Homework 9

Brian Detweiler

April 24, 2017

```r
set.seed(48548493)
```

1. Consider an AR(1) model with ϕ = -0.5 and μ = 14. If our last observation occurred at time 50 and the value observed was 12, calculate $\hat{Y}_{50}(1)$, $\hat{Y}_{50}(2)$, and $\hat{Y}_{50}(10)$.

From equation 9.3.6, we get

$$\begin{aligned}\hat{Y}_t(1) &= \mu + \phi(Y_t - \mu)\\ &= 14 - 0.5(12 - 14)\\ &= 15\end{aligned}$$

Once we have $\hat{Y}_t(1)$, we can use equation 9.3.7 in conjunction with this and obtain the general values for arbitrary $\hat{Y}_t(l)$.

$$\begin{aligned}\hat{Y}_t(2) &= \mu + \phi(\hat{Y}_{2-1} - \mu)\\ &= 14 - 0.5(15 - 14)\\ &= 13.5\end{aligned}$$

$$\begin{aligned}\hat{Y}_t(10) &= \mu + \phi(\hat{Y}_{10-1} - \mu)\\ &= 14 - 0.5(14.0039062 - 14)\\ &= 13.9980469\end{aligned}$$

■

2. Consider the AR(2) model $Y_t = 5 + 1.1Y_{t-1} - 0.5Y_{t-2} + e_t$ with $\sigma_e^2 = 2$.

(a) If the last 3 observed values of the sequence are $Y_{90} = 9$, $Y_{91} = 11$, and $Y_{92} = 10$, find the forecasts $\hat{Y}_{92}(1)$, $\hat{Y}_{92}(2)$, and $\hat{Y}_{92}(3)$.

For an ARMA(2, 0), we can use equation 9.3.28,

$$
\begin{aligned}
\hat{Y}_t(l) &= \phi_1 \hat{Y}_t(l-1) + \phi_2 \hat{Y}_t(l-2) \\
\hat{Y}_{92}(1) &= \phi_1 \hat{Y}_{92} + \phi_2 \hat{Y}_{91} \\
&= 1.1(10) - 0.5(11) \\
&= 5.5 \\
\hat{Y}_{92}(2) &= \phi_1 \hat{Y}_{93} + \phi_2 \hat{Y}_{92} \\
&= 1.1(5.5) - 0.5(10) \\
&= 1.05 \\
\hat{Y}_{92}(3) &= \phi_1 \hat{Y}_{94} + \phi_2 \hat{Y}_{93} \\
&= 1.1(1.05) - 0.5(5.5) \\
&= -1.595
\end{aligned}
$$

(b) Find the variances of $e_{92}(1)$, $e_{92}(2)$, $e_{92}(3)$.

$$
\begin{aligned}
Var(e_t(l)) &= \sigma_e^2 \sum_{j=0}^{l-1} \psi_j^2 = l\sigma_e^2 \\
Var(e_t(1)) &= 1(2) = 2 \\
Var(e_t(2)) &= 2(2) = 4 \\
Var(e_t(3)) &= 3(2) = 6
\end{aligned}
$$

(c) Find the 95% prediction limits for the forecasts $\hat{Y}_{92}(1)$, $\hat{Y}_{92}(2)$, and $\hat{Y}_{92}(3)$.

