ASSIGNMENT 8 SOLUTIONS

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

Fall 2020

Simon Fraser University

1. The "color" dataset in the TSA package gives the values of a colour property from 35 consecutive batches in an industrial process. Suppose we decide to fit an AR(1) model to this dataset.
(Note: If we were to investigate more closely, we would see that this might not be the most appropriate model for this dataset. However, we will consider it here.)

a) Fit the AR(1) model to this dataset, using the Maximum Likelihood Estimation approach within the ar() function. Give the estimates of ϕ and μ.
(Hint: We will need the argument method="mle" in the ar( ) function.)

Solution:

```
library(TSA)
```

```
data(color)
ar1.mle <- ar(color, order.max=1, aic=FALSE, method='mle')
ar1.mle
```

```
##
## Call:
## ar(x = color, aic = FALSE, order.max = 1, method = "mle")
##
## Coefficients:
##      1
## 0.5703
##
## Order selected 1  sigma^2 estimated as  24.83
```

```
ar1.mle$x.mean
```

```
## intercept
##  74.34332
```

The parameter estimates are: $\hat{\phi} = 0.570$ and $\hat{\mu} = 74.343$ (we note that this is obtained by calling the "x.mean" object).

b) Write out the full equation(s) you could use to estimate σ_e^2, using the estimates of ϕ and μ. Make sure to plug the estimates of ϕ and μ into the equation. You do not have to actually evaluate this estimate, since the dataset is somewhat large.

Solution:

From the notes on slide 6 of Video 31, the MLE approach to estimating σ_e^2 is by obtaining

$$\hat{\sigma}_e^2 = \frac{S(\hat{\phi}, \hat{\mu})}{n-2},$$

where

$$\begin{aligned}
S(\hat{\phi}, \hat{\mu}) &= \sum_{t=2}^{n}[(Y_t - \hat{\mu}) - \hat{\phi}(Y_{t-1} - \hat{\mu})]^2 + (1 - \hat{\phi}^2)(Y_1 - \hat{\mu})^2 \\
&= \sum_{t=2}^{n}[(Y_t - 74.34332) - 0.5703(Y_{t-1} - 74.34332)]^2 + (1 - (0.5703)^2)(Y_1 - 74.34332)^2
\end{aligned}$$

We can obtain $Y_1, Y_2, \ldots, Y_n$ from the dataset, and $n = 35$ (this can be obtained by calling length(color)).

We also note that an estimate of σ_e^2 is also given in the AR(1) model output in R. These values will be very similar.

2.

a) Fit the AR(1) model to the "color" dataset, using the Method of Moments approach within the ar() function. Give the estimates of ϕ and μ.
(Hint: We will need the argument method="yw" in the ar() function. This stands for "Yule-Walker", because the Yule-Walker equations need to be solved to get the MOM parameter estimates.)

Solution:

```
ar1.mom <- ar(color, order.max=1, aic=FALSE, method='yw')
ar1.mom
```

```
##
## Call:
## ar(x = color, aic = FALSE, order.max = 1, method = "yw")
##
## Coefficients:
##      1
## 0.5282
##
## Order selected 1  sigma^2 estimated as  27.56
```

```
ar1.mom$x.mean
```

```
## [1] 74.88571
```

The parameter estimates are: $\hat{\phi} = 0.528$ and $\hat{\mu} = 74.886$ (this is again obtained by calling the "x.mean" object).

b) Using equation(s) we have learned about in Video 29, obtain an estimate of the process variance γ_0.
(Hint: You may have to explore a bit to find a function in R that can give you the sample variance of a dataset.)

Solution:

From the notes on slide 19 of Video 29, an unbiased estimate of $\gamma_0 = Var(Y_t)$ is the sample variance,

$$s^2 = \frac{1}{n-1}\sum_{t=1}^{n}(Y_t - \bar{Y})^2$$

In R, the "var" function calculates the sample variance of a dataset. Therefore:

```
s2 <- var(color)
s2
```

```
## [1] 37.1042
```

The estimate of the process variance is therefore: $\hat{\gamma}_0 = s^2 = 37.104$.

c) Using the above results, and equation(s) we have learned about in Video 29, obtain an estimate of the white noise variance σ_e^2.
(Hint: If you need some sample correlations from the dataset: For any dataset mydata, the vector of r_k-values is given by acf(mydata)$acf.)

Solution:

The notes on slide 19 of Video 29 tell us that we estimate σ_e^2 using the relationships between it and the other parameters $(\gamma_0, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$.

For an AR(1) process, we know from Video 16 that

$$\gamma_0 = \frac{\sigma_e^2}{1 - \phi\rho_1},$$

so therefore,

$$\begin{aligned}
\hat{\sigma}_e^2 &= \hat{\gamma}_0(1 - \hat{\phi}r_1) \\
&= s^2(1 - \hat{\phi}r_1) \\
&= 37.1042(1 - 0.5282 \times 0.5282) \\
&= 26.752
\end{aligned}$$

where the lag-1 autocorrelation can be obtained by calling acf(color)$acf[1], or by simply recognizing the fact that it is equal to the MOM estimator of ϕ that we have already obtained above. (The reason why they are equal was discussed on slide 10 of Video 29.)

