ASSIGNMENT 9 SOLUTIONS

STAT 485/685 E100/G100: Applied Time Series Analysis

Fall 2020

Simon Fraser University

1. The "gold" dataset in the TSA package gives the daily price of gold (in \$ per troy ounce) for the 252 trading days of 2005.

a) Suppose we were to try fitting an AR(1) model to this dataset. Fit this model using the arima() function in R. Give the estimates of ϕ and μ.
(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 θ_0.)

Solution:

```r
library(TSA)
```

```r
data(gold)
gold.ar1 <- arima(gold, order=c(1,0,0))
gold.ar1
```

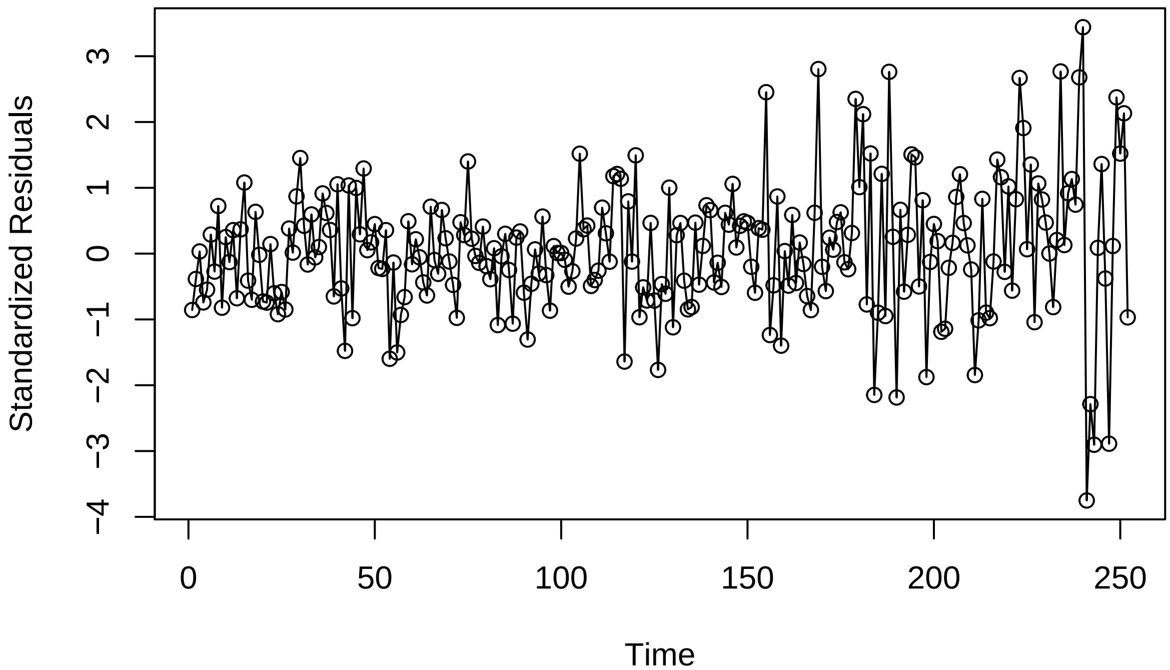
```
##
## Call:
## arima(x = gold, order = c(1, 0, 0))
##
## Coefficients:
##          ar1  intercept
##       0.9947   458.5493
## s.e.  0.0055    29.0241
##
## sigma^2 estimated as 14.5:  log likelihood = -696.78,  aic = 1397.56
```

From the output, we see that the estimates are: $\hat{\phi} = 0.995$ and $\hat{\mu} = 458.55$. If we wished, we could also easily obtain the standard errors of these estimates (as seen in the output).

b) Create a plot of the (standardized) residuals vs. time for this model. Interpret what you see in the plot.