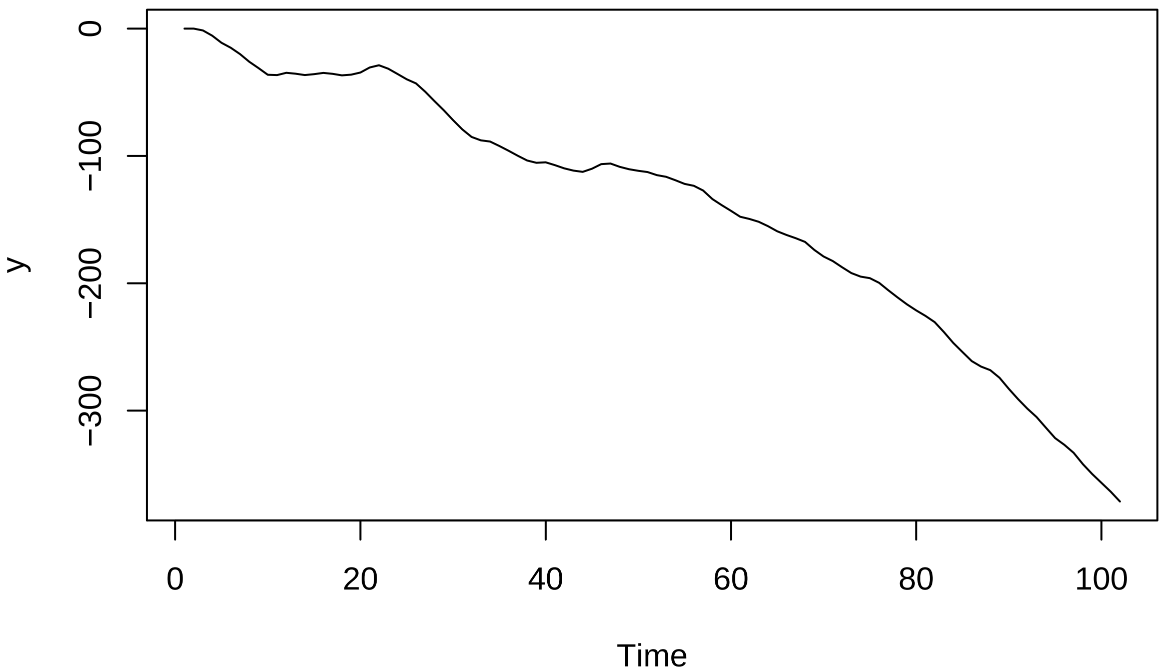
The 95% prediction limits for the forcasts are at $\pm 2\sqrt{\sigma_e^2}$ for each value of $Var(e_t(l))$.

$$
\begin{aligned}
\hat{Y}_{92}(1) &: (3.5, 7.5) \\
\hat{Y}_{92}(2) &: (-2.95, 5.05) \\
\hat{Y}_{92}(3) &: (-7.595, 4.405)
\end{aligned}
$$

■

3. Use `arima.sim` with $n = 100$ to simulate an ARIMA(0,2,2) with $\theta_1 = 1$ and $\theta_2 = -0.75$ (Use your NUID in `set.seed`). Store the data as `y`. Look at the data, notice that there are 102 observations and the first two are both zero. Remove the first two (zero) observations.

```
y <- arima.sim(n=100, list(order=c(0, 2, 2), ma = c(1, -0.75)))
plot(y)
```




```
y <- ts(y[3:102])
```

