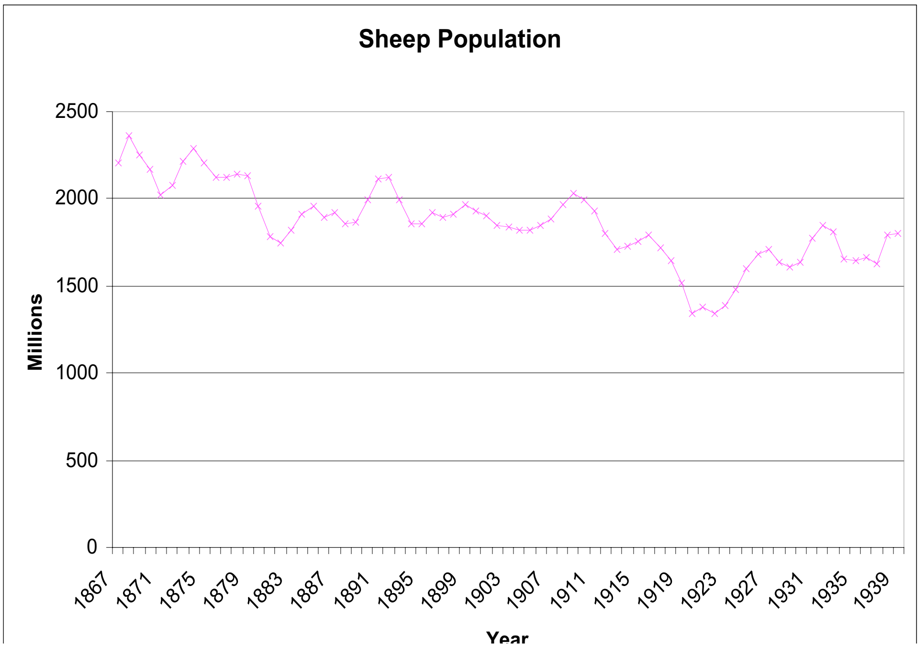
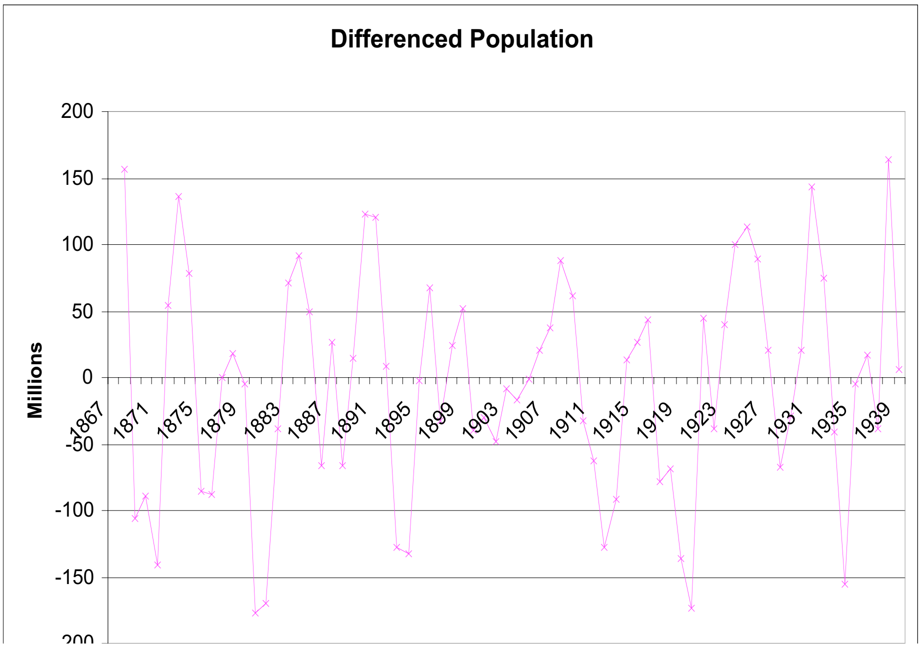
Time Series and Forecasting: Exercises 2 Solutions