3.

a) Fit the AR(1) model to the "color" dataset, using the (Conditional) Least Squares approach within the ar() function. Give the estimates of ϕ and μ.
(Hint: We will need the argument method="ols" in the ar() function.)

Solution:

```
ar1.ols <- ar(color, order.max=1, aic=FALSE, method='ols')
ar1.ols
```

```
##
## Call:
## ar(x = color, aic = FALSE, order.max = 1, method = "ols")
##
## Coefficients:
##      1
## 0.5549
##
## Intercept: 0.1032 (0.8474)
##
## Order selected 1  sigma^2 estimated as  24.38
```

```
ar1.ols$x.mean
```

```
## [1] 74.88571
```

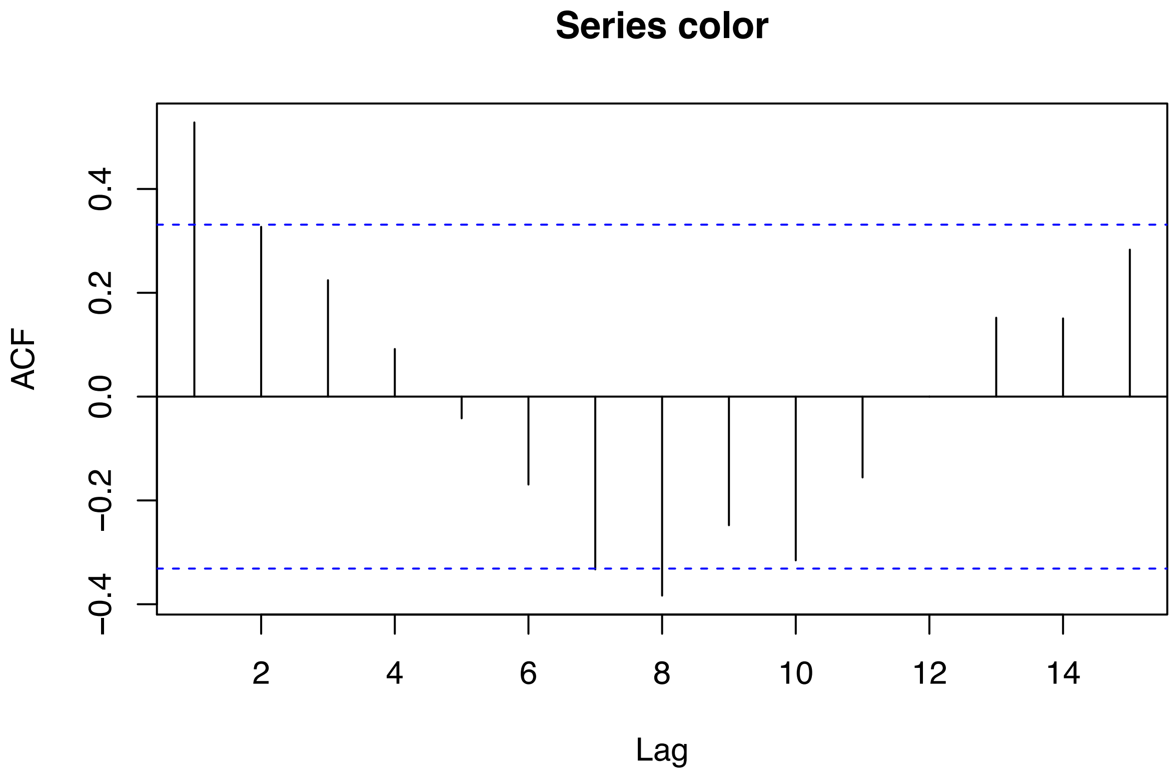
The parameter estimates are: $\hat{\phi} = 0.555$ and $\hat{\mu} = 74.886$ (this is again obtained by calling the "x.mean" object).

b) How does the estimate of ϕ compare to the lag-1 autocorrelation for this dataset?

Solution:

We can obtain the sample lag-1 autocorrelation from R as follows:

```
r1 <- acf(color)$acf[1]
```




```r
r1
```

```
## [1] 0.5282091
```

So, the sample lag-1 autocorrelation is $r_1 = 0.528$.

We saw on slide 8 of Video 30 that, for large n, we should expect that $\hat{\phi} \approx r_1$ (in other words, the LSE estimator of ϕ is approximately equal to the MOM estimator of ϕ).

For this dataset, the values are similar but not all that close. We suspect that the reason for this is that our dataset is not very large: $n = 35$.

c) Obtain the estimate of μ in a different way – by calculating the sample mean of the data. How does this compare to the estimate of μ you obtained in part (a)?

Solution:

```r
mean(color)
```

```
## [1] 74.88571
```

The sample mean of the data is $\bar{Y} = 74.886$. This is equal to our estimate above.

