1.

5.1 The sunspot numbers $\{X_t, t = 1, \ldots, 100\}$, filed as SUNSPOTS.TSM, have sample autocovariances $\hat{\gamma}(0) = 1382.2$, $\hat{\gamma}(1) = 1114.4$, $\hat{\gamma}(2) = 591.73$, and $\hat{\gamma}(3) = 96.216$. Use these values to find the Yule–Walker estimates of ϕ_1, ϕ_2, and σ^2 in the model

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2),$$

for the mean-corrected series $Y_t = X_t - 46.93$, $t = 1, \ldots, 100$. Assuming that the data really are a realization of an AR(2) process, find 95 % confidence intervals for ϕ_1 and ϕ_2.

for AR(2): $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t$, $Z_t \sim WN(0,\sigma^2)$

to get YW: $E(Y_t Y_{t-h}) = E[\phi_1 Y_{t-1} Y_{t-h}] + E(\phi_2 Y_{t-2} Y_{t-h}) + E(Y_{t-h} Z_t)$, $h \geq 1$

$\gamma(h) = \phi_1 \gamma(h-1) - \phi_2 \gamma(h-2)$

$$\begin{cases} \gamma(1) = \phi_1 \gamma(0) - \phi_2 \gamma(1) \\ \gamma(2) = \phi_1 \gamma(1) - \phi_2 \gamma(0) \end{cases}$$

plug in

$\hat{\rho}(0) = \frac{\hat{\gamma}(0)}{\hat{\gamma}(0)} = 1$

$\hat{\rho}(1) = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = 0.8063$

$\hat{\rho}(2) = \frac{\hat{\gamma}(2)}{\hat{\gamma}(0)} = 0.4281$

we have $\hat{\Phi} = \begin{bmatrix} \hat{\rho}(0) & \hat{\rho}(1) \\ \hat{\rho}(1) & \hat{\rho}(0) \end{bmatrix}^{-1} \begin{pmatrix} \hat{\gamma}(1) \\ \hat{\gamma}(2) \end{pmatrix} = \begin{pmatrix} 1.318 \\ -0.634 \end{pmatrix}$

$\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}_1 \hat{\gamma}(1) - \hat{\phi}_2 \hat{\gamma}(2)$

$= 289.18$

another way: $\frac{\hat{\sigma}^2}{n} \hat{\Gamma}_2^{-1} = \frac{289.18}{100} \begin{pmatrix} 0.0021 & -0.0017 \\ -0.0017 & 0.0021 \end{pmatrix}$

95% CI for ϕ_1 is given by $\hat{\phi}_1 \pm 1.96 \sqrt{\frac{1-\hat{\phi}_2^2}{n}} = 1.318 \pm 0.15$

for ϕ_2 is given by $\hat{\phi}_2 \pm 1.96 \sqrt{\frac{1-\hat{\phi}_2^2}{n}} = -0.634 \pm 0.15$

1.

5.2 From the information given in the previous problem, use the Durbin–Levinson algorithm to compute the sample partial autocorrelations $\hat{\phi}_{11}$, $\hat{\phi}_{22}$, and $\hat{\phi}_{33}$ of the sunspot series. Is the value of $\hat{\phi}_{33}$ compatible with the hypothesis that the data are generated by an AR(2) process? (Use significance level 0.05.)

$\hat{\rho}(1) = \hat{\gamma}(1) / \hat{\gamma}(0) = 1114.4 / 1382.2 = 0.8063$

$\hat{\rho}(2) = \hat{\gamma}(2) / \hat{\gamma}(0) = 591.73 / 1382.2 = 0.4281$

$\hat{\rho}(3) = \hat{\gamma}(3) / \hat{\gamma}(0) = 96.216 / 1382.2 = 0.0696$

Based on Durbin-Levinson algorithm

$$\begin{cases} \phi_{11} = \rho(1) \\ \phi_{22} = \frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)} \\ \phi_{21} = \phi_{11} - \phi_{22}\phi_{11} \\ \phi_{33} = \frac{\rho(3) - \phi_{21}\rho(2) - \phi_{22}\rho(1)}{1 - \phi_{21}\rho(1) - \phi_{22}\rho(2)} \end{cases}$$

$\Rightarrow \hat{\phi}_{11} = 0.8063$

$\hat{\phi}_{22} = (0.4281 - 0.8063^2) / (1 - 0.8063^2)$

$= -0.6346$

$\hat{\phi}_{21} = 0.8063 + 0.5117 = 1.318$

$\hat{\phi}_{33} = \frac{0.0696 - 1.318 \times 0.4281 + 0.6346 \times 0.8063}{1 - 1.318 \times 0.8063 + 0.6346 \times 0.4281}$

≈ 0.08 $\boxed{0.082}$

Consider: $H_0: \phi_{33} = 0$,

$H_1: \phi_{33} \neq 0$

the 95% CI for $\hat{\phi}_{33}$ is $\hat{\phi}_{33} \pm \frac{1.96}{\sqrt{n}} = 0.0804 \pm \frac{1.96}{\sqrt{100}} = (-0.1156, 0.2764)$

include zero, so we can not reject H_0, that means our model is more likely to be AR(2) Process.