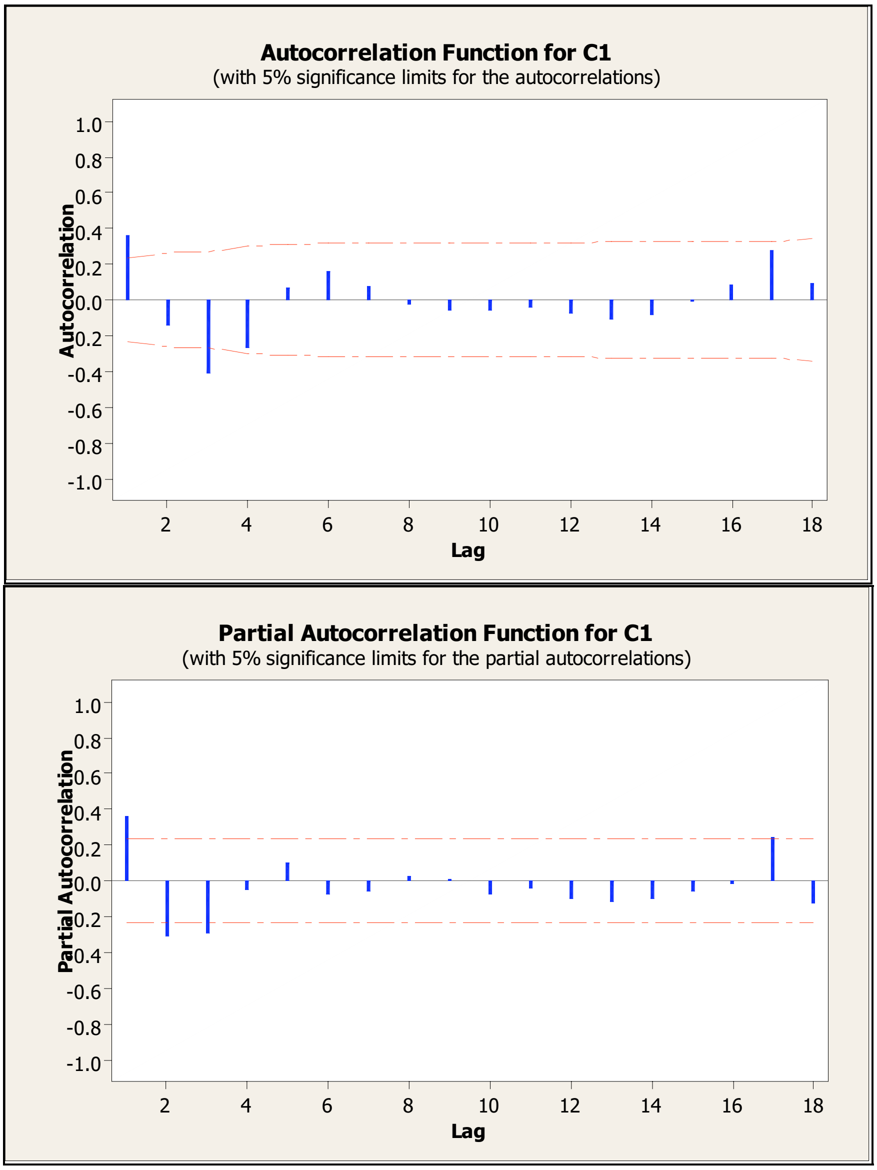
9. Original time series X_t:



There is a clear downward trend, whence we take first order differences $Y_t = X_t - X_{t-1}$. Note that we expect the Y_t to have a non-zero mean μ. The differenced data does not have any notable trends or seasonal behaviour, so we assume it to be stationary.



We use the sample acf and partial acf to help determine the model and its order. We can calculate them using Minitab (for example).