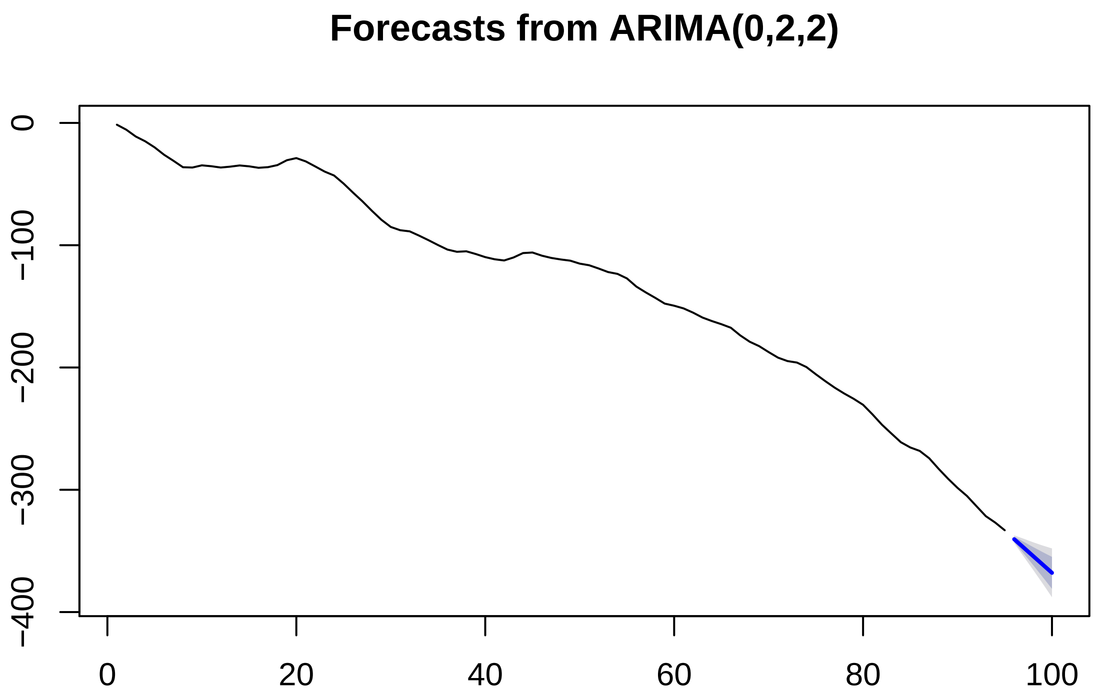
(a) Fit an ARIMA(0,2,2) model to `y[1:95]`, the first 95 observations of the simulated series, and find the maximum likelihood estimates of θ_1 and θ_2 .

```
y.95 <- ts(y[1:95])
fit <- arima(y.95, order=c(0, 2, 2), method = "ML")
fit
```

```
##
## Call:
## arima(x = y.95, order = c(0, 2, 2), method = "ML")
##
## Coefficients:
##          ma1      ma2
##       0.1232  -0.4785
## s.e.  0.0977   0.1013
##
## sigma^2 estimated as 2.512:  log likelihood = -175.08,  aic = 354.16
```

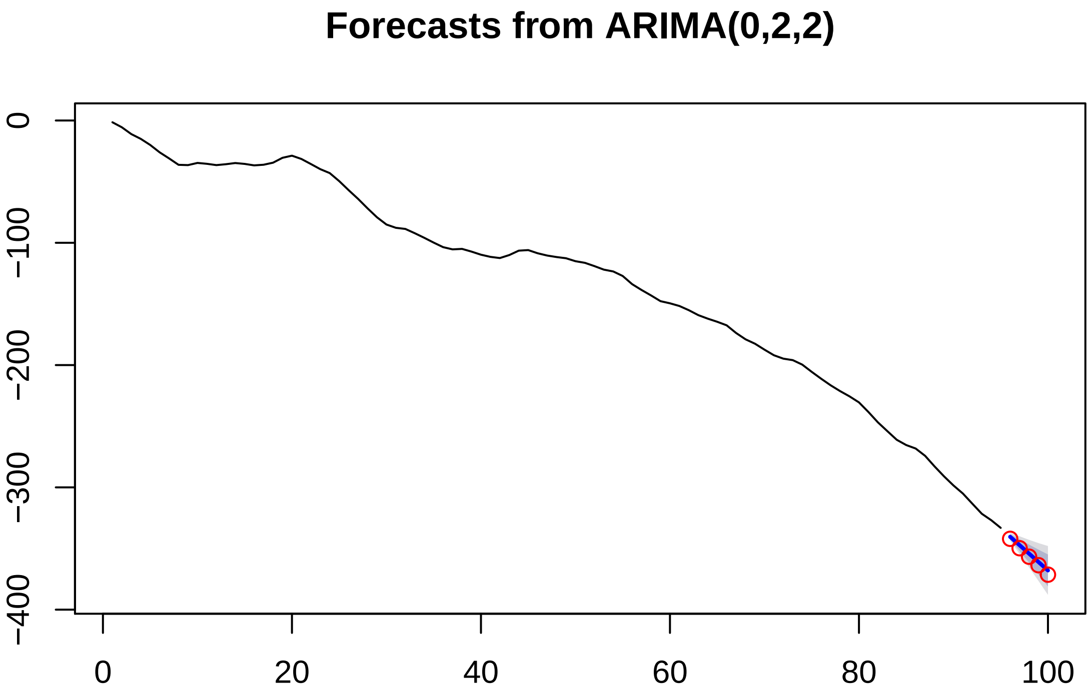
(b) Construct a time series plot that shows observations 91 through 95, and the forecasts (with prediction limits) for observations 96 through 100.

```
fore <- forecast(y.95, model = fit, h = 5)
plot(fore)
```