1

5.4 Two hundred observations of a time series, $X_1, \ldots, X_{200}$, gave the following sample statistics:

sample mean: $\bar{x}_{200} = 3.82$;

sample variance: $\hat{\gamma}(0) = 1.15$;

sample ACF: $\hat{\rho}(1) = 0.427$;

$\hat{\rho}(2) = 0.475$;

$\hat{\rho}(3) = 0.169$.

a. Based on these sample statistics, is it reasonable to suppose that $\{X_t - \mu\}$ is white noise?

b. Assuming that $\{X_t - \mu\}$ can be modeled as the AR(2) process

$$X_t - \mu - \phi_1(X_{t-1} - \mu) - \phi_2(X_{t-2} - \mu) = Z_t,$$

where $\{Z_t\} \sim \mathrm{IID}(0, \sigma^2)$, find estimates of μ, ϕ_1, ϕ_2, and σ^2.

c. Would you conclude that $\mu = 0$?

d. Construct 95 % confidence intervals for ϕ_1 and ϕ_2.

e. Assuming that the data were generated from an AR(2) model, derive estimates of the PACF for all lags $h \geq 1$.

a. if $X_t \sim WN(0, \sigma^2)$, the 95% CI for sample ACF is $\hat\rho(h) \pm \frac{1.96}{\sqrt{n}} = \hat\rho(h) \pm 0.139$

so all the CI intervals for $\hat\rho(1)$, $\hat\rho(2)$, $\hat\rho(3)$ not include 0 which means they are significant, so $\{X_t - \mu\}$ is not a WN process.

b. $\hat\mu$ = sample mean = $\bar{X}_{200} = 3.82$

based on YW equations, for AR(2)

$$\hat\Phi = \begin{bmatrix} \hat\rho(0) & \hat\rho(1) \\ \hat\rho(1) & \hat\rho(0) \end{bmatrix}^{-1} \begin{pmatrix} \hat\gamma(1) \\ \hat\gamma(2) \end{pmatrix} = \begin{pmatrix} 1 & 0.427 \\ 0.427 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0.427 \times 1.15 \\ 0.475 \times 1.15 \end{pmatrix} = \begin{pmatrix} 0.315 \\ 0.412 \end{pmatrix} \begin{matrix} \hat\phi_1 \\ \hat\phi_2 \end{matrix}$$

$$\hat\sigma^2 = \hat\gamma(0) - \hat\Phi_1 \hat\gamma(1) - \hat\Phi_2 \hat\gamma(2)$$

$$= 1.15 - 0.315 \times 0.475 - 0.412 \times 0.475 = 0.7703$$

c. $H_0: \mu = 0$. $H_1: \mu \neq 0$

the CI for μ is $\bar{X} \pm \sqrt{\frac{\sum_{h=-\infty}^{\infty} \gamma(h)}{n}} \approx \bar{X} \pm \sqrt{\frac{\hat\gamma(-3) + \hat\gamma(-2) + \cdots + \hat\gamma(3)}{n}}$

not include 0, so we reject H_0, which means $\mu \neq 0$.

d. for ϕ_1 and ϕ_2, the 95% CI is

$$\hat\Phi_1 \pm 1.96\sqrt{\frac{1-\hat\Phi_2^2}{n}} = 0.315 \pm 1.96 \times \sqrt{\frac{1-0.797^2}{200}} = 1.678 \pm 0.084$$

$$\hat\Phi_2 \pm 1.96\sqrt{\frac{1-\hat\Phi_2^2}{n}} = 0.412 \pm 0.084$$

e. Based on the Durbin-Levinso Algorithm:

$$\begin{cases} \phi_{11} = \rho(1) \\ \phi_{22} = \dfrac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2} \end{cases}$$

for AR(2) model

$\hat\Phi_{11} = \hat\rho(1) = 0.427$

$\hat\Phi_{22} = 0.358$

$\hat\phi_{hh} = 0$, $h > 2$.

2. Derive the ACF of a SARIMA $(0,0,0)\times(1,0,0)_{12}$ model.

the model is $X_t - \Phi X_{t-12} = e_t$

$\gamma(0) = cov(X_t, X_t) = cov(\Phi X_{t-12} + e_t,\ \Phi X_{t-12} + e_t)$

$= \Phi^2 \gamma(0) + \sigma^2$

$\because X_{t-12}, e_t$ is independent $\therefore cov(X_{t-12}, e_t) = 0$

$\Rightarrow \gamma(0) = \dfrac{\sigma^2}{1-\Phi^2}$

$\gamma(1) = cov(X_t, X_{t+1}) = cov(\Phi X_{t-12} + e_t,\ \Phi X_{t-11} + e_{t+1}) = 0$

......