Solution:

```r
plot(rstandard(gold.ar1), type='o', ylab='Standardized Residuals', xlab='Time')
```



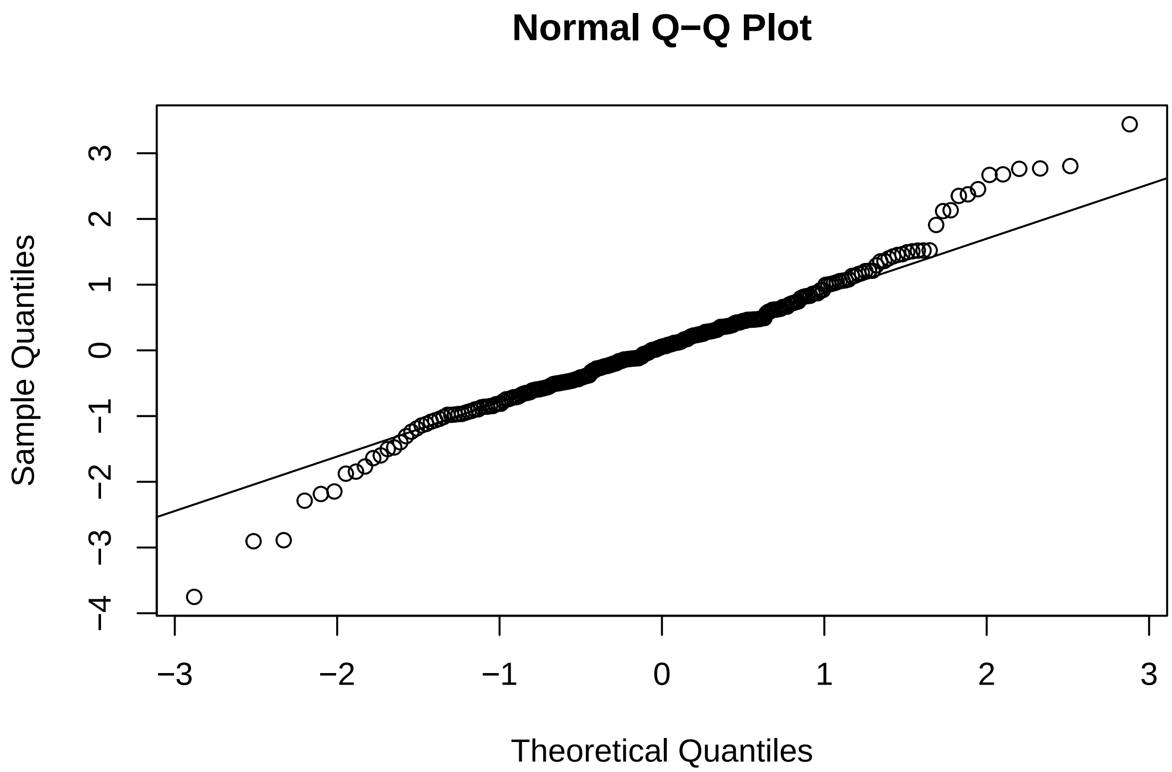
From this plot, we see that the residuals do not appear to resemble white noise very closely. In particular, there is evidence of unequal variances: the values are much more variable at later times.

However, there doesn't appear to be much evidence of significant correlations (either positive or negative), and the mean does appear to be constant about zero. There are also no worrying patterns in the plot, such as U-shapes or upside-down U-shapes.

c) Create a Q-Q plot of the (standardized) residuals for this model. Interpret what you see in the plot.