(c) Use `points(96:100, y[96:100], col="red")` to add the actual observations to the plot. Compare the forecasts with the actual observations.

```
plot(fore)
points(96:100, y[96:100], col="red")
```


Forecasts from ARIMA(0,2,2)
0
−100
−200
−300
−400
0
20
40
60
80
100

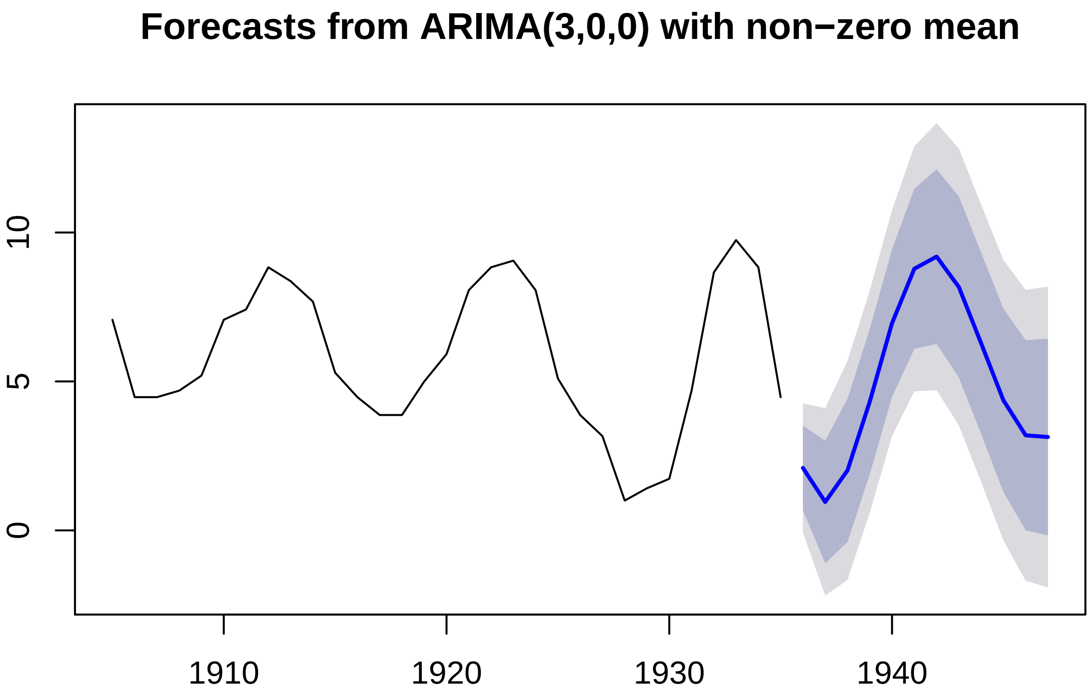
■

4. We have previously fitted an AR(3) to the square-root of the `hare` data.

```
data(hare)
```

(a) Fit the model and plot the time series along with the forecasts (with prediction limits) for the next 12 observations of the square root of hare abundance.

```
fit <- Arima(y = sqrt(hare), order=c(3, 0, 0))
fore <- forecast(object=sqrt(hare), model=fit, h=12)
plot(fore)
```