$\gamma(12) = cov(X_t, X_{t+12}) = cov(X_t,\ \Phi X_t + e_{t+12})$

$= \Phi\gamma(0)$

$\because X_t, e_{t+12}$ is independent $\therefore cov(X_t, e_{t+12}) = 0$

$\Rightarrow$ In General,

$$\begin{cases} \gamma(0) = \dfrac{\sigma^2}{1+\Phi^2} \\ \gamma(12k) = \Phi^k \gamma(0) \\ \gamma(h) = 0 \quad \text{o.w.} \end{cases} \Rightarrow \text{ACF } \rho(h) = \begin{cases} 1, & h=0 \\ \Phi^k, & h=12k,\ k=1,2,\cdots \\ 0, & \text{ow.} \end{cases}$$

5. 5. Suppose X_t satisfies an $AR(1)$ model with parameter ϕ. How long a series do you need to estimate ϕ so that with 95% confidence the estimation error within ± 0.1 of the true value?

AR(1): $X_t - \phi X_{t-1} = e_t$

Based on the Yule-Walker Estimation: $\hat\phi \sim N(\phi, \frac{1-\phi^2}{n})$

with 95% CI: $\hat\phi \pm 1.96\sqrt{\frac{1-\hat\phi^2}{n}}$

$1.96\sqrt{\frac{1-\hat\phi^2}{n}} = 0.1$

$\Rightarrow \dfrac{1-\hat\phi^2}{n} = \dfrac{1}{19.6^2} \Rightarrow n = 384.16(1-\hat\phi^2)$

so if $\hat\phi = 0.7$ we need $n = \lceil 384.16 \times 0.51 \rceil = \lceil 195.9 \rceil = 196$

3. For the following SARIMA models, write the models in their standard forms, and find the ACF and PACF using R (choose your own parameters). Describe the ACF and PACF behavior in words.

a. $SARIMA(1,0,0)\times(0,0,1)_{12}$

b. $SARIMA(0,0,1)\times(0,0,1)_{12}$

c. $SARIMA(1,0,0)\times(1,0,0)_{12}$

a. $SARIMA(1,0,0)\times(0,0,1)_{12}$

$$\begin{cases} \phi(B)\,\Phi(B^{12})\,Y_t = \theta(B)\,\Theta(B^{12})\,e_t \\ Y_t = X_t \end{cases}$$

$$\Rightarrow (1-\phi B)X_t = (1+\Theta B^{12})e_t$$

$$X_t - \phi X_{t-1} = e_t + \Theta e_{t-12}$$

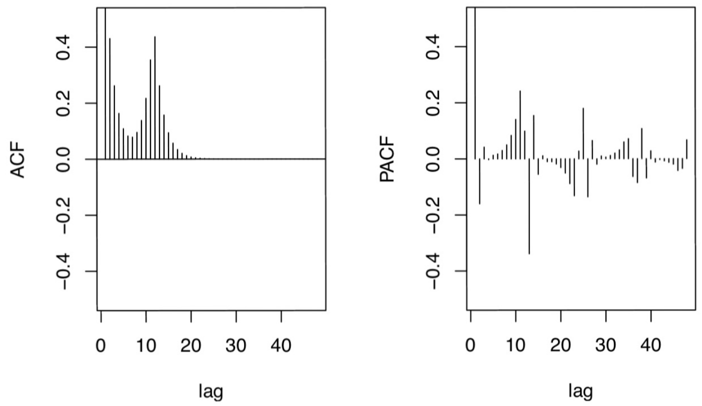
take $\phi = 0.4$, $\Theta = 0.8$ for example

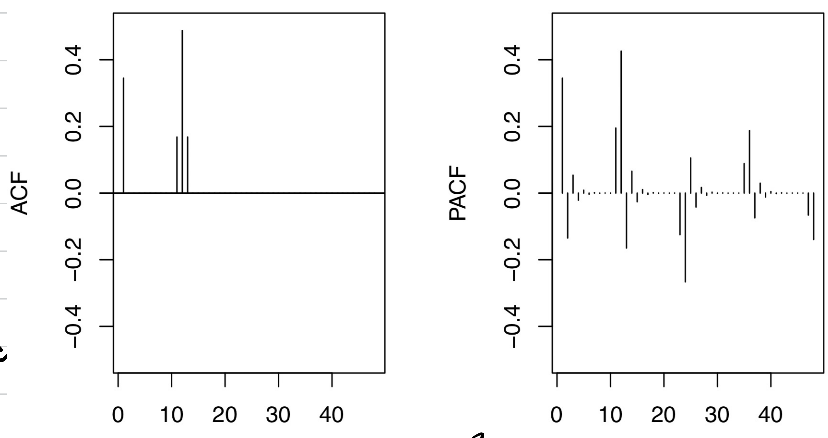
we can see ACF and PACF has large values at seasonal lag 12, within each period $d=12$. for non-seasonal part, ACF decays to zero, PACF cuts-of after lag 1

for seasonal part, ACF cuts-off after lag 12, PACF decays to zero.

