Lab 5

Pstat 174/274

Seasonal ARIMA models

A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA models we have seen so far. It is written as follows:

$$\text{SARIMA} \quad \underbrace{(p,d,q)}_{\substack{\text{non-seasonal}\\ \text{part of}\\ \text{the model}}} \times \underbrace{(P,D,Q)_s}_{\substack{\text{seasonal}\\ \text{part of}\\ \text{the model}}}$$

$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t$ with $Z_t \sim WN(0, \sigma_Z^2)$ and $Y_t := \underbrace{(1-B)^d(1-B^s)^D}_{\text{differencing on original series}} X_t$

Example 1

Consider a SARIMA $(2,0,1) \times (1,1,1)_6$

a. Write the model's equation

- $p = 2$ then: $\phi(B) = (1 - \phi_1 B - \phi_2 B^2)$
- $d = 0$ then: $(1-B)^0 = 1$
- $q = 1$ then: $\theta(B) = (1 + \theta_1 B)$
- $s = 6$
- $P = 1$ then: $\Phi(B) = (1 - \Phi_1 B^6)$
- $D = 1$ then: $(1 - B^s)^D = (1 - B^6)^1$
- $Q = 1$ then: $\Theta(B) = (1 + \Theta_1 B^6)$

Finally, we write: $\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D Y_t = \theta(B)\Theta(B^s)Z_t$ as:

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_1 B^6)(1 - B^6)X_t = (1 + \theta_1 B)(1 + \Theta_1 B^6)Z_t$$

Expanding the terms you get:

$$\big(1 - \phi_1 B - \phi_2 B^2 - \Phi_1 B^6 + \phi_1 \Phi_1 B^7 + \phi_2 \Phi_1 B^8\big)(1 - B^6)X_t = (1 + \theta_1 B + \Theta_1 B^6 + \theta_1 \Theta_1 B^7)Z_t$$

b. How many parameters do you need to estimate for this model?

For the AR components: ϕ_1, ϕ_2, Φ_1, three; For the MA components: θ_1, Θ_1, two; And the white noise variance σ_Z^2, one. In total: 6.

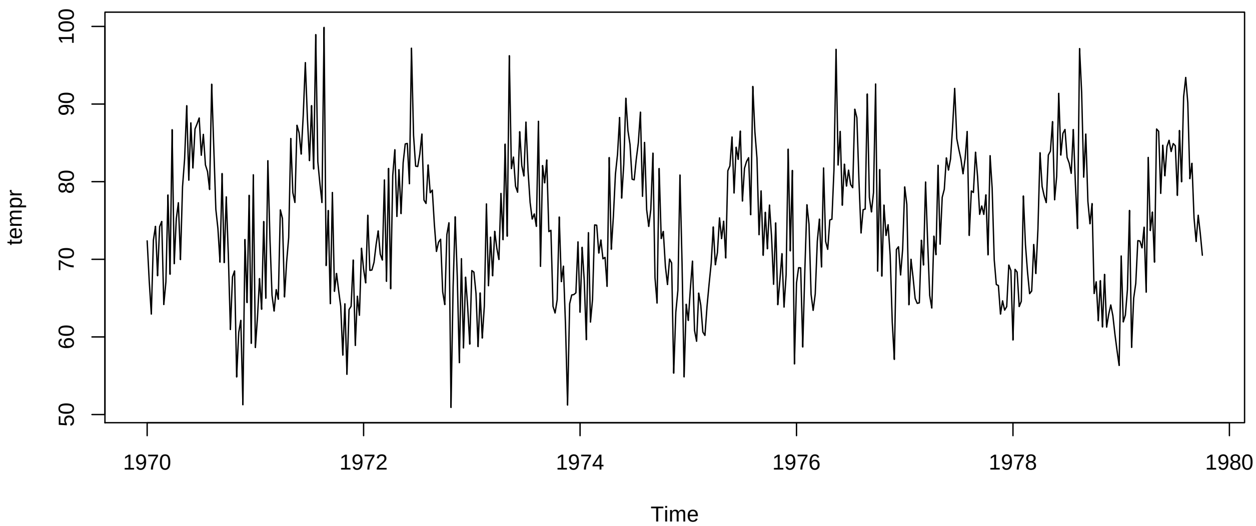
c. Let $Y_t := (1 - B^6)X_t$ and suppose you would like to fit an ARMA(8,7) to Y_t. How many parameters would you have to estimate?

