L8 Revision

June 1, 2020

0.1 Lecture 8: More Diagnostics & Intro to Forecasting

Diagnostics so far: * ACF plots: * If linear --> Non-stationary * Can check to see if process is white noise -> $r_t \sim \text{NID}(0, \frac{1}{T})$ -> $\pm \frac{2}{\sqrt{T}}$ * Can check to see if MA(q) process is within confidence intervals -> $\pm 2\sqrt{\frac{1+2(r_1^2+r_2^2+...+r_q^2)}{T}}$

`* Use ACF to estimate q`

- PACF plots:
  - Can read off the p of an AR(p) process -> $\hat{\phi}_{uu} \sim \text{NID}(0, \frac{1}{T})$ -> $\hat{\phi}_{uu} \pm \frac{2}{\sqrt{T}}$
  - If both ACF and PACF plots are geometric, then you're dealing with an ARMA process
- AIC

Outline for the additional diagnostics: * Overfitting * Residuals * ACF of residuals * Ljung-Box statistic (portmanteau statistic)

1 1. Overfitting

- Once we have identified what we believe is an adequate model, **we fit another model with one extra parameter**, and check whether the estimate for the additional parameter differs significantly from zero.
- If the additional parameter does not differ significantly from zero, then we have additional evidence to support our model.
- Get a **t-statistic** via $\frac{\text{parameter estimate}}{\text{standard error in the parameter estimate}}$
- Compare that **t-statistic** with the **t-distribution** with **T - number of parameters - 1 for the mean - 1 extra ?**

1.0.1 Example

In previous lecture, we were confident that an AR(1) model was appropriate to model a given process, so we checked with an AR(2) model.

The second AR parameter: $\phi_2 = -0.1235$ with a standard error of 0.0892

We get a t-stat of -1.38:

```
In [1]: t <- -0.1235 / 0.0892
        t
```

-1.38452914798206

$T = 124$, and so in this three parameter model (μ, ϕ_1, ϕ_2), we have $t - 4 = 120$ degrees of freedom.

```
In [6]: # this is a symmetric two-tailed test, so the p-value is 0.1686
        2*pt(t, 124, lower.tail=TRUE)
```

0.168682876019897

With a p-value of 0.17, we cannot reject the null hypothesis that the observed parameter is different to zero, therefore we accept the alternative that they are the same.

The secondary piece of diagnostic information we had was that the AIC has increased as we went from AR(1) to AR(2).

2 2. Residuals

To examine the adequecy of the fitted model, examine the residuals / standardised residuals.

2.1 Residuals should appear to be similar to white noise observations:

- mean zero;
- constant variance;
- pairwise uncorrelated

Slightly stronger assumption for the distribution of white noise: > $\epsilon_t \sim \mathrm{NID}(0, \sigma^2)$.
We expect the residuals to be approximately normally distributed.

2.2 Residual definition

Residuals, e_t: > $e_t = y_t - \hat{y}_t$,

where $\hat{y}_t$ is the fitted value of y_t at t. You get $\hat{y}_t$ from the systematic part of the model equation, which for AR(1) would be:

$$\hat{y}_t - \hat{\mu} = \hat{\phi}(y_{t-1} - \hat{\mu})$$

For a general ARIMA model, the residuals e_t are not always easy to express in simple terms.

Since the residuals, e_t are assumed to be $e_t \sim \mathrm{NID}(0, \sigma^2)$, then you can center and scale the residuals by subtracting the mean and dividing through by the standard deviation. The mean is assumed to be zero, so we just divide e_t by the square root of our estimate for the variance, which R produces.

$$d_t = \frac{e_t}{\hat{\sigma}}.$$

2.2.1 We plot:

```
1. The residuals / standardised residuals;
2. The ACF of the residuals,
```

As diagnostic checks to see whether the residuals are indeed behaving like a white noise process.