3(a)

```
par(mfrow=c(1,2))
plot(ARMAacf(ar = c(0.6), ma = c(0.4, rep(0,10), 0.8),lag.max=48)[-1],type="h",ylim=c(-0.5,0.5),ylab="A
abline(h=0)
plot(ARMAacf(ar = c(0.6), ma = c(0.4, rep(0,10), 0.8),lag.max=48,pacf=T),type="h",ylim=c(-0.5,0.5),ylab
```





b. $SARIMA(0,0,1)\times(0,0,1)_{12}$

$$\begin{cases} \phi(B)\,\Phi(B^{12})\,Y_t = \theta(B)\,\Theta(B^{12})\,e_t \\ Y_t = X_t \end{cases}$$

$$\Rightarrow X_t = (1-\theta B)(1+\Theta B^{12})e_t$$

$$X_t = e_t - \theta e_{t-1} - \Theta e_{t-12} + \theta\Theta e_{t-13}$$

take $\theta = 0.4$, $\Theta = 0.8$, $\theta\Theta = 0.32$ for example

we can see ACF and PACF has large values at seasonal lag 12, within each period $d=12$. for both non-seasonal part and seasonal part

ACF cuts-of after lag 1, PACF decays to zero

c. SARIMA$(1,0,0) \times (1,0,0)_{12}$

$$\begin{cases} \phi_1(B)\,\Phi_1(B^{12})\,y_t = \theta_0(B)\,\Theta_0(B^{12})\,e_t \\ y_t = x_t \end{cases}$$

$$\Rightarrow (1-\phi B)(1-\Phi B^{12})X_t = e_t$$

$$X_t - \phi X_{t-1} - \Phi X_{t-12} + \phi\Phi X_{t-13} = e_t$$

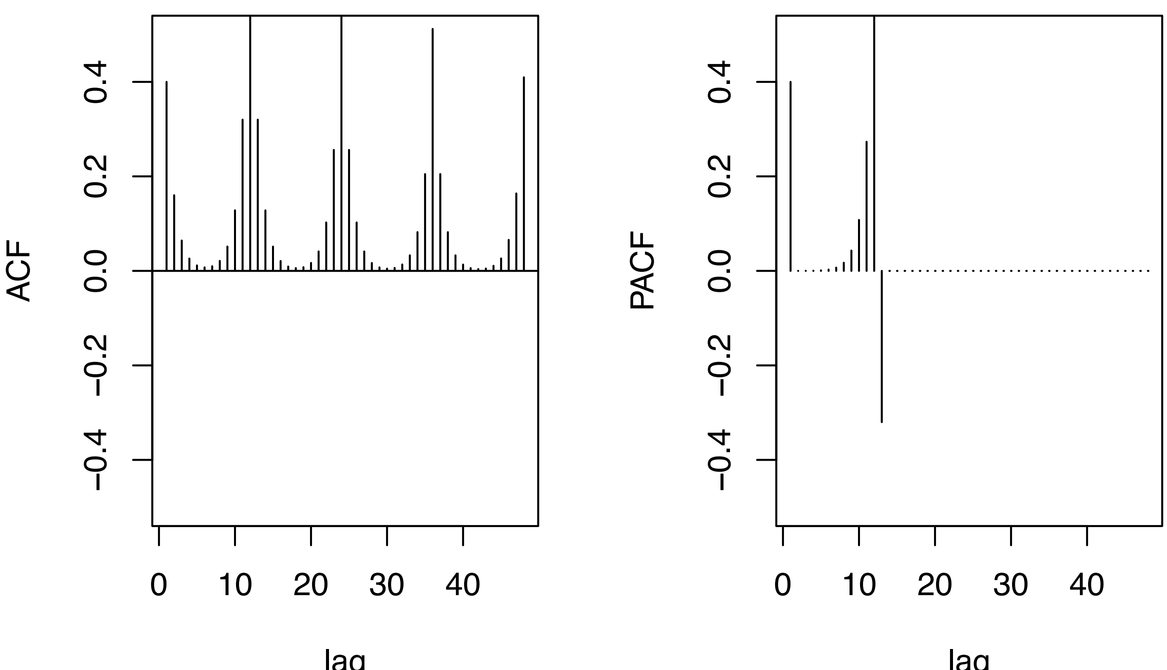
take $\phi = 0.4$, $\Phi = 0.8$ $\phi\Phi = 0.32$ for example

we can see ACF and PACF has large values at seasonal lag 12, within each period $d=12$, for both non-seasonal part and seasonal part

ACF decays to zero, PACF cuts-of after lag 1

3(c)

```
par(mfrow=c(1,2))
plot(ARMAacf( ar = c(0.4, rep(0,10), 0.8, -0.32), lag.max=48)[-1],type="h",ylim=c(-0.5,0.5),ylab="ACF",
abline(h=0)
plot(ARMAacf( ar = c(0.4, rep(0,10), 0.8, -0.32), lag.max=48,pacf=T),type="h",ylim=c(-0.5,0.5), ylab="P
```

