9.7 Exercises

1. Consider monthly sales and advertising data for an automotive parts company (data set `advert`).

   a. Plot the data using `autoplot`. Why is it useful to set `facets=TRUE`?

   b. Fit a standard regression model $y_t = a + bx_t + \eta_t$ where y_t denotes sales and x_t denotes advertising using the `tslm()` function.

   c. Show that the residuals have significant autocorrelation.

   d. What difference does it make you use the `Arima` function instead:

   ```
   Arima(advert[,"sales"], xreg=advert[,"advert"],
     order=c(0,0,0))
   ```

   e. Refit the model using `auto.arima()`. How much difference does the error model make to the estimated parameters? What ARIMA model for the errors is selected?

   f. Check the residuals of the fitted model.

   g. Assuming the advertising budget for the next six months is exactly 10 units per month, produce and plot sales forecasts with prediction intervals for the next six months.

2. This exercise uses data set `huron` giving the level of Lake Huron from 1875–1972.

   a. Fit a piecewise linear trend model to the Lake Huron data with a knot at 1920 and an ARMA error structure.
   b. Forecast the level for the next 30 years.

3. This exercise concerns `motel`: the total monthly takings from accommodation and the total room nights occupied at hotels, motels, and guest houses in Victoria, Australia, between January 1980 and June 1995. Total monthly takings are in thousands of Australian dollars; total room nights occupied are in thousands.

a. Use the data to calculate the average cost of a night's accommodation in Victoria each month.
b. Use `cpimel` to estimate the monthly CPI.
c. Produce time series plots of both variables and explain why logarithms of both variables need to be taken before fitting any models.
d. Fit an appropriate regression model with ARIMA errors. Explain your reasoning in arriving at the final model.
e. Forecast the average price per room for the next twelve months using your fitted model. (Hint: You will need to produce forecasts of the CPI figures first.)

4. We fitted a harmonic regression model to part of the `gasoline` series in Exercise 6 in Section 5.10. We will now revisit this model, and extend it to include more data and ARMA errors.

   a. Using `tslm()`, fit a harmonic regression with a piecewise linear time trend to the full `gasoline` series. Select the position of the knots in the trend and the appropriate number of Fourier terms to include by minimising the AICc or CV value.
   b. Now refit the model using `auto.arima()` to allow for correlated errors, keeping the same predictor variables as you used with `tslm()`.
   c. Check the residuals of the final model using the `checkresiduals()` function. Do they look sufficiently like white noise to continue? If not, try modifying your model, or removing the first few years of data.
   d. Once you have a model with white noise residuals, produce forecasts for the next year.

5. Electricity consumption is often modelled as a function of temperature. Temperature is measured by daily heating degrees and cooling degrees. Heating degrees is 18°C minus the average daily temperature when the daily average is below 18°C; otherwise it is zero. This provides a measure of our need to heat ourselves as temperature falls. Cooling degrees measures our need to cool ourselves as the temperature rises. It is defined as the average daily temperature minus 18°C when the daily average is above 18°C; otherwise it is zero. Let y_t denote the monthly total of kilowatt-hours of electricity used, let $x_{1,t}$ denote the monthly total of heating degrees, and let $x_{2,t}$ denote the monthly total of cooling degrees.

   An analyst fits the following model to a set of such data:

$$y_t^* = \beta_1 x_{1,t}^* + \beta_2 x_{2,t}^* + \eta_t,$$

where

$$(1-B)(1-B^{12})\eta_t = \frac{1-\theta_1 B}{1-\phi_{12}B^{12}-\phi_{24}B^{24}}\varepsilon_t$$

and $y_t^* = \log(y_t)$, $x_{1,t}^* = \sqrt{x_{1,t}}$ and $x_{2,t}^* = \sqrt{x_{2,t}}$.

a. What sort of ARIMA model is identified for η_t?

b. The estimated coefficients are

Parameter	**Estimate**	**s.e.**	Z	P**-value**
β_1	0.0077	0.0015	4.98	0.000
β_2	0.0208	0.0023	9.23	0.000
θ_1	0.5830	0.0720	8.10	0.000
ϕ_{12}	-0.5373	0.0856	-6.27	0.000
ϕ_{24}	-0.4667	0.0862	-5.41	0.000

Explain what the estimates of β_1 and β_2 tell us about electricity consumption.

c. Write the equation in a form more suitable for forecasting.
d. Describe how this model could be used to forecast electricity demand for the next 12 months.
e. Explain why the η_t term should be modelled with an ARIMA model rather than modelling the data using a standard regression package. In your discussion, comment on the properties of the estimates, the validity of the standard regression results, and the importance of the η_t model in producing forecasts.

6. For the retail time series considered in earlier chapters:

   a. Develop an appropriate dynamic regression model with Fourier terms for the seasonality. Use the AIC to select the number of Fourier terms to include in the model. (You will probably need to use the same Box-Cox transformation you identified previously.)
   b. Check the residuals of the fitted model. Does the residual series look like white noise?
   c. Compare the forecasts with those you obtained earlier using alternative models.