Recall: The ACF of a white noise process should have 19/20 plots within the 95% confidence interval plotted by R.

We're also looking to see whether there's trend in the residuals, and whether the residuals appear to be normally distributed in a qq-plot.

In general, $\frac{1}{T}$ provides an upper bound on the variances of the r_τ, and therefore the standard error is estimated as $\frac{1}{\sqrt{T}}$.

3 3. Portmanteau Statistics

Null hypothesis: the model is correct

Can produce a statistic based on the sample autocorrelations **of the residuals**.

I.e. fit a model, and then compare observed values with fitted values, to produce a series of residuals.

If the fitted model is an **ARIMA(p,q)** model, then **if the model is correct**, then:

$$Q_K^* = T\sum_{\tau=1}^K r_\tau^2,$$

The Box-Pierce test statistic, follows a χ^2_{K-p-q} distribution:

$$Q_K^* \sim \chi^2_{K-p-q}$$

3.1 Improve statistic: Modified Box-Pierce statistic / Ljung-Box statistic

$$Q_K = T(T+2)\sum_{\tau=1}^K \frac{r_\tau^2}{T-\tau},$$

which is closer to the χ^2_{K-p-q} distribution.

3.2 Diagnostics

YOU GET A Ljung-Box STATISTIC FOR EACH LAG

If any of the chi-square values is significant, then we may reject the null hypothesis that the model is correct.

3.3 Bread Price Diagnostics

```
In [14]: price <- c(5.8, 6.1, 5.4, 6.2, 5.0, 4.6, 5.8, 5.1, 5.3, 5.1, 4.8, 5.3, 6.8, 9.0, 8.6,
         9.0, 7.4, 6.4, 4.8, 3.9, 3.9, 5.6, 5.7, 7.5, 7.3, 7.4, 7.5, 9.7, 6.1, 6.0, 5.7, 5.0,
         4.2, 4.6, 5.9, 5.4, 5.4, 5.4, 5.6, 7.6, 7.4, 5.4, 5.1, 6.9, 7.5, 5.9, 6.2, 5.6, 5.8,
         5.6, 6.6, 4.8, 5.2, 4.5, 4.4, 5.3, 5.0, 6.4, 7.8, 8.5, 5.6, 7.1, 7.1, 8.0, 7.3, 5.7,
```

```
        4.8, 4.3, 4.4, 5.7, 4.7, 4.1, 4.1, 4.7, 7.0, 8.7, 6.2, 5.9, 5.4, 6.3, 4.9, 5.5, 5.4,
        4.7, 4.1, 4.6, 4.8, 4.5, 4.7, 4.8, 5.4, 6.0, 5.1, 6.5, 6.2, 4.6, 4.5, 4.0, 4.1, 4.7,
        5.1, 5.2, 5.3, 4.8, 5.0, 6.2, 6.4, 4.7, 4.1, 3.9, 4.0, 4.9, 4.9, 4.8, 5.0, 4.9, 4.9,
        5.4, 5.6, 5.0, 4.5, 5.0, 7.2, 6.1)
```

```
In [15]: # use time series function to transform it into time series format (auto indexes the
         BP.ts <- ts(price, start=1634, frequency = 1)
```

```
In [22]: ar.1 <- arima(BP.ts, order=c(1,0,0))
         ar.1
```

```
Call:
arima(x = BP.ts, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.6429     5.6608
s.e.  0.0678     0.2307

sigma^2 estimated as 0.8655:  log likelihood = -167.26,  aic = 340.52
```