(b) Use your answer to part (a) to find the numerical values of the forecasts for the hare abundance for the next 12 years.

```
hares <- (coredata(fore$mean))^2
upper <- (coredata(fore$upper[,2]))^2
lower <- (coredata(fore$lower[,2]))^2

for (i in 1:12) {
  print(paste0("Year: ", (i + 1935), " - ", hares[i], " hares"))
}
```

```
## [1] "Year: 1936 - 4.39174950543336 hares"
## [1] "Year: 1937 - 0.911565638251505 hares"
## [1] "Year: 1938 - 4.05207464586828 hares"
## [1] "Year: 1939 - 18.6767034337167 hares"
## [1] "Year: 1940 - 48.3854755465944 hares"
## [1] "Year: 1941 - 77.1203480714941 hares"
## [1] "Year: 1942 - 84.477180002393 hares"
## [1] "Year: 1943 - 66.7127553523153 hares"
```

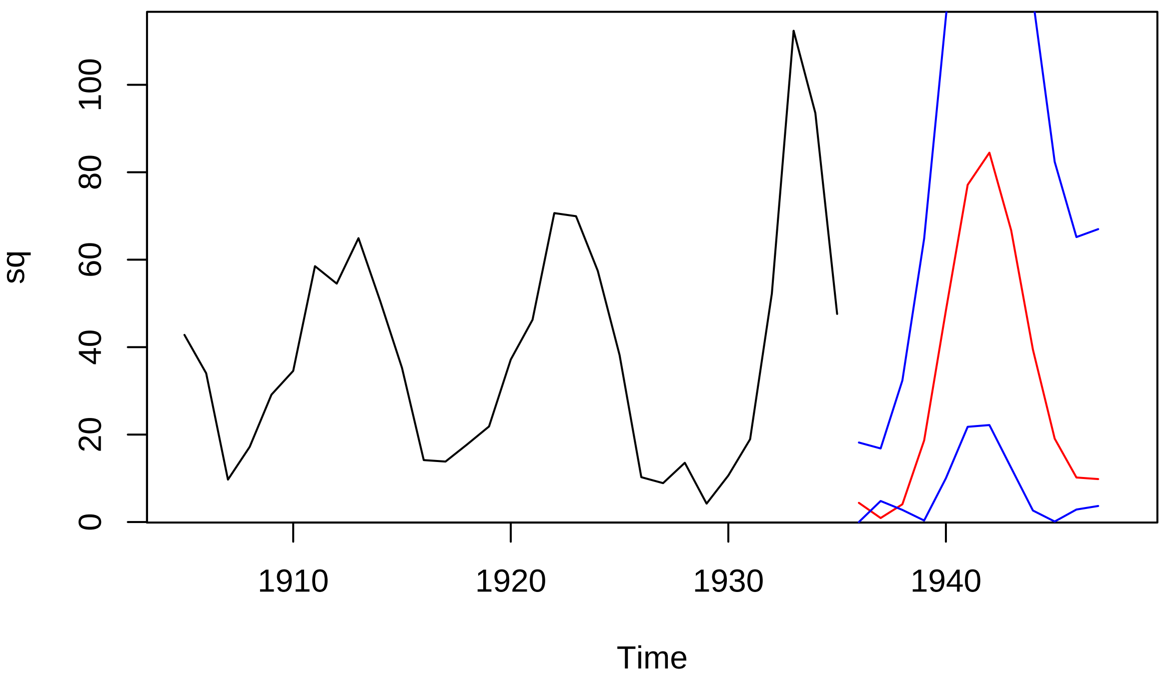
```
## [1] "Year: 1944 - 39.436797771036 hares"
## [1] "Year: 1945 - 19.0768504577495 hares"
## [1] "Year: 1946 - 10.1844639713439 hares"
## [1] "Year: 1947 - 9.8233753927957 hares"
```

(c) How do we plot the values you found in (b) along with the corresponding intervals? The plot command has a `transform` option that allows the data to be transformed.

```r
square <- function(x) {
  y <- x^2
}

sq <- square(fore$fitted)

plot(sq, xlim=c(1905, 1948))
lines(1936:1947, hares, col="red")
lines(1936:1947, upper, col="blue")
lines(1936:1947, lower, col="blue")
```