Solution:

```
qqnorm(rstandard(gold.ar1))
qqline(rstandard(gold.ar1))
```

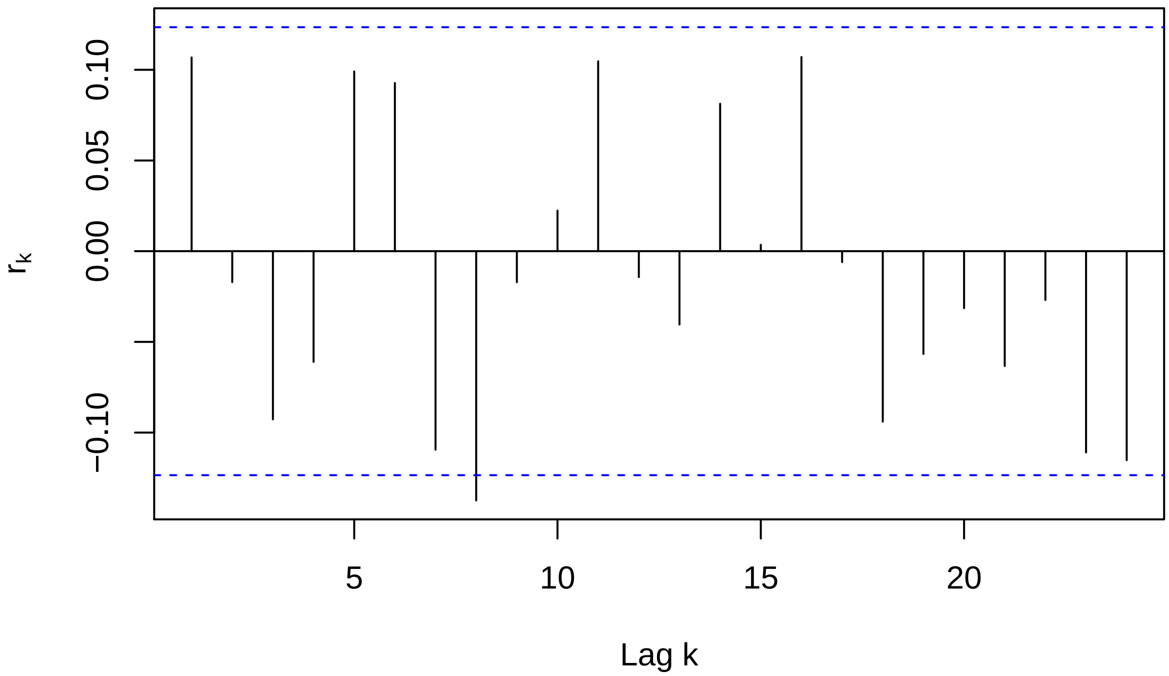


The relationship between the sample quantiles of the (standardized) residuals and the theoretical standard normal quantiles does not fall closely on a straight line. In particular, the points at the left and right tails are quite far from the line. Therefore, we conclude that the standardized residuals do not appear to follow a normal distribution.

d) Create the sample ACF plot of the (standardized) residuals for this model. Interpret what you see in the plot.

Solution:

```
acf(rstandard(gold.ar1), xlab='Lag k', ylab=expression('r'[k]), main='')
```



From the sample ACF plot, we see that the residuals appear to be mostly uncorrelated. At almost all lags k, the sample autocorrelation r_k between residuals is deemed to be not significantly different from zero. There is a single significant negative correlation at lag 8; however, this may be just due to random chance since all the other correlations are insigificant.

This plot gives us evidence in favour of the residuals "looking like" white noise. However, as we saw from the plot of residuals vs. time, more than correlation needs to be considered – the variances were clearly unequal, which is also important.

2. 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 θ_1, θ_2 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 $-\theta_1$ and $-\theta_2$!)
(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 θ_0.)

Solution:

```
data(units)
```

As was pointed out in an Announcement to the class, the package will struggle to obtain forecasts in part (f) if we use the “units” dataset directly. Instead, we take the following roundabout approach:

```
my.dataset <- units
units.ma2 <- arima(my.dataset, order=c(0,0,2))
units.ma2
```

```
##
## Call:
## arima(x = my.dataset, order = c(0, 0, 2))
##
## Coefficients:
##          ma1     ma2  intercept
##       1.7078  1.0000   137.1767
## s.e.  0.1786  0.1666     8.8527
##
## sigma^2 estimated as 147.9:  log likelihood = -97.23,  aic = 200.46
```

From the output, we see that the estimates are: $\hat{\theta}_1 = -1.71$, $\hat{\theta}_2 = -1.00$ and $\hat{\mu} = 137.18$. If we wished, we could also easily obtain the standard errors of these estimates (as seen in the output).