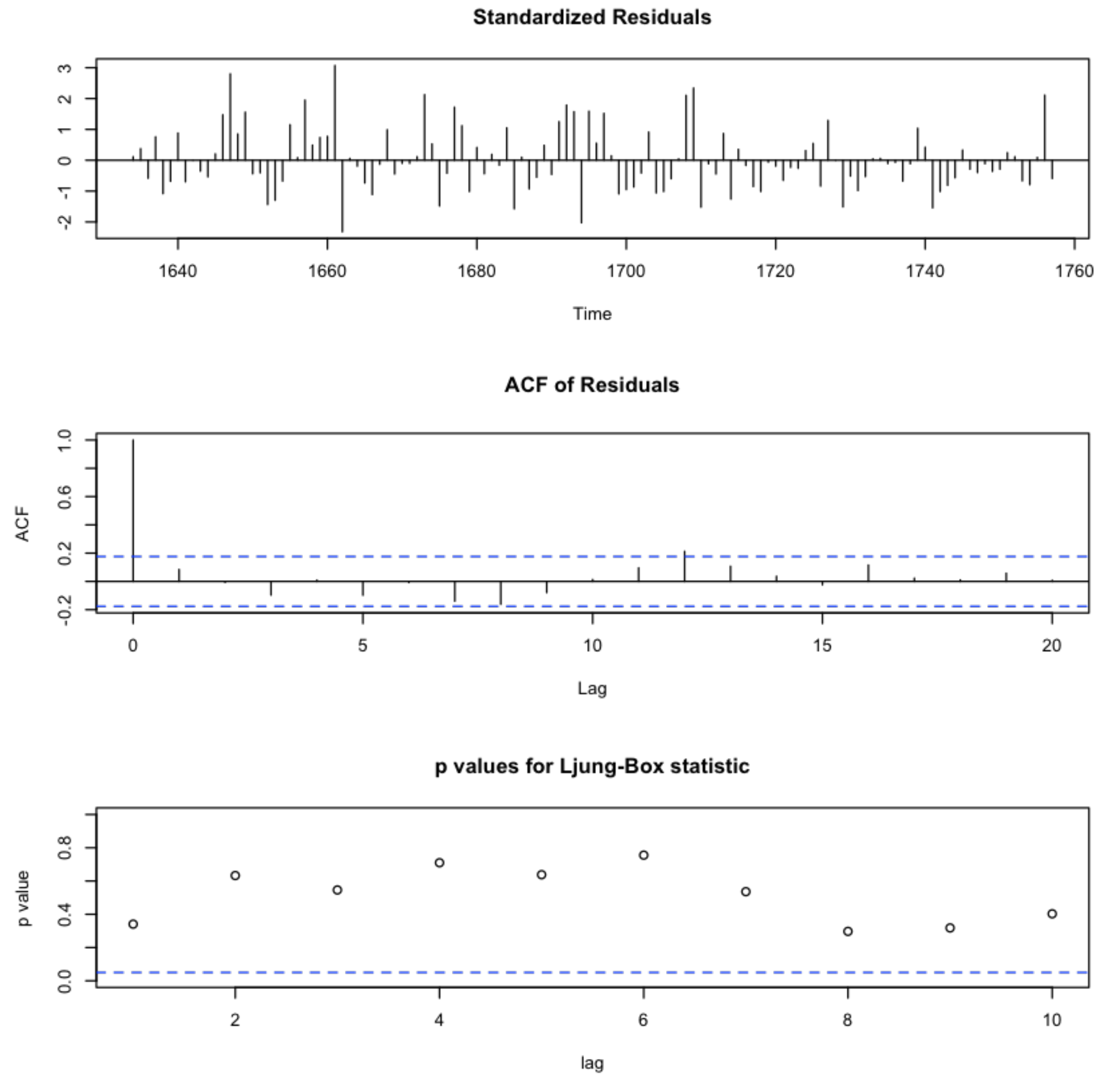
```
In [28]: ar.2 <- arima(BP.ts, order=c(1,0,0))
         ar.2
```

```
Call:
arima(x = BP.ts, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.6429     5.6608
s.e.  0.0678     0.2307

sigma^2 estimated as 0.8655:  log likelihood = -167.26,  aic = 340.52
```

```
In [29]: tsdiag(ar.1)
```