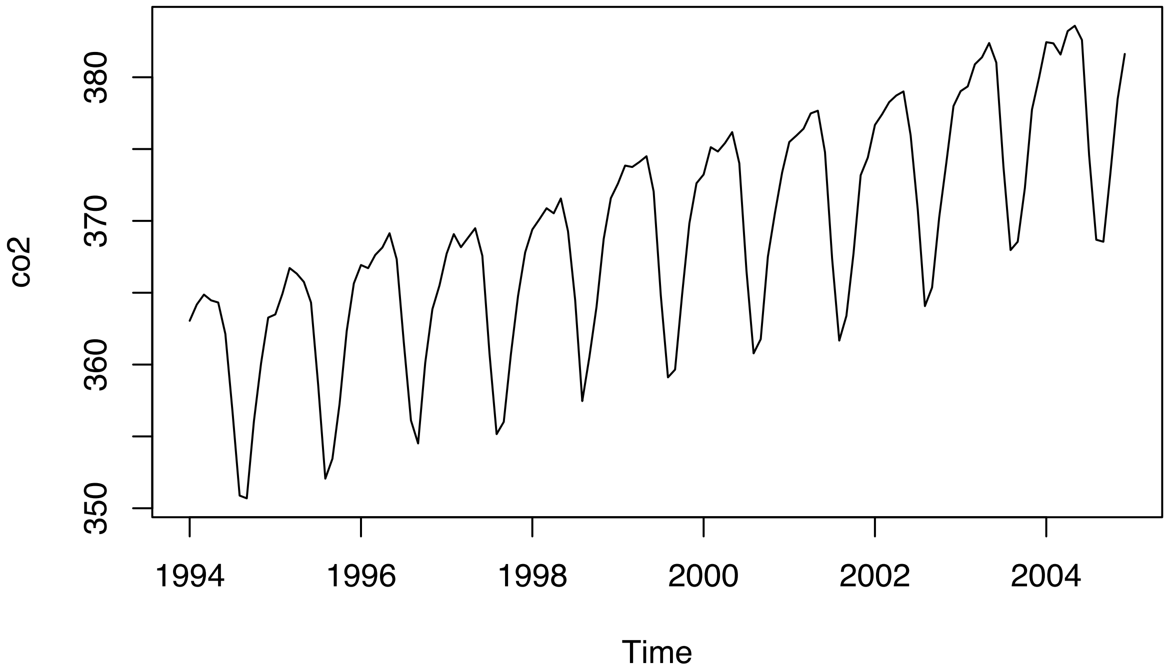


4.

4. Consider the co2 data in the dataset pacakage in R, which is Mauna Loa atmospheric CO2 Concentration. Set aside the last 24 observations as the test data and the rest as the training data.

(a) Plot the data and apply Box-Cox transformation, if necessary.

```r
data(co2)
# co2
plot(co2)
```



The stable patterns indicates that there is no need to apply Box-Cox transformation.

(b) Forecast 1: Use subset selection method to fit an ARIMA model to the data. Verify if the model is adequate. Forecast the 24 values along with the forecast intervals.

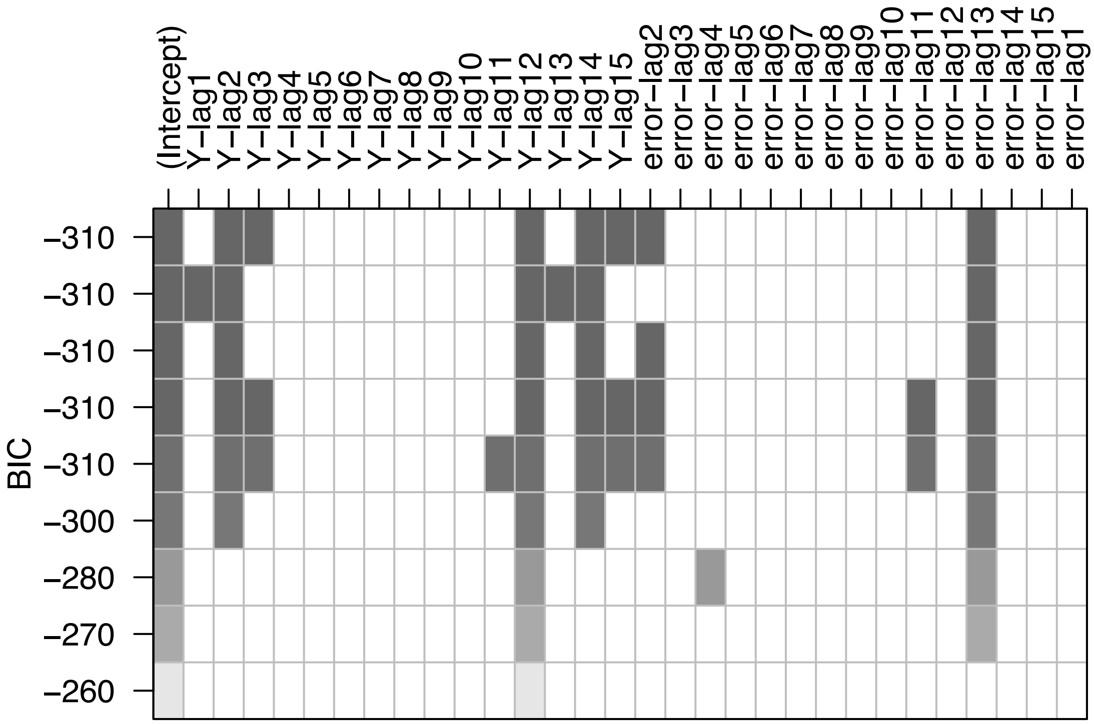
```r
# split the data
n <- length(co2)
train <- co2[1: (n-24)]
test  <- co2[(n-23): n]

# subset selection
set.seed(585)
fit0 <- armasubsets(train, nar = 15, nma = 15)
```

```
## Warning in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax, force.in =
## force.in, : 1 linear dependencies found

## Reordering variables and trying again:
```