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

Solution:

From slide 9 of Video 33, we recall that the forecast of $Y_{t+\ell}$ for any ARMA model is defined as

$$\hat{Y}_t(\ell) = \theta_0 + \phi_1 \hat{Y}_t(\ell-1) + \phi_2 \hat{Y}_t(\ell-2) + \cdots + \phi_p \hat{Y}_t(\ell-p)$$
$$- \theta_1 E(e_{t+\ell-1}|Y_1, Y_2, \ldots, Y_t) - \theta_2 E(e_{t+\ell-2}|Y_1, Y_2, \ldots, Y_t) - \cdots - \theta_q E(e_{t+\ell-q}|Y_1, Y_2, \ldots, Y_t)$$

where

$$\hat{Y}_t(j) = \begin{cases} \text{the actual forecast } \hat{Y}_t(j) & \text{if } j > 0 \\ \text{the observed value } Y_{t+j} & \text{if } j \leq 0 \end{cases}$$

and

$$E(e_{t+j}|Y_1, Y_2, \ldots, Y_t) = \begin{cases} e_{t+j} & \text{if } j \leq 0 \\ 0 & \text{if } j > 0 \end{cases}$$

Therefore, for the MA(2) model, the forecasts are:

$$\hat{Y}_t(\ell) = \theta_0 - \theta_1 E(e_{t+\ell-1}|Y_1, Y_2, \ldots, Y_t) - \theta_2 E(e_{t+\ell-2}|Y_1, Y_2, \ldots, Y_t)$$

where

$$E(e_{t+\ell-1}|Y_1, Y_2, \ldots, Y_t) = \begin{cases} e_{t+\ell-1} & \text{if } \ell \leq 1 \\ 0 & \text{if } \ell > 1 \end{cases} = \begin{cases} e_t & \text{if } \ell = 1 \\ 0 & \text{if } \ell > 1 \end{cases}$$

and

$$E(e_{t+\ell-2}|Y_1, Y_2, \ldots, Y_t) = \begin{cases} e_{t+\ell-2} & \text{if } \ell \leq 2 \\ 0 & \text{if } \ell > 2 \end{cases} = \begin{cases} e_{t-1} & \text{if } \ell = 1 \\ e_t & \text{if } \ell = 2 \\ 0 & \text{if } \ell > 2 \end{cases}$$

In other words:

$$\hat{Y}_t(\ell) = \begin{cases} \theta_0 - \theta_1 e_t - \theta_2 e_{t-1} & \text{if } \ell = 1 \\ \theta_0 - \theta_2 e_t & \text{if } \ell = 2 \\ \theta_0 & \text{if } \ell > 2 \end{cases}$$

Since the R output has given us an estimate of μ, rather than θ_0, we would probably prefer the "alternative" formulation, where we remove the θ_0 and replace all $\hat{Y}$'s by $(\hat{Y} - \mu)$. Since this is just an MA model, this is very easy to do: Replacing $\hat{Y}_t(\ell)$ by $(\hat{Y}_t(\ell) - \mu)$, and bringing the μ over to the right, we get:

$$\hat{Y}_t(\ell) = \begin{cases} \mu - \theta_1 e_t - \theta_2 e_{t-1} & \text{if } \ell = 1 \\ \mu - \theta_2 e_t & \text{if } \ell = 2 \\ \mu & \text{if } \ell > 2 \end{cases}$$

The final step is to replace all the parameters by their estimates:

$$\hat{Y}_t(\ell) = \begin{cases} 137.1767 + 1.7078 e_t + 1.0000 e_{t-1} & \text{if } \ell = 1 \\ 137.1767 + 1.0000 e_t & \text{if } \ell = 2 \\ 137.1767 & \text{if } \ell > 2 \end{cases}$$

c) Derive the equation for the forecast error variance for $Y_{t+\ell}$, denoted by $Var(e_t(\ell))$. Make sure to include each possible case of values that ℓ can take on.

Solution:

From slide 10 of Video 33, we know that the forecast error variance can be expressed as

$$Var(e_t(\ell)) = \sigma_e^2 \sum_{j=0}^{\ell-1} \psi_j^2,$$

where $\psi_0 = 1$, and $\psi_1, \psi_2, \ldots, \psi_{\ell-1}$ are the first $\ell - 1$ coefficients in the general linear process formulation of the ARMA model.

We saw on slide 14 of Video 17 that an MA(q) process is a special case of a general linear process, which truncates after q terms, and whose coefficients are just $-\theta$'s. In other words:

$$\psi_j = \begin{cases} -\theta_j & \text{for } j \leq q \\ 0 & \text{for } j > q \end{cases}$$

This is evident just from looking at the definition of the general linear process (first seen in Video 14).