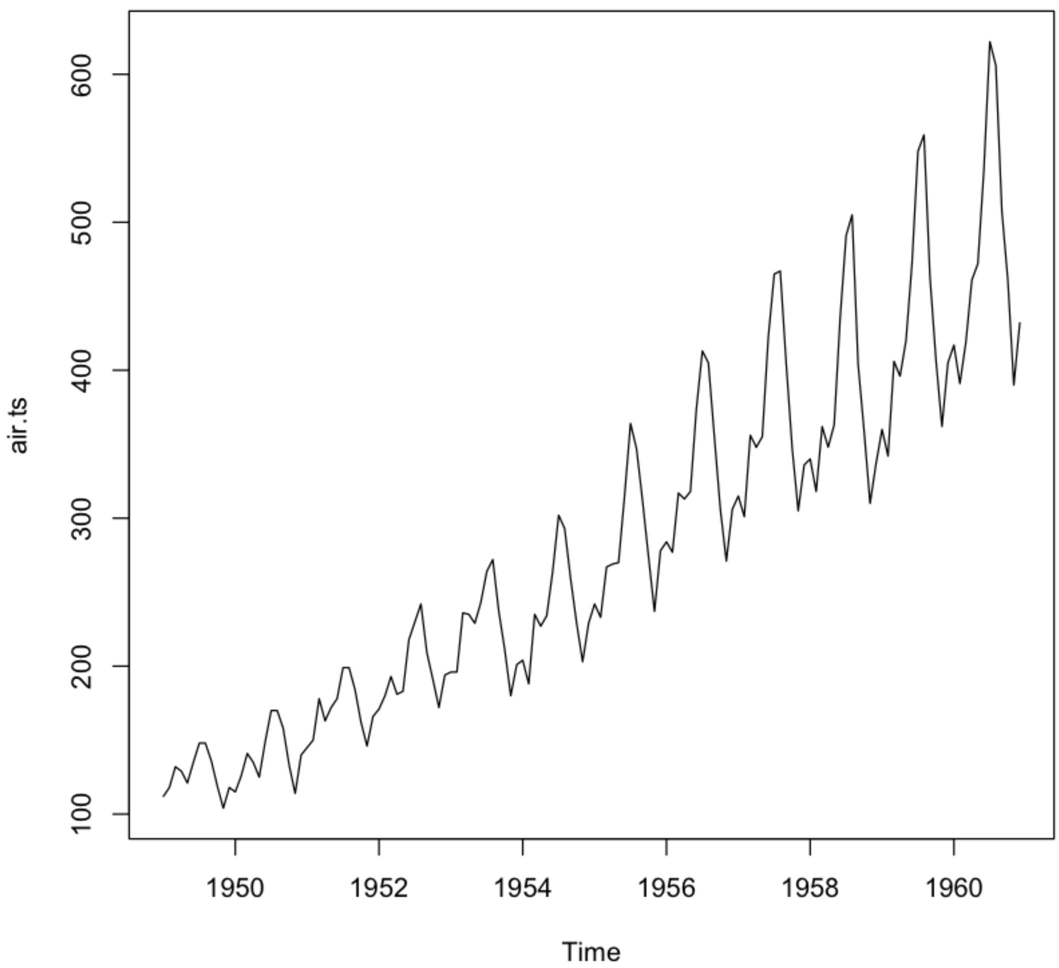


4 Modelling Air Passenger Data

```
In [31]: passengers<-c(112, 118, 132, 129, 121, 135, 148, 148, 136, 119, 104, 118, 115, 126,
                       +141, 135, 125, 149, 170, 170, 158, 133, 114, 140, 145, 150, 178, 163,
                       +199, 199, 184, 162, 146, 166, 171, 180, 193, 181, 183, 218, 230, 242,
                       +172, 194, 196, 196, 236, 235, 229, 243, 264, 272, 237, 211, 180, 201,
                       +235, 227, 234, 264, 302, 293, 259, 229, 203, 229, 242, 233, 267, 269,
                       +364, 347, 312, 274, 237, 278, 284, 277, 317, 313, 318, 374, 413, 405,