In this case we need to estimate 8 terms for the AR component, 7 terms for the MA component and the white noise variance. In total: 16.

Example 2: Some time series with seasonal cycles

Case 1: LA temperature measured from 1970 to 1980

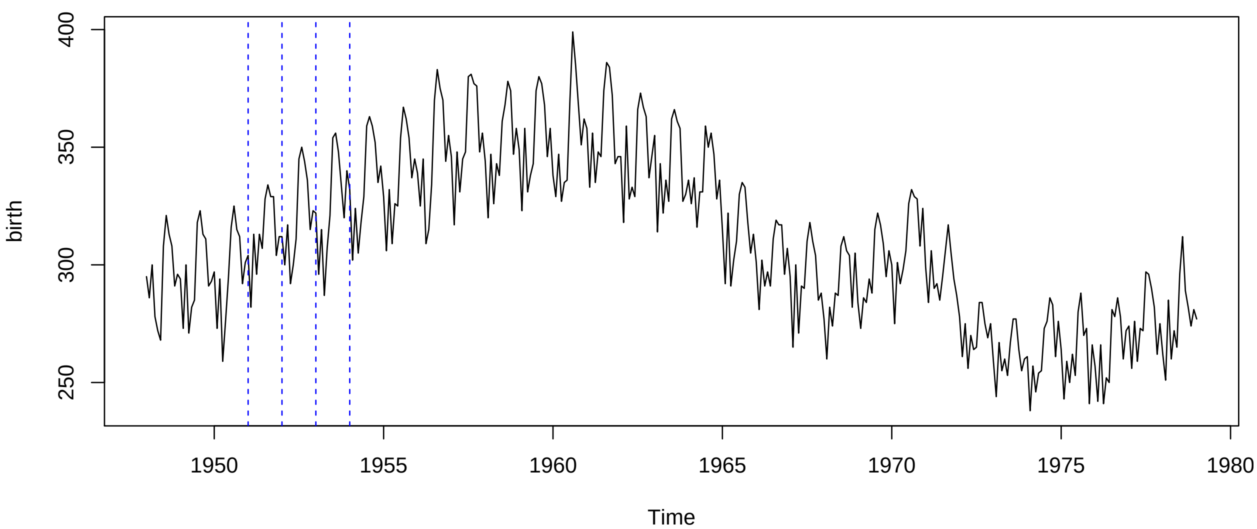
```
library(astsa)
plot(tempr) # LA temperature measured from 1970 to 1980
```



Case 2: Monthly live births (adjusted) in thousands for the United States, 1948-1979.

The dotted blue lines in the figure highlights where we can identify seasonal cycles in the US Monthly Brith Rates. The red vertical lines high

```
library(astsa)
plot(birth) # US Monthly birth rates from 1948 to 1979
abline(v = ts(c(1951,1952,1953,1954)), col = "blue", lty = 2)
```