The pacf shows a clear cut off after lag 3. The acf is less clear, with a small correlation at lag 2 and a borderline correlation at lag 4. Thus an AR(3) model would seem reasonable. The spike at lag 17 can be attributed to chance.

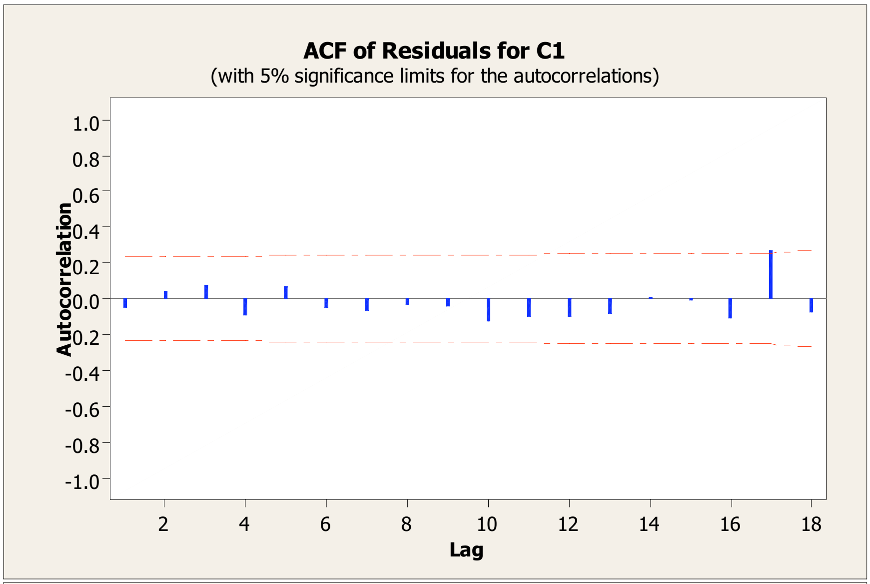
Our model is thus $Y_t = \mu + a_1(Y_{t-1} - \mu) + a_2(Y_{t-2} - \mu) + a_3(Y_{t-3} - \mu) + Z_t$ where the Z_t are i.i.d. mean 0. From Minitab we get the following estimates (using the mean for μ):

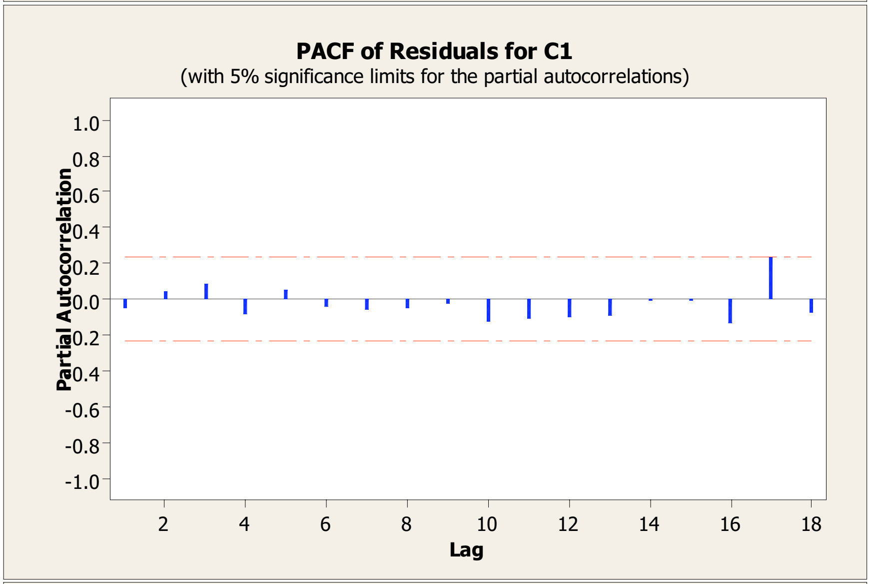
$$\mu = -5.893, a_1 = 0.4201, a_2 = -0.2105, a_3 = -0.3274.$$