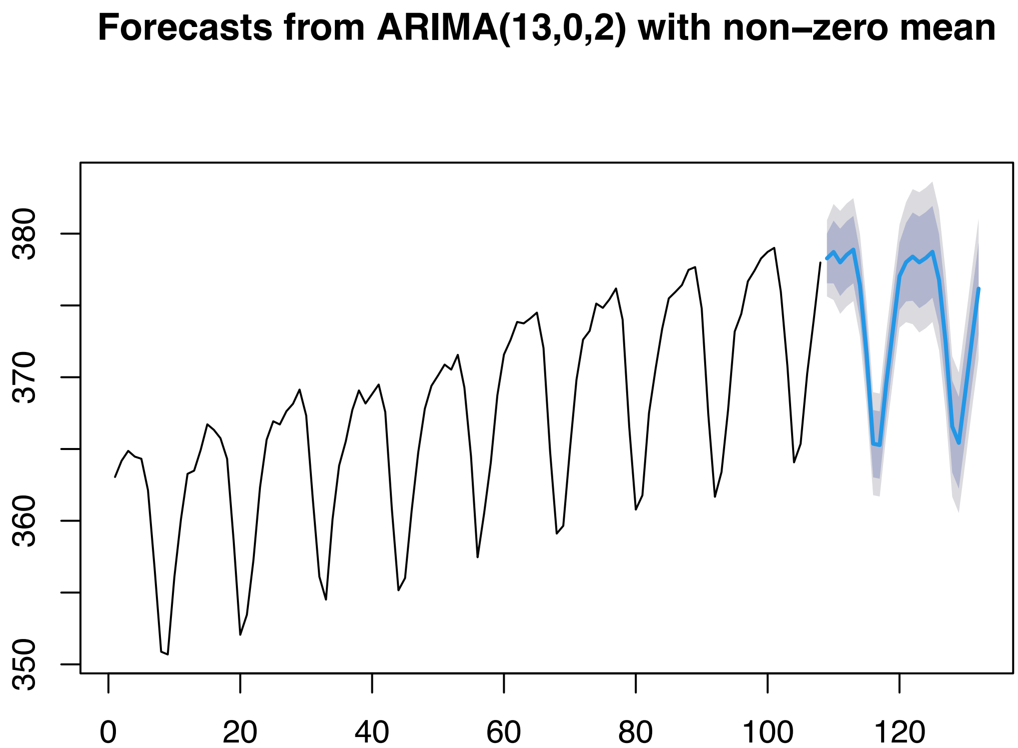
```r
plot(fit0)
```




```r
# based on the first line of BIC plot
fit1 <- Arima(train, order=c(13,0,2), fixed = c(0,rep(NA,2),rep(0,8),rep(NA,5)))
fit1
```

```
## Series: train
## ARIMA(13,0,2) with non-zero mean
##
## Coefficients:
##          ar1      ar2     ar3  ar4  ar5  ar6  ar7  ar8  ar9  ar10  ar11    ar12
##            0  -0.0473  0.0398    0    0    0    0    0    0     0     0  0.8098
## s.e.       0   0.0825  0.0530    0    0    0    0    0    0     0     0  0.0577
##           ar13     ma1     ma2      mean
##         0.1888  0.7656  0.5358  367.2276
## s.e.    0.0849  0.0895  0.1088    6.3304
##
## sigma^2 estimated as 1.834:  log likelihood=-192.1
## AIC=400.2   AICc=401.66   BIC=421.66
```

```r
fc1 <- forecast(fit1, h=24)
plot(fc1)
```



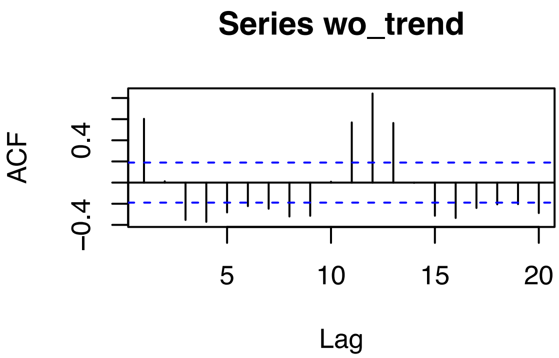

```
accuracy(fc1, test)
```

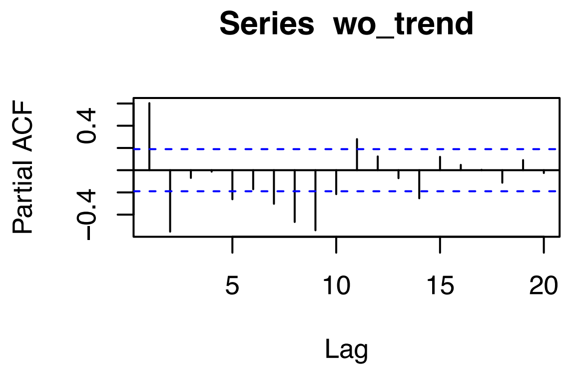
```
##                     ME     RMSE      MAE       MPE      MAPE      MASE
## Training set 0.6559915 1.309733 1.025558 0.1774836 0.2796191 0.3837414
## Test set     3.4955006 3.761315 3.495501 0.9235274 0.9235274 1.3079401
##                     ACF1
## Training set -0.04204224
## Test set              NA
```