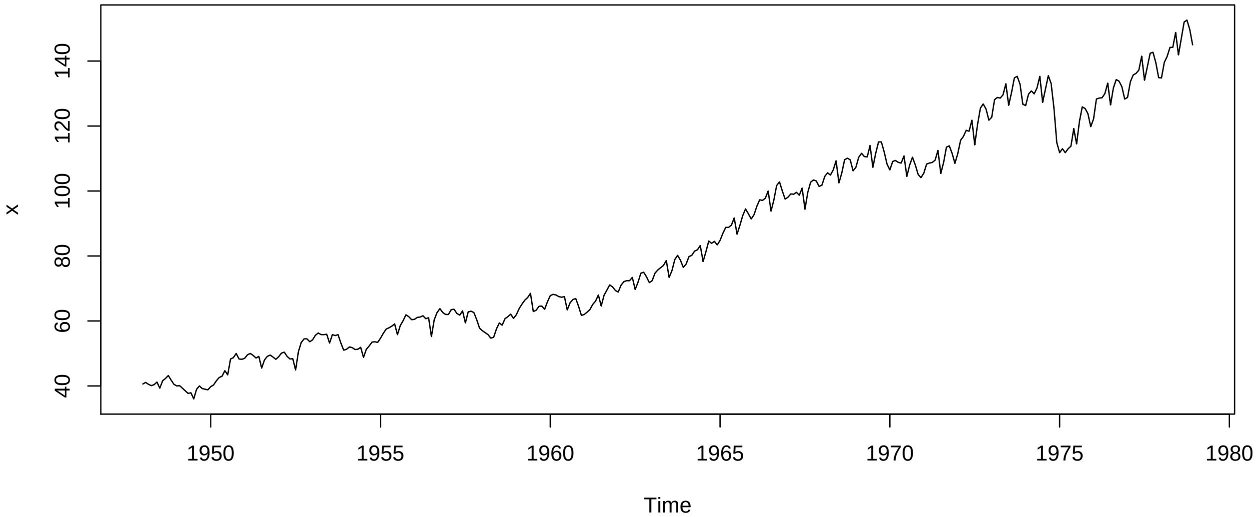
Example 3: Federal Reserve Board Index

In this case we work with the Monthly Federal Reserve Board Production Index (1948-1978, $n = 372$ months).

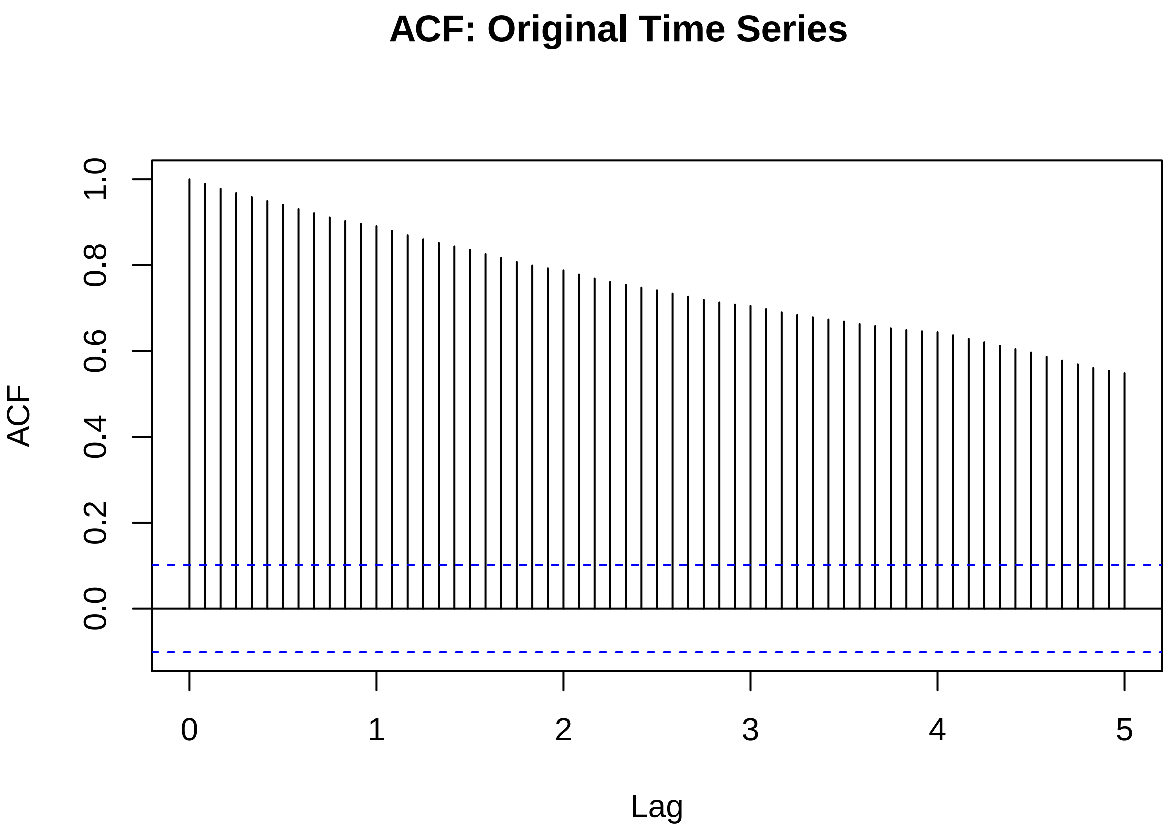
a. Get the production time series and plot it.