                       +271, 306, 315, 301, 356, 348, 355, 422, 465, 467, 404, 347, 305, 336,
                       +362, 348, 363, 435, 491, 505, 404, 359, 310, 337, 360, 342, 406, 396,
                       +548, 559, 463, 407, 362, 405, 417, 391, 419, 461, 472, 535, 622, 606,
                       +390, 432)
```

```
In [32]: air.ts <- ts(passengers, start = 1949, frequency = 12)
```

```
In [34]: plot(air.ts)
```

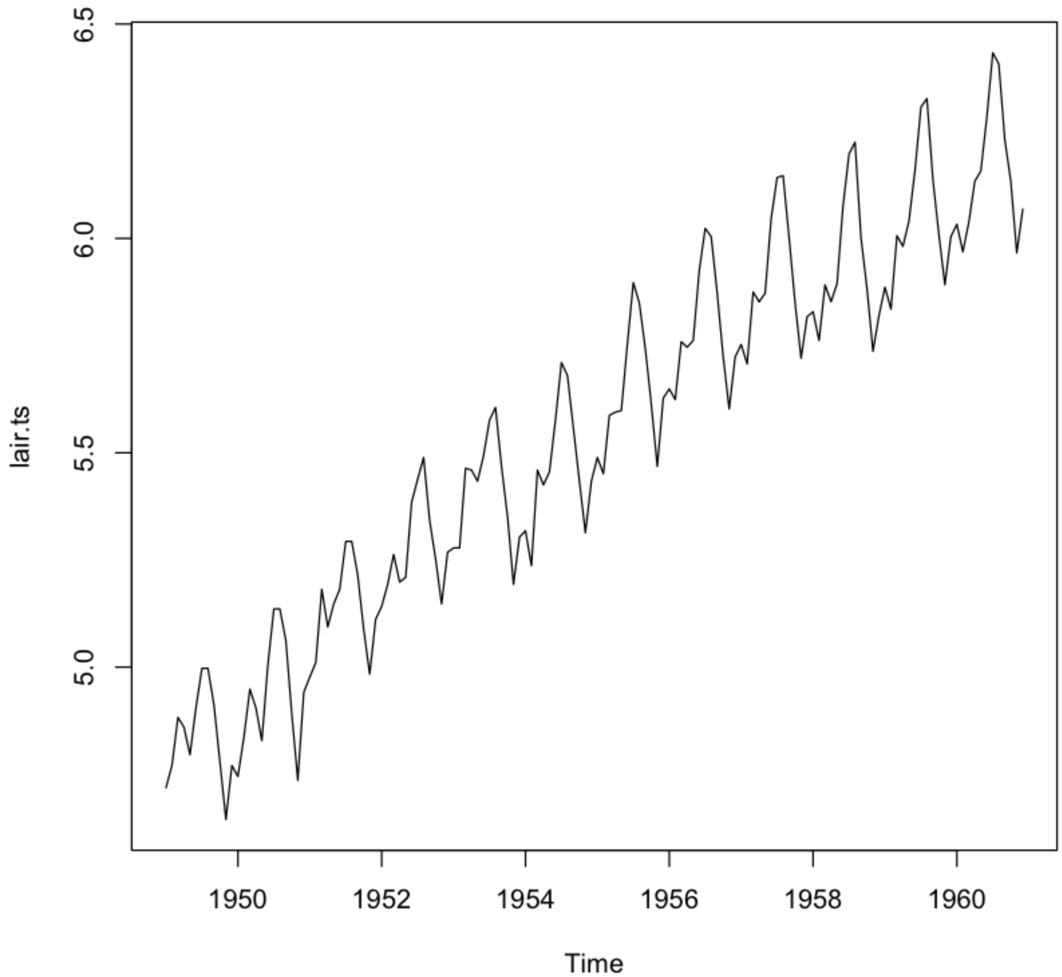


4.1 Firstly, logging the process

We see the seasonal trend is increasing in variance, so logging this will make the variance more consistent.

```
In [37]: lair.ts <- log(air.ts)
```

```
In [38]: plot(lair.ts)
```



