) 

We rewrite the equation as:



The IMA(1) scheme for *Z* is



Substituting back to the original terms:

.

The result follows.

### Minicase 6.1

This minicase is an extension of Exercise 6.14 and the same general approach may be followed to develop suitable models. As with the quarterly data, a good starting point for the monthly series is the airline model. The annual data clearly exhibit a trend, suggesting ARIMA(0,1,1)+C or similar.

### Minicase 6.2

A variety of analyses is possible with this data set, such as:

1. Standard ARIMA analysis with data up to September 1996

2. Separate analysis for June 1997 on, and compare with overall model for entire series.

3. Adjustment for outliers.

4. Regression analysis to allow for marathons, injuries etc. followed by time series analysis of the residuals.

5. GARCH analysis of the residuals.