The resulting figures indicates the presence of a positive time trend.

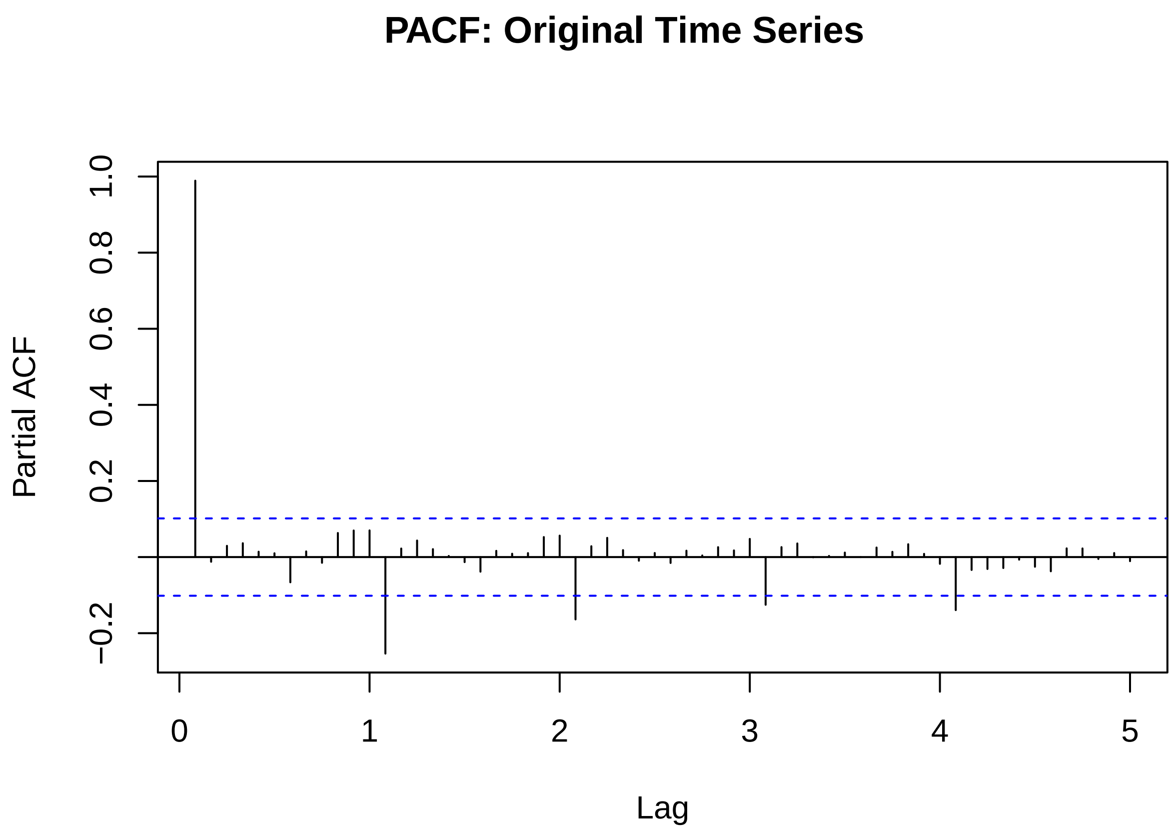
```r
x = prodn
plot(x)
```




```r
#op <- par(mfrow = c(2,1))
acf(x,lag.max = 60,main = "")
title("ACF: Original Time Series", line = -1, outer=TRUE)
```




```r
pacf(x,lag.max = 60,main = "")
title("PACF: Original Time Series", line = -1, outer=TRUE)
```