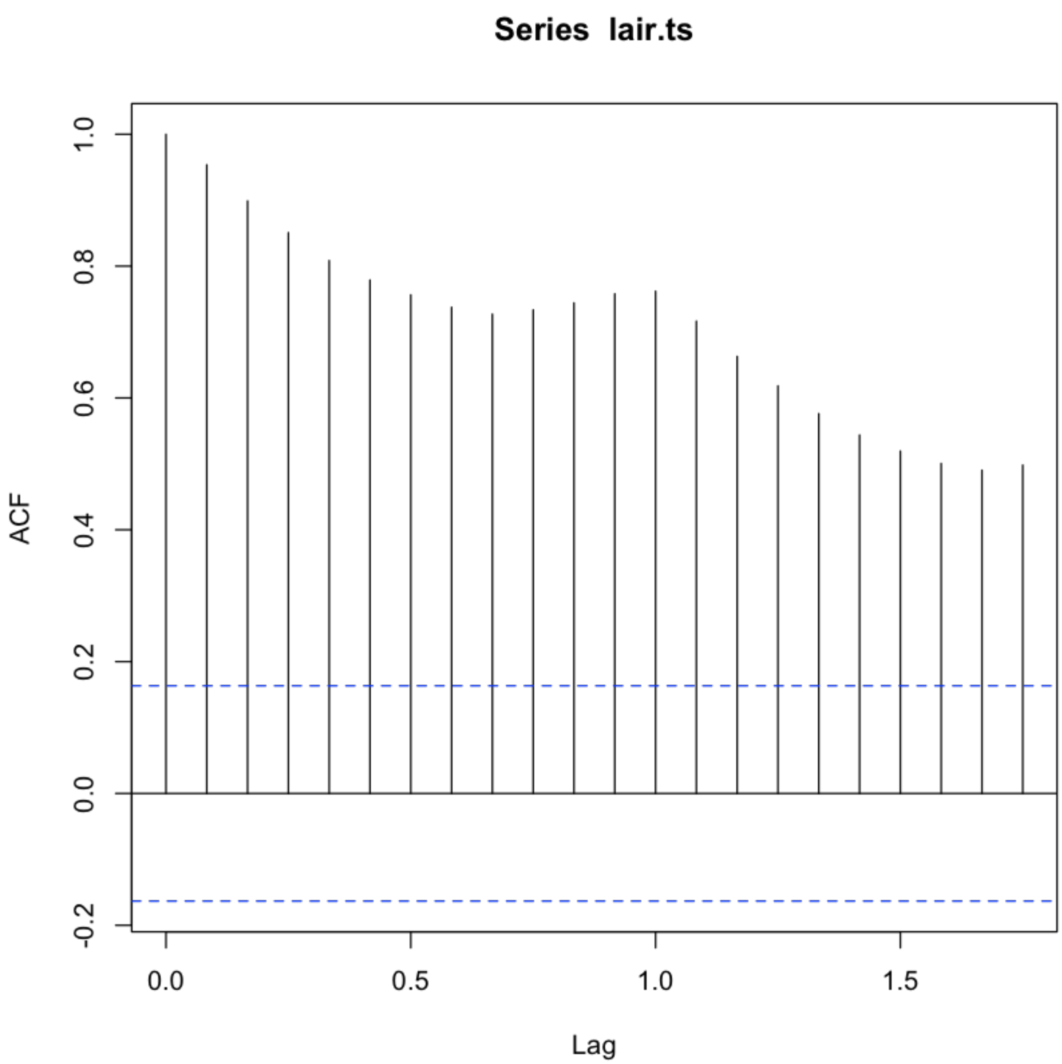

```
In [39]: length(lair.ts)
```

144

4.2 Looking at ACF of $\{\log Y_t\}$

- Linear decrease trend - signature of non-stationary process.
- Will need to difference

```
In [42]: acf(lair.ts)
```