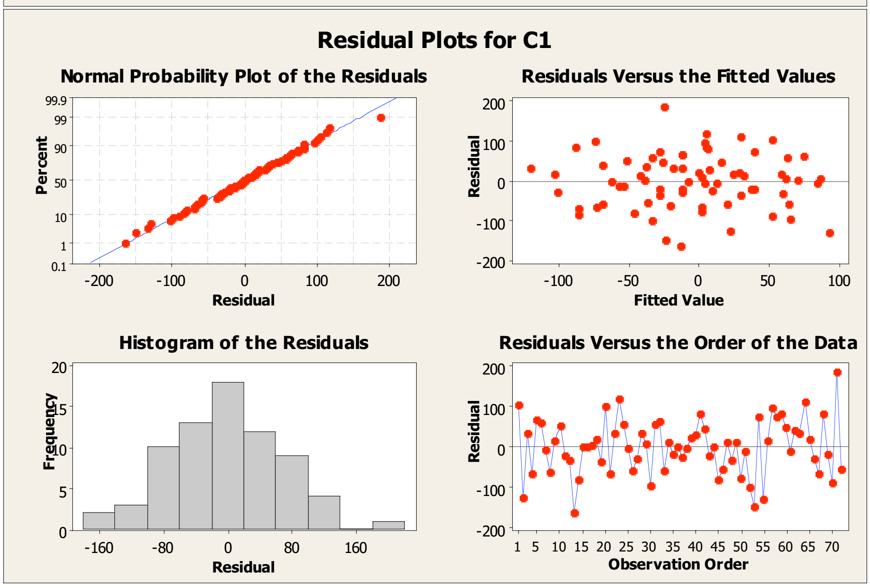
Given these estimates we can calculate the residuals

$$R_t = Y_t - \mu - a_1(Y_{t-1} - \mu) - a_2(Y_{t-2} - \mu) - a_3(Y_{t-3} - \mu).$$

You can calculate the residuals using a spreadsheet, or have Minitab calculate them for you. We use the residuals to check whether or not the model gives a good fit.







The acf and pacf are consistent with an i.i.d. sequence. The marginal spike at lag 17 remains because we did not specifically model this. The QQ plot and histogram are consistent with normally distributed residuals, and the two plots of residuals don't show any trends or heteroscedasticity. Thus we accept the model.

We forecast the Y_t as follows

$$\hat{Y}_{n+1} = \mu + a_1(Y_n - \mu) + a_2(Y_{n-1} - \mu) + a_3(Y_{n-2} - \mu) = -26.16$$
$$\hat{Y}_{n+2} = \mu + a_1(\hat{Y}_{n+1} - \mu) + a_2(Y_n - \mu) + a_3(Y_{n-1} - \mu) = -72.53$$
$$\hat{Y}_{n+3} = \mu + a_1(\hat{Y}_{n+2} - \mu) + a_2(\hat{Y}_{n+1} - \mu) + a_3(Y_n - \mu) = -33.52$$

Given these we forecast the X_t using

$$\hat{X}_{n+1} = \hat{Y}_{n+1} + X_n = 1770.85$$
$$\hat{X}_{n+2} = \hat{Y}_{n+2} + \hat{X}_{n+1} = 1698.32$$
$$\hat{X}_{n+3} = \hat{Y}_{n+3} + \hat{X}_{n+2} = 1664.81$$

Calculating confidence intervals requires writing the model in MA form, which we cannot sensibly do by hand. Also note that the $\hat{X}_t$ and $\hat{Y}_t$ will be dependent, so we cannot use confidence intervals for the $\hat{Y}_t$ to get confidence intervals for the $\hat{X}_t$ directly. However fitting an ARIMA(3,1,0) model to the original data we obtain from Minitab the following forecasts with 95% confidence intervals

Period	Forecast	Lower	Upper
74	1770.85	1634.03	1907.66
75	1698.31	1460.68	1935.95
76	1664.80	1360.77	1968.83