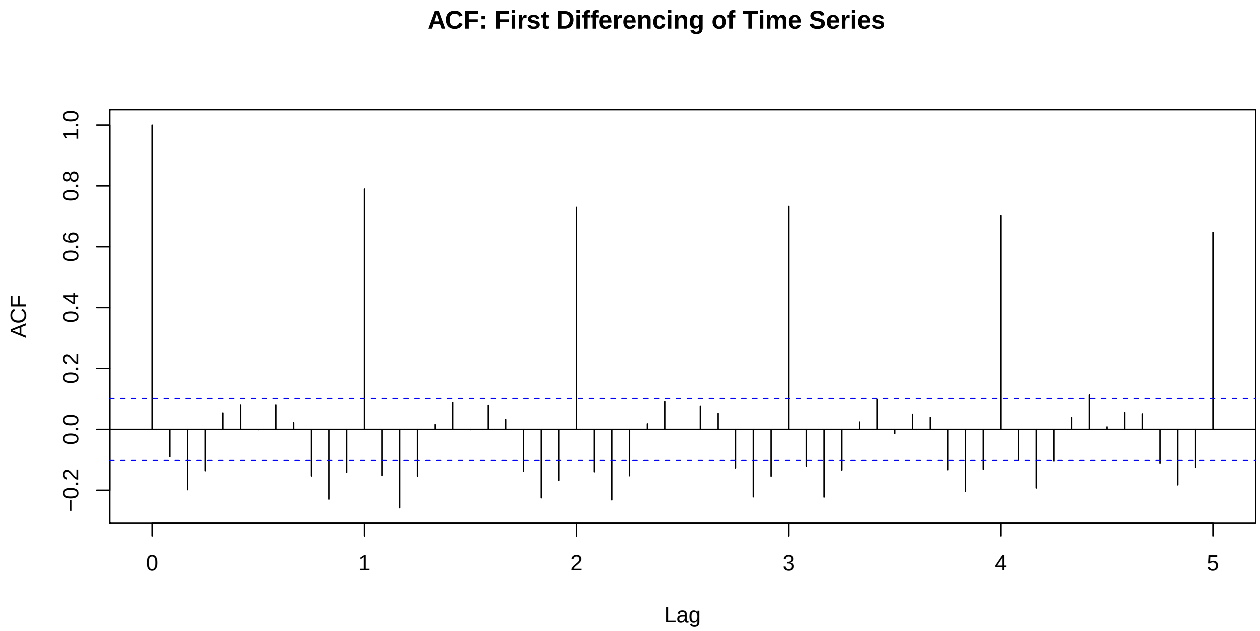



```
#par(op)
```

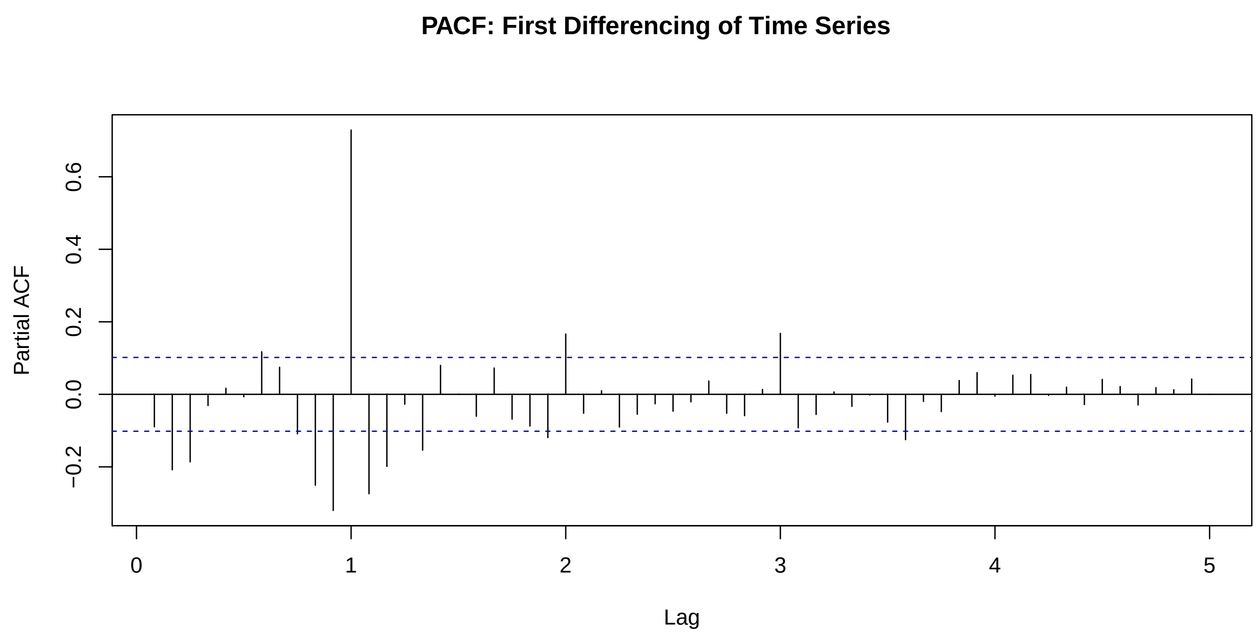
c. Apply a first differencing and plot ACF and PACF