■

5. Consider the multiplicative seasonal ARIMA $(0,0,2)\times(0,1,0)_4$ model.

(a). Write down the model $Y_t =$ ________________________________.

$$W_t = \nabla_4 Y_t = Y_{t-4} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2})$$

(b). Find the first four ψ-weights for this model.

$$\begin{aligned} Y_t &= Y_{t-4} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \\ &= (Y_{t-8} + e_{t-4} - \theta_1 e_{t-5} - \theta_2 e_{t-6}) + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \\ &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} + e_{t-4} - \theta_1 e_{t-5} - \theta_2 e_{t-6} + Y_{t-8} \end{aligned}$$

Thus the ψ-weights here are $\psi_1 = 1$, $\psi_2 = -\theta_1$, $\psi_3 = 0$, and $\psi_4 = 1$.

(c). Suppose that $\theta_1 = 0.5$, $\theta_2 = -0.25$, and $\sigma_e^2 = 1$, and that the last four observed values were: 25, 20, 25, 40, with corresponding residuals: 2, 1, 2, 3. Predict the next 4 values.

We have

$$\begin{aligned} \hat{Y}_t(1) &= Y_{t-3} - \theta_1 e_t - \theta_2 e_{t-1} \\ &= 25 - (0.5)(3) - (-0.25)(2) \\ &= 24 \\ \hat{Y}_t(2) &= Y_{t-2} - \theta_2 e_t \\ &= 20 - (-0.25)(3) \\ &= 20.75 \\ \hat{Y}_t(3) &= Y_{t-1} = 25 \\ \hat{Y}_t(4) &= 40 \end{aligned}$$

(d) Construct prediction intervals for the predictions found in (b).

$$\begin{aligned} &\hat{Y}_t(1) : 24 \pm 2\sqrt{1} \Rightarrow (20, 26) \\ &\hat{Y}_t(2) : 20.75 \pm 2\sqrt{1 + (0.5)^2} \Rightarrow (18.513932, 22.986068) \\ &\hat{Y}_t(3) : 24 \pm 2\sqrt{1 + (0.5)^2 + (0.25)^2} \Rightarrow (21.7087122, 26.2912878) \\ &\hat{Y}_t(4) : 40 \pm 2\sqrt{1 + (0.5)^2 + (0.25)^2 + 0^2} \Rightarrow (37.7087122, 42.2912878) \end{aligned}$$