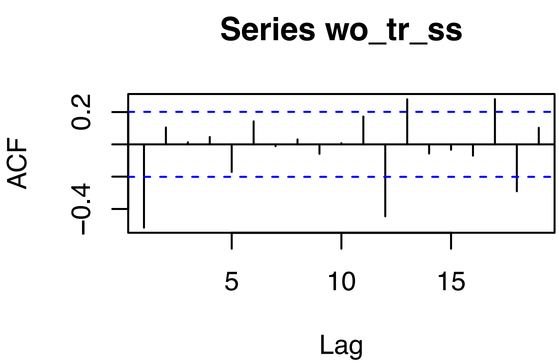
(b) Forecast 2: Now identify potential SARIMA models from ACF and PACF plots. Fit the candidate models and compare AICC to choose your final model. Use the model to Forecast the 24 values along with the forecast intervals.

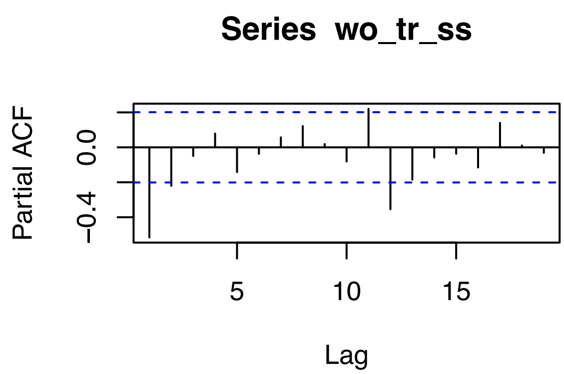
```
# eliminate the trend
wo_trend <- diff(train)
# eliminate the seasonal component
wo_tr_ss <- diff(wo_trend, 12)

par(mfrow=c(2,2))
acf(wo_trend)
pacf(wo_trend)
acf(wo_tr_ss)
pacf(wo_tr_ss)
```









Based on the ACF and PACF plots, I choose to compare the following models:

$SARIMA(3,1,0) \times (1,0,0)_{12}$

$SARIMA(1,1,1) \times (0,1,1)_{12}$

$SARIMA(1,1,0) \times (0,1,1)_{12}$

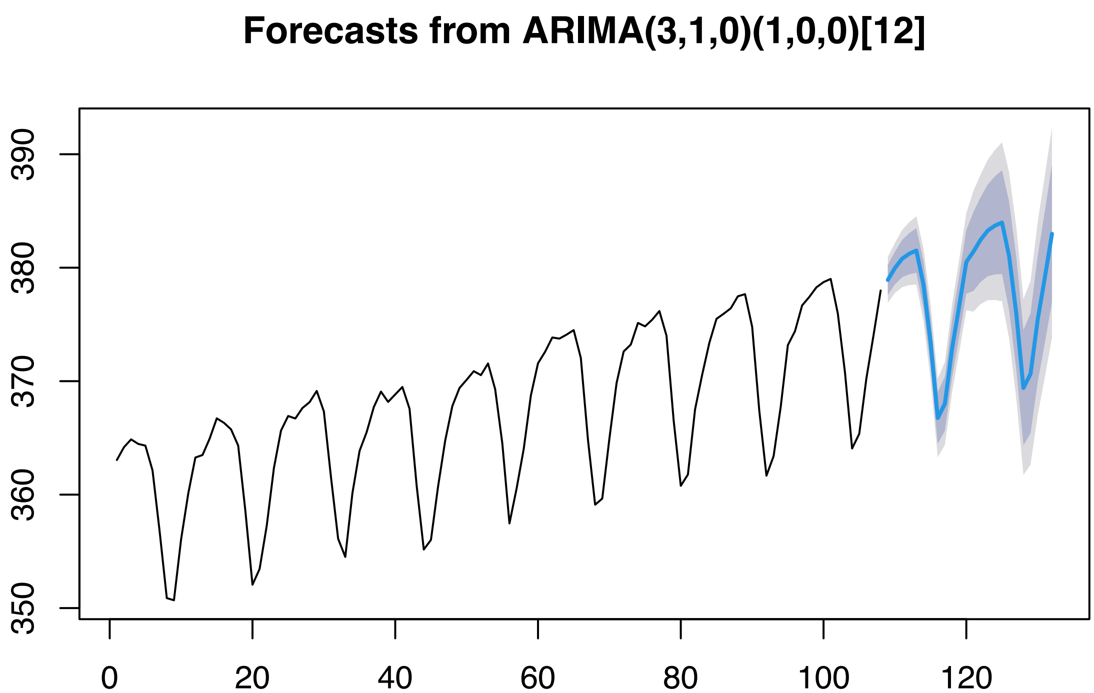
$SARIMA(0,1,1) \times (0,1,1)_{12}$

```r
fit2.1 <- Arima(train, order=c(3,1,0), seasonal=list(order=c(1,0,0),period=12), lambda=0)
fit2.2 <- Arima(train, order=c(1,1,1), seasonal=list(order=c(0,1,1),period=12), lambda=0)
fit2.3 <- Arima(train, order=c(1,1,0), seasonal=list(order=c(0,1,1),period=12), lambda=0)
fit2.4 <- Arima(train, order=c(0,1,1), seasonal=list(order=c(0,1,1),period=12), lambda=0)

which.min(c(fit2.1$aicc,fit2.2$aicc,fit2.3$aicc,fit2.4$aicc))
```