We plot ACF and PACF plots for $Y_t = \nabla X_t = (1 - B)X_t$, hoping to remove the time trend with the first differencing.

```
y_1 = diff(x, 1)
acf( y_1, lag.max = 60, main = "")
title("ACF: First Differencing of Time Series", line = -1, outer = TRUE)
```




```
pacf( y_1, lag.max = 60, main = "")
title("PACF: First Differencing of Time Series", line = -1, outer = TRUE)
```

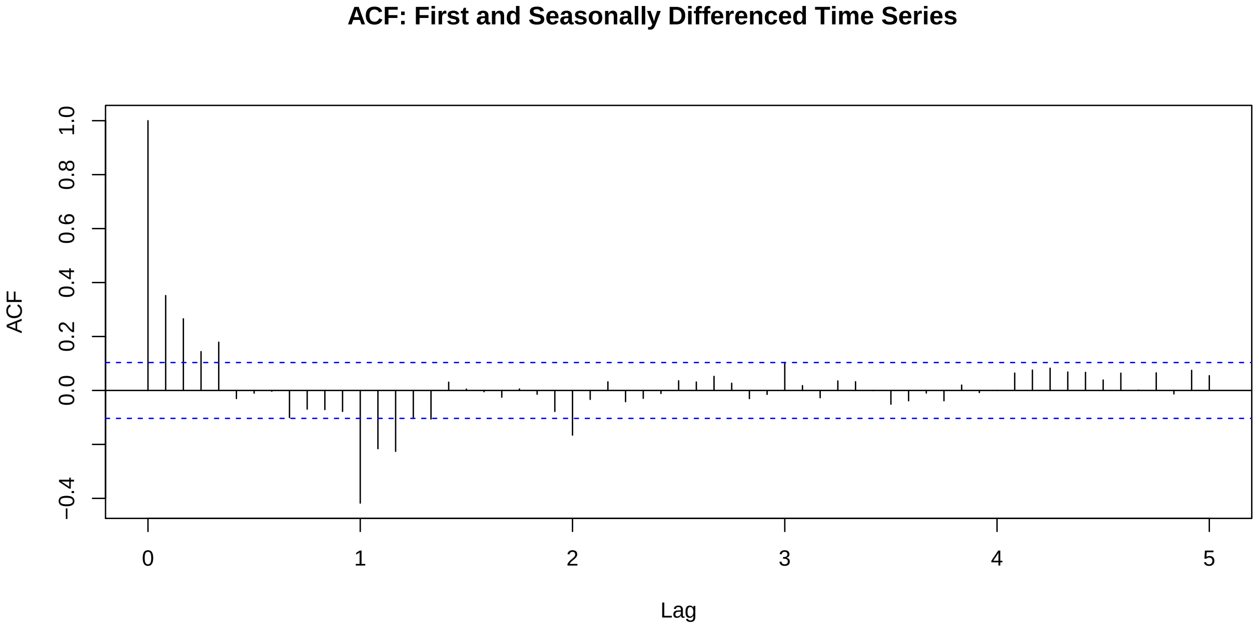


Note the peaks at seasonal lags, $h = 1s, 2s, 3s, 4s$ where $s = 12$ (i.e., $h = 12$, 24, 36, 48) with relatively slow decay in the ACF plot suggests a seasonal difference $(1 - B^s)$ with $s = 12$.