Therefore, we consider all possible values that ℓ can take on, and how these will affect the expression of the forecast error variance.

If $\ell = 1$:

$$Var(e_t(\ell)) = \sigma_e^2 \sum_{j=0}^{0} \psi_j^2 = \sigma_e^2 \psi_0^2 = \sigma_e^2,$$

since $\psi_0 = 1$.

If $\ell = 2$:

$$Var(e_t(\ell)) = \sigma_e^2 \sum_{j=0}^{1} \psi_j^2 = \sigma_e^2 (\psi_0^2 + \psi_1^2) = \sigma_e^2\,(1 + (-\theta_1)^2) = \sigma_e^2\,(1 + \theta_1^2),$$

since $\psi_1 = -\theta_1$.

If $\ell = 3$:

$$Var(e_t(\ell)) = \sigma_e^2 \sum_{j=0}^{2} \psi_j^2 = \sigma_e^2 (\psi_0^2 + \psi_1^2 + \psi_2^2) = \sigma_e^2\,(1 + (-\theta_1)^2 + (-\theta_2)^2) = \sigma_e^2\,(1 + \theta_1^2 + \theta_2^2),$$

since $\psi_2 = -\theta_2$.

Finally, for all $\ell > 3$:

$$Var(e_t(\ell)) = \sigma_e^2 \sum_{j=0}^{\ell-1} \psi_j^2 = \sigma_e^2 (\psi_0^2 + \psi_1^2 + \psi_2^2) = \sigma_e^2\,(1 + \theta_1^2 + \theta_2^2),$$

since $\psi_j = 0$ for all $j > 2$.

Therefore, we can summarize these expressions as:

$$Var(e_t(\ell)) = \begin{cases} \sigma_e^2 & \text{if } \ell = 1 \\ \sigma_e^2\,(1 + \theta_1^2) & \text{if } \ell = 2 \\ \sigma_e^2\,(1 + \theta_1^2 + \theta_2^2) & \text{if } \ell > 2 \end{cases}$$

d) Using your equation in part (b), obtain the forecast of Y_{t+1}. Show your calculations. *(Note: You can use R's estimates of the noise terms to help you out. The estimates of $e_1, \ldots, e_t$ can be found in the object name.of.your.ma2.model$residuals.)*

Solution:

From the equation in part (b), the forecast of Y_{t+1} is:

$$\hat{Y}_t(1) = 137.1767 + 1.7078 e_t + 1.0000 e_{t-1}.$$

Therefore, we need estimates of e_t and e_{t-1} in order to evaluate this forecast. We obtain them as follows:

```
resids <- units.ma2$residuals
t <- length(my.dataset)
e_t <- resids[t]
e_t
```

```
## [1] 17.8631
```

```r
e_tminus1 <- resids[t-1]
e_tminus1
```

```
## [1] 29.00726
```

Therefore:

$$\hat{Y}_t(1) = 137.1767 + 1.7078(17.8631) + 1.0000(29.00726) = 196.6912 \approx 196.69.$$

e) Using your forecast in part (d), and the equation in part (c), calculate the 95% prediction limits for Y_{t+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.)

Solution:

From slide 8 of Video 34, we know that the $100(1-\alpha)\%$ prediction limits for any $Y_{t+\ell}$ are

$$\hat{Y}_t(\ell) \pm z_{\alpha/2}\sqrt{Var(e_t(\ell))},$$

where $z_{\alpha/2}$ is the $(1-\alpha/2)^{\text{th}}$ percentile of the standard normal distribution, i.e. it is the point on the distribution such that $Pr(Z < z_{\alpha/2}) = 1 - \alpha/2$. Since we are interested in constructing 95% prediction limits, we have $\alpha = 0.05$, and therefore we can obtain $z_{\alpha/2}$ using R as follows:

```r
alpha <- 0.05
z_alpha2 <- qnorm(p=(1-alpha/2), mean=0, sd=1, lower.tail=TRUE)
z_alpha2
```

```
## [1] 1.959964
```

So, $z_{\alpha/2} \approx 1.96$.