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

```r
fc2 <- forecast(fit2.1, h = 24)
plot(fc2)
```

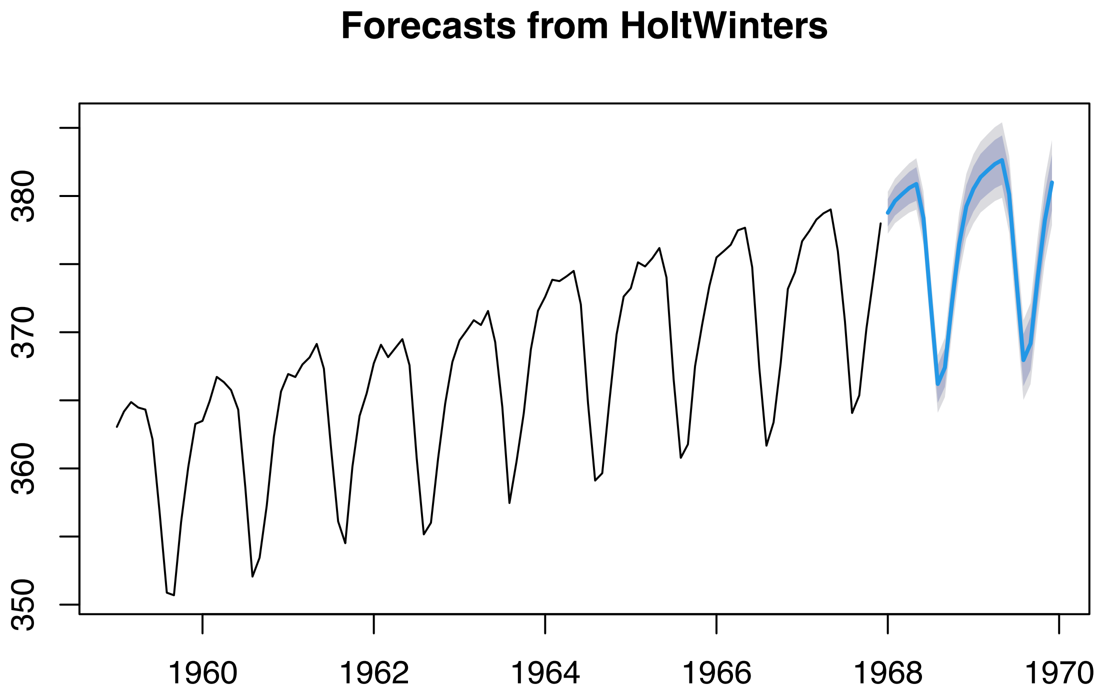



```
accuracy(fc2, test)
```

```
##                        ME      RMSE       MAE         MPE      MAPE      MASE
## Training set  0.05801004 0.9713285 0.7483229  0.01518981 0.2046792 0.2800061
## Test set     -0.12245416 1.1464115 0.9297257 -0.03329397 0.2466061 0.3478831
##                       ACF1
## Training set -0.0004915294
## Test set                NA
```

(b) Forecast 3: Use Holt-Winters seasonal forecasting method to predict the 24 values along with the forecast intervals.

```
train_ts <- ts(train, start = c(1959,1), frequency = 12)
fit3 <- HoltWinters(train_ts, seasonal='additive')
fc3 <- forecast(fit3, 24)
plot(fc3)
```




```r
accuracy(fc3, test)
```

```
##                      ME      RMSE       MAE         MPE      MAPE      MASE
## Training set 0.03293655 0.7820849 0.6110082 0.008599474 0.1665901 0.2286259
## Test set     0.84283622 1.1975182 0.9924440 0.222444572 0.2624521 0.3713509
##                    ACF1
## Training set 0.01306432
## Test set             NA
```

(c) Now complete the following table to compare between the forecasts: What is your conclusion?

Criteria	Forecast1	Forecast2	Forecast3
RMSE (Root Mean Squared Error)	3.761315	1.1464115	1.1975182
MAPE (Mean Average Percent Error)	1.3079401	0.2466061	0.3713509

Based on the results of accuracy, $SARIMA(3,1,0) \times (1,0,0)_{12}$ seems to be the best fit model.