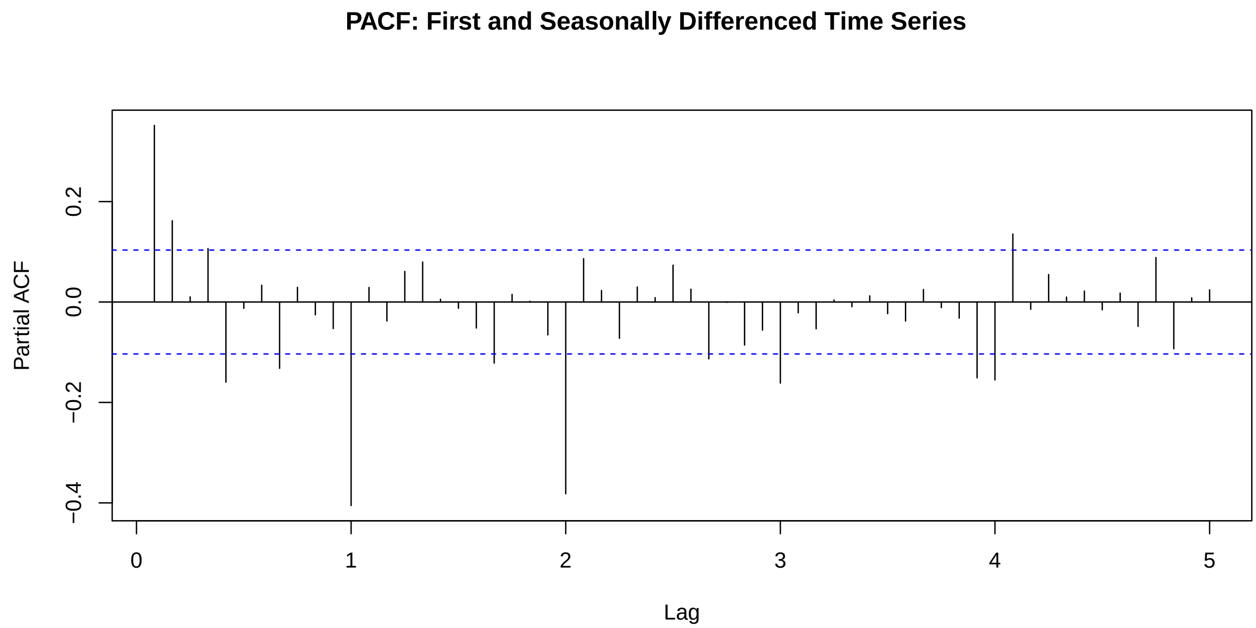
d. Apply a first seasonal differencing and plot ACF and PACF

In this case, we work with $Y_t = \nabla_{12}\nabla X_t = (1 - B^{12})(1 - B)X_t$

```
y_12 = diff(y_1, 12)
acf( y_12, lag.max = 60, main = "")
title("ACF: First and Seasonally Differenced Time Series", line = -1, outer = TRUE)
```




```r
pacf( y_12, lag.max = 60, main = "")
title("PACF: First and Seasonally Differenced Time Series", line = -1, outer = TRUE)
```



e. Based on part d), suggest some models to fit.

Modeling the seasonal part (P, D, Q): For this part, focus on the seasonal lags $h = 1s, 2s$, etc.

- We applied one seasonal differencing so $D = 1$ at lag $s = 12$.
- The ACF shows a strong peak at $h = 1s$ and smaller peaks appearing at $h = 2s, 3s$.

  A good choice for the MA part could be $Q = 1$ or $Q = 3$.
- The PACF shows two stronk peaks at $h = 1s, 2s$ and smaller peaks at $h = 3s, 4s$.

  A good choice for the AR part could be $P = 2$ or $P = 4$.

