2. (6 marks) The units dataset in the TSA package gives the annual sales of certain large equipment, 1983-2005.

(a) Fit an MA(2) model (with a potentially non-zero constant mean) to this data using the arima() function in R. Give the estimates of the parameters θ1, θ2 and μ. (IMPORTANT: The way the arima( ) function defines the MA model is by placing plus signs, instead of minus signs, in front of the MA parameters. Therefore, the values of the MA parameters given in this output are actually −θ1 and −θ2!) (Note: As we saw in Video 32, the coefficient named “intercept” in the arima( ) output is actually referring to the mean μ, NOT the intercept θ0.)

```
# 2. (6 marks) The units dataset in the TSA package gives the annual sales of certain large
# equipment, 1983-2005.
data(units)
# (a) Fit an MA(2) model (with a potentially non-zero constant mean) to this data
# using the arima() function in R. Give the estimates of the parameters θ1, θ2 and μ.
# (IMPORTANT: The way the arima( ) function defines the MA model is by placing plus signs, instead of minus
#in front of the MA parameters. Therefore,
#   the values of the MA parameters given in this output are actually −θ1 and −θ2!)
# (Note: As we saw in Video 32, the coefficient named “intercept” in the arima( )
#   output is actually referring to the mean μ, NOT the intercept θ0.)
units.ma2.model = arima(units, order = c(0, 0, 2))
units.ma2.model
```

θ1 = -1.71

θ2 = -1

mu = 137.18

(b) Using the methods practiced in Video 33, derive the equation for the forecast of Yt+l at any lead time l. Make sure to replace any parameters with the estimates you obtained in part (a).

$\mu = 137.18 \quad \theta_1 = -1.71 \quad \theta_2 = -1$

Date/ Page/

MA(2) model

$$Y_t - \mu = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

$$\hat{Y}_t(l) - \mu = -\theta_1 E(e_{t+l-1} \mid Y_1, Y_2, \dots Y_t) - \theta_2 E(e_{t+l-2} \mid Y_1, Y_2, \dots, Y_t)$$

Recall

$$E(e_{t+j} \mid Y_1, Y_2, \dots Y_t) = \begin{cases} e_{t+j} & \text{if } j \le 0 \\ 0 & \text{if } j > 0. \end{cases}$$

Therefore.

$$\hat{Y}_t(l) = \begin{cases} \mu - \theta_1 e_{t+l-1} - \theta_2 e_{t+l-2} & \text{for } l \le 2 \\ \mu & \text{for } l > 1 \end{cases}$$

$$\hat{Y}_t(l) = \begin{cases} \mu - \theta_2 e_{t+l} & \text{for } l = 1 \\ \mu - \theta_1 e_t - \theta_2 e_{t+2} & \text{for } l = 0 \\ \mu & \text{for } l \ge 2. \end{cases}$$

$$\hat{Y}_t(l) = \begin{cases} 137.18 + e_{t+1} & \text{for } l = 1 \\ 137.18 + 1.71 e_t + e_{t+2} & \text{for } l = 0 \\ 137.18 & \text{for } l \ge 2. \end{cases}$$

(c) Derive the equation for the forecast error variance for Yt+l , denoted by V ar(et(l)). Make sure to include each possible case of values that l can take on.

(d) Using your equation in part (b), obtain the forecast of Yt+1. Show your calculations. (Note: You can use R's estimates of the noise terms to help you out. The estimates of e1, . . . , et can be found in the object name_ of_ your_ ma2_ model $residuals .)

```
# (d) Using your equation in part (b), obtain the forecast of Yt+1. Show your calculations.
# (Note: You can use R's estimates of the noise terms to help you out. The estimates
#   of e1, . . . , et can be found in the object name_ of_ your_ ma2_ model $residuals .)
units.ar2.model$residuals
```

```
> units.ar2.model$residuals
Time Series:
Start = 1982
End = 2005
Frequency = 1
 [1] -21.3326712  -1.7776049  11.4187820 -18.6147653  -7.5682668   0.8588146  -2.8471222  -6.5924036  -5.2826071 -23.6825588  -0.3551033
[12]  12.2194644   1.6687669 -11.3095093   1.4284394   7.2128921  17.6696294  -8.0242493  -5.1961215  -2.4503521   6.2925833  12.3559089
[23]  38.3138988   9.1431551
```

$$\hat{Y}_t(l) = \begin{cases} 137.13 - 1.78 \\ 137.13 + 1.71 \times (-21.33) + 11.42 \\ 137.13 \end{cases}$$

(e) Using your forecast in part (d), and the equation in part (c), calculate the the 95% prediction limits for Yt+1. (Note: You can use R's estimate of the white noise variance if you need it. It can be found in the object name_ of_ your_ ma2_ model $sigma2 .)

(f) Create a plot of the predictions of Yt+` out to 20 time points in the future. Does the forecast for ` = 1 match your results above? (Hint: You can directly read the values off the plot or, if you'd like exact values, you can extract them by adding $pred , $lpi or $upi after the plot( ) function. This will give you the forecasts, and lower and upper 95% prediction limits, respectively.)