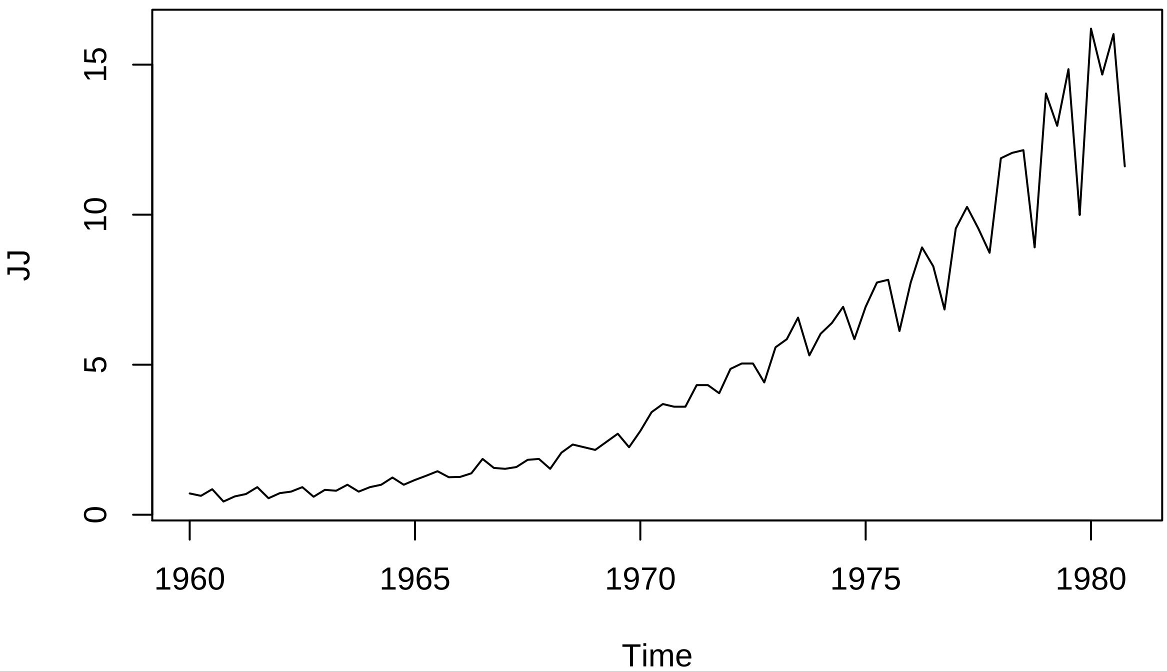
■

6. The dataset `JJ` contains the earnings per share for each quarter from 1960 to 1980 for Johnson and Johnson.

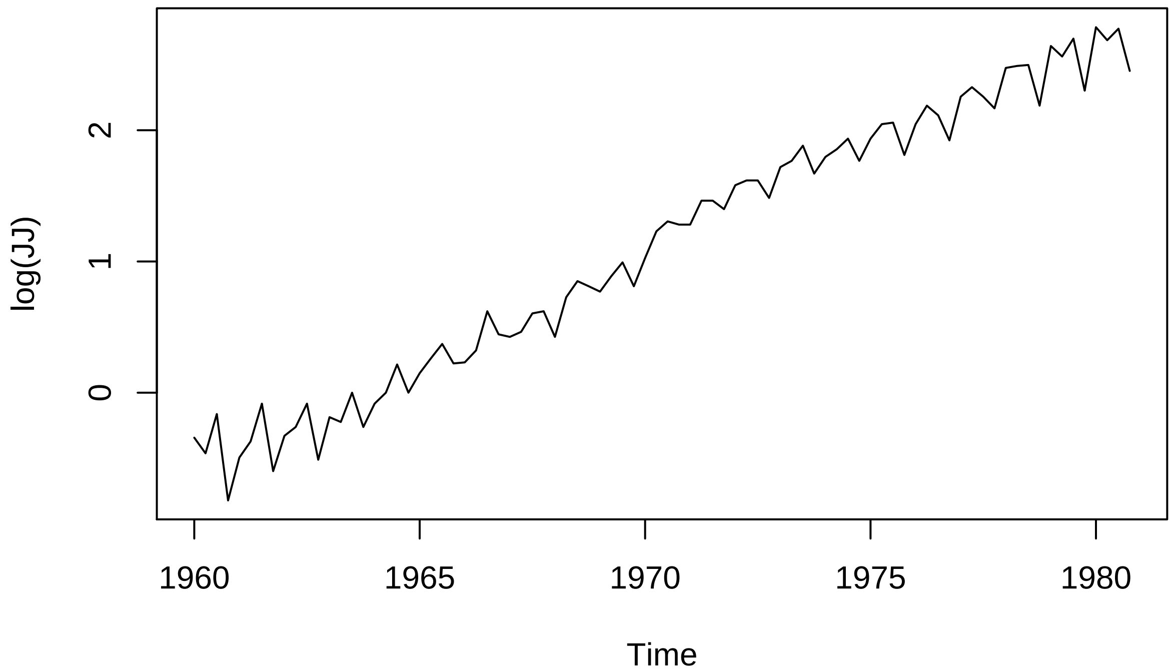
```r
data(JJ)
```

(a) Plot the time series and also the logarithm of the series. Whether should we use the log transformation to model this series? Explain.

```r
plot(JJ)
```




```r
plot(log(JJ))
```



We'll use the log, because it creates a linear pattern.