Modeling the non-seasonal part (p , d, q): In this case focus on the within season lags, $h = 1,\ldots,11$.

- We applied one differencing to remove the trend: $d = 1$
- The ACF seems to be tailing off. Or perhaps cuts off at lag.

  A good choice for the MA part could be $q = 0$ or $q = 4$ respectively.
- The PACF cuts off at lag h=2.

  A good choice for the AR part could be $p = 2$.

f. Fit a couple of candidate models

As an illustration we fit the following two models:

i. SARMA $(p = 2, d = 1, q = 0) \times (P = 2, D = 1, Q = 1)_{s=12}$
ii. SARMA $(p = 2, d = 1, q = 4) \times (P = 4, D = 1, Q = 3)_{s=12}$

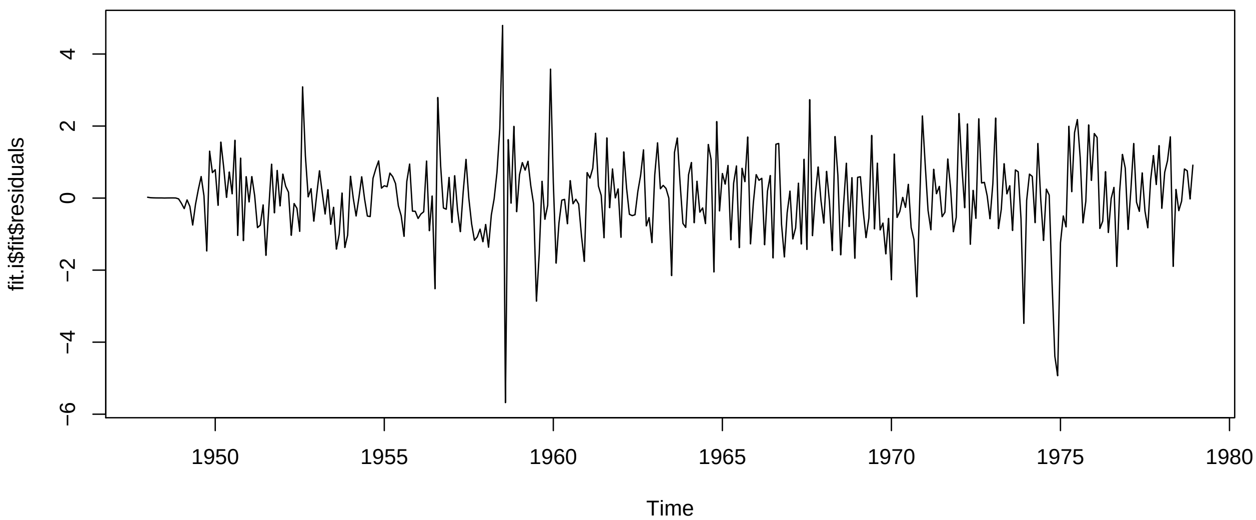
```
fit.i <- sarima( xdata = prodn,
                 p = 2, d = 1, q = 0,
                 P = 2 , D = 1, Q = 3, S = 12,
                 details = F)
print('Coefficients')
```

```
## [1] "Coefficients"
```

```
fit.i$fit$coef
```

```
##        ar1        ar2       sar1       sar2       sma1       sma2
##  0.3023102  0.1125150  0.0241631  0.1987099 -0.7607446 -0.3206817
##       sma3
##  0.4091256
```

```
plot(fit.i$fit$residuals)
```




```
fit.ii <- sarima( xdata = prodn,
                  p = 2, d = 1, q = 4,
                  P = 4 , D = 1, Q = 3, S = 12,
                  details = F)

print('Coefficients')
```

```
## [1] "Coefficients"
```

```r
fit.ii$fit$coef
```

```
##         ar1         ar2         ma1         ma2         ma3         ma4
## -0.63074026 -0.39527565  0.95307388  0.79678450  0.28870963  0.22697521
##        sar1        sar2        sar3        sar4        sma1        sma2
## -0.03318587 -0.45244730 -0.25475241 -0.27015453 -0.69571212  0.29637713
##        sma3
##  0.21573810
```

```r
plot(fit.ii$fit$residuals)
```