Then, we also need $Var(e_t(\ell))$, which we found in part (c) to be equal to σ_e^2 when $\ell = 1$. From R, we obtain:

```r
sigma2e <- units.ma2$sigma2
sigma2e
```

```
## [1] 147.9096
```

Therefore, $\hat{\sigma}_e^2 = 147.9096$.

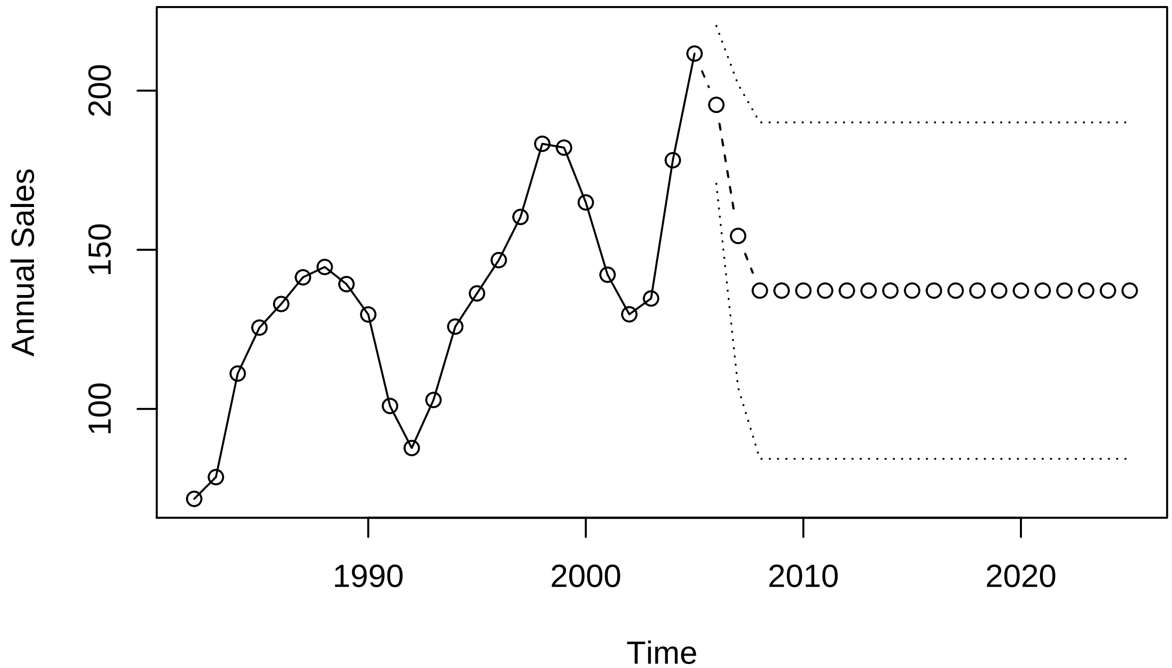
Then, we can calculate:

$$\begin{aligned}
&\hat{Y}_t(\ell) \pm z_{\alpha/2}\sqrt{Var(e_t(\ell))} \\
&\hat{Y}_t(1) \pm z_{0.025}\sqrt{Var(e_t(1))} \\
&196.6912 \pm 1.96\sqrt{147.9096} \\
&196.6912 \pm 23.83714 \\
\implies &(172.85, 220.53)
\end{aligned}$$

f) Create a plot of the predictions of $Y_{t+\ell}$ out to 20 time points in the future. Does the forecast for $\ell = 1$ 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.)

Solution:

```r
my.forecasts <- plot(units.ma2, n.ahead=20, type='b', xlab='Time', ylab='Annual Sales')
```



From the plot above, the forecast at $\ell = 1$ appears to be approximately 190, with 95% prediction limits approximately between 170 and 220. However, it is somewhat difficult to read these numbers from the plot. Instead, we can obtain the exact values as follows:

```r
my.forecasts$pred[1]
```

```
## [1] 195.5453
```

```r
my.forecasts$lpi[1]
```

```
## [1] 170.7646
```

```r
my.forecasts$upi[1]
```

```
## [1] 220.3261
```

Therefore, the forecast at $\ell = 1$ is 195.55, with prediction limits $(170.76, 220.33)$. This approximately matches our results above, and the minor differences are likely just due to different choices of estimates for e_t and e_{t-1}.