4. Suppose we have a time series dataset of size $n = 3$, as follows: $Y_1 = 4$, $Y_2 = 3$ and $Y_3 = 7$.

(a) Evaluate the lag-1 autocorrelation for this dataset. Show your work.

Solution:

From slide 5 of Video 23 (the sample ACF video):

$$r_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}.$$

Therefore, for our dataset:

$$\begin{aligned} r_1 &= \frac{\sum_{t=2}^{3}(Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=1}^{3}(Y_t - \bar{Y})^2} \\ &= \frac{(Y_2 - \bar{Y})(Y_1 - \bar{Y}) + (Y_3 - \bar{Y})(Y_2 - \bar{Y})}{(Y_1 - \bar{Y})^2 + (Y_2 - \bar{Y})^2 + (Y_3 - \bar{Y})^2} \\ &= \frac{(3 - 4.6667)(4 - 4.6667) + (7 - 4.6667)(3 - 4.6667)}{(4 - 4.6667)^2 + (3 - 4.6667)^2 + (7 - 4.6667)^2} \\ &= -0.321. \end{aligned}$$

We obtained the sample mean as:

$$\bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t = \frac{Y_1 + Y_2 + Y_3}{3} = \frac{4 + 3 + 7}{3} = 4.6667.$$

(b) Suppose we wish to fit an MA(1) model to this dataset. Calculate the Method of Moments estimate of the parameter θ, by hand. Show your work.

Solution:

From slides 14-16 of Video 29, we recall that the solution for the MOM estimate of θ will depend on the value of r_1. In part (a), we found that $r_1 = -0.321$. Therefore, $|r_1| = 0.321 < 0.5$, and so we consider "Scenario A" on slide 15.

"Scenario A" tells us that there are two unique, real solutions for $\hat{\theta}$, and that only one of them will satisfy the invertibility condition. The one that satisfies it is:

$$\hat{\theta} = \frac{-1 + \sqrt{1 - 4r_1^2}}{2r_1} = \frac{-1 + \sqrt{1 - 4(-0.320513)^2}}{2(-0.320513)} = 0.362670 \approx 0.36$$

(c) Is this solution invertible? Explain your reasoning.

Solution:

We already concluded that this solution is invertible, on slide 15 of Video 29. However, we can check it again here:

$$|\hat{\theta}| = |0.362670| = 0.362670 < 1,$$

therefore the solution is invertible.

(d) Calculate an estimate of the process variance γ_0. Show your work.

Solution:

From the notes on slide 19 of Video 29, an unbiased estimate of $\gamma_0 = Var(Y_t)$ is the sample variance,

$$\begin{aligned} s^2 &= \frac{1}{n-1}\sum_{t=1}^{n}(Y_t - \bar{Y})^2 \\ &= \frac{(Y_1 - \bar{Y})^2 + (Y_2 - \bar{Y})^2 + (Y_3 - \bar{Y})^2}{2} \\ &= \frac{(4 - 4.6667)^2 + (3 - 4.6667)^2 + (7 - 4.6667)^2}{2} \\ &= 4.333333 \\ &\approx 4.33 \end{aligned}$$

Therefore, the estimate of the process variance is: $\hat{\gamma}_0 = s^2 = 4.33.$

(e) Calculate an estimate of the white noise variance σ_e^2. Show your work.

Solution:

The notes on slide 19 of Video 29 tell us that we estimate σ_e^2 using the relationships between it and the other parameters $(\gamma_0, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$.

For an MA(1) process, we know from Video 15 that

$$\gamma_0 = \sigma_e^2\,(1 + \theta^2)$$

so therefore,

$$\hat{\sigma}_e^2 = \frac{\hat{\gamma}_0}{1 + \hat{\theta}^2} = \frac{s^2}{1 + \hat{\theta}^2} = \frac{4.333333}{1 + (0.362670)^2} = 3.829626 \approx 3.83$$

(f) Calculate an estimate of the process mean μ. Show your work.

Solution:

In the `hare` data example on slide 22 of Video 29, we discussed how an estimate of the process mean μ can be obtained using the sample mean, $\bar{Y}$. Therefore,

$$\hat{\mu} = \bar{Y} = 4.666667 \approx 4.67,$$

which we first calculated in part (a).

(g) Using the parameter estimates you have obtained above, write out an equation for the model for $\{Y_t\}$.

Solution:

The model for $\{Y_t\}$ is an MA(1) process with a non-zero constant mean. This can be written as follows (note that we are substituting parameters with their estimates):

$$(Y_t - 4.67) = e_t - 0.36\,e_{t-1}.$$

(h) Re-write the equation from part (g), in terms of an intercept term θ_0 (instead of the mean value μ).

Solution:

Re-writing the equation, we obtain:

$$Y_t = 4.67 + e_t - 0.36\, e_{t-1}.$$

We note that, for MA models in general,

$$\theta_0 = \mu(1 - \phi_1 - \cdots - \phi_p) = \mu,$$

and so it was very easy to re-write the model formulation in terms of θ_0.