(b) Based on your decision of part (a), find the most appropriate ARIMA $(p, d, q) \times (P, D, Q)_s$ to fit the data or transformed data.

```
fit <- auto.arima(log(JJ))
```

We will use an ARIMA $(2, 0, 0) \times (1, 1, 0)_4$.

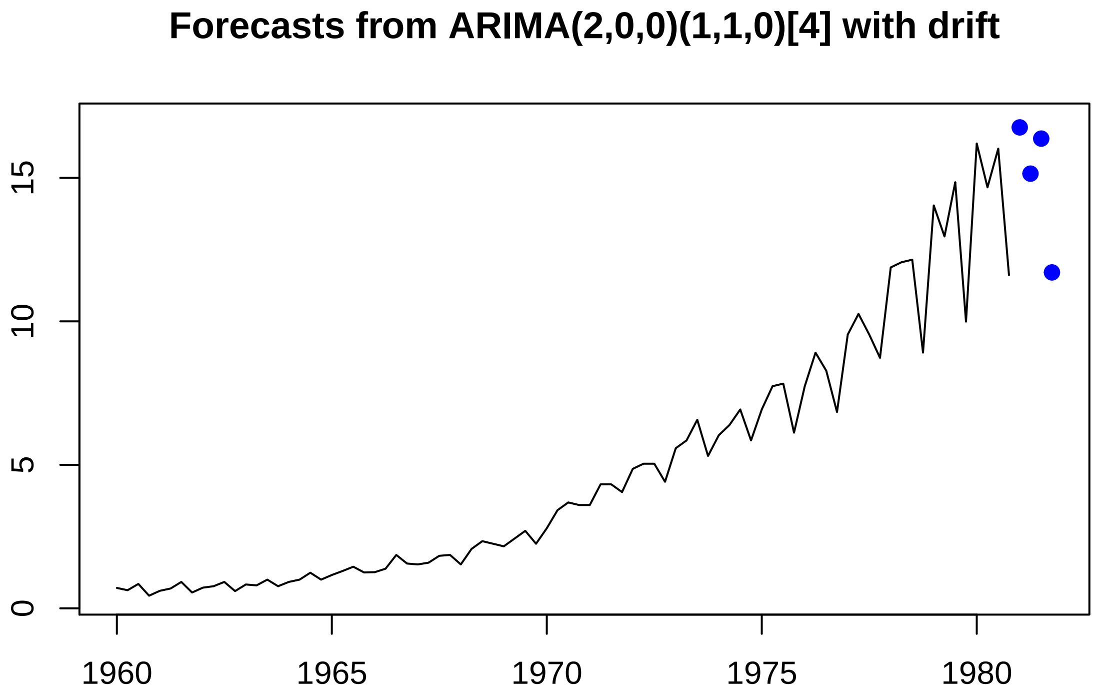
(c) Estimate the parameters of the chosen model.

```
fit
```

```
## Series: log(JJ)
## ARIMA(2,0,0)(1,1,0)[4] with drift
##
## Coefficients:
##          ar1     ar2     sar1   drift
##       0.2686  0.2855  -0.2695  0.0382
## s.e.  0.1137  0.1214   0.1212  0.0042
##
## sigma^2 estimated as 0.007793:  log likelihood=82.47
## AIC=-154.95   AICc=-154.14   BIC=-143.04
```

(d) Use your model to predict the next 4 values of the series.

```
fore <- forecast(object = JJ, model=fit, h = 4)
plot(fore)
```




```r
fore
```

```
##         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 1981 Q1       16.75936 16.64623 16.87249 16.58634 16.93238
## 1981 Q2       15.14774 15.03060 15.26488 14.96859 15.32689
## 1981 Q3       16.36907 16.24514 16.49300 16.17954 16.55861
## 1981 Q4       11.70623 11.58077 11.83169 11.51435 11.89810
```

■