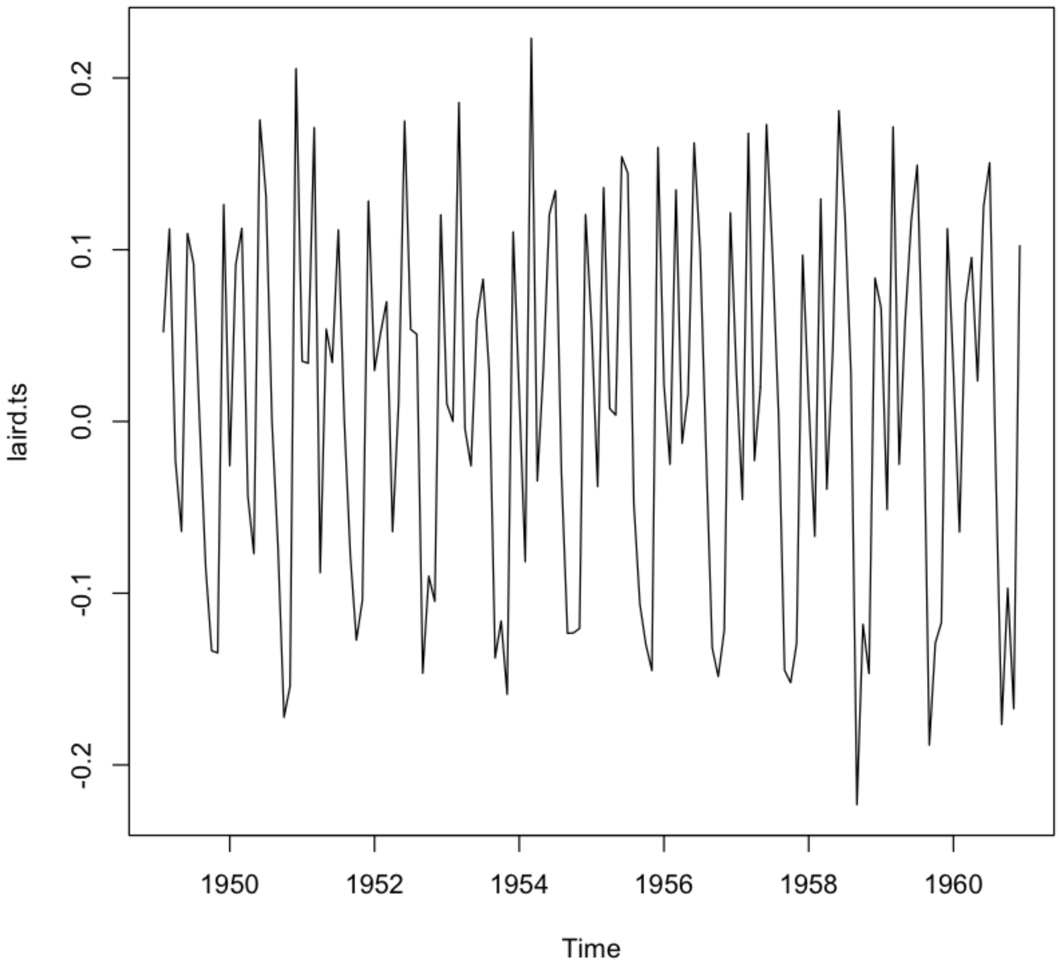
4.3 First round of differencing

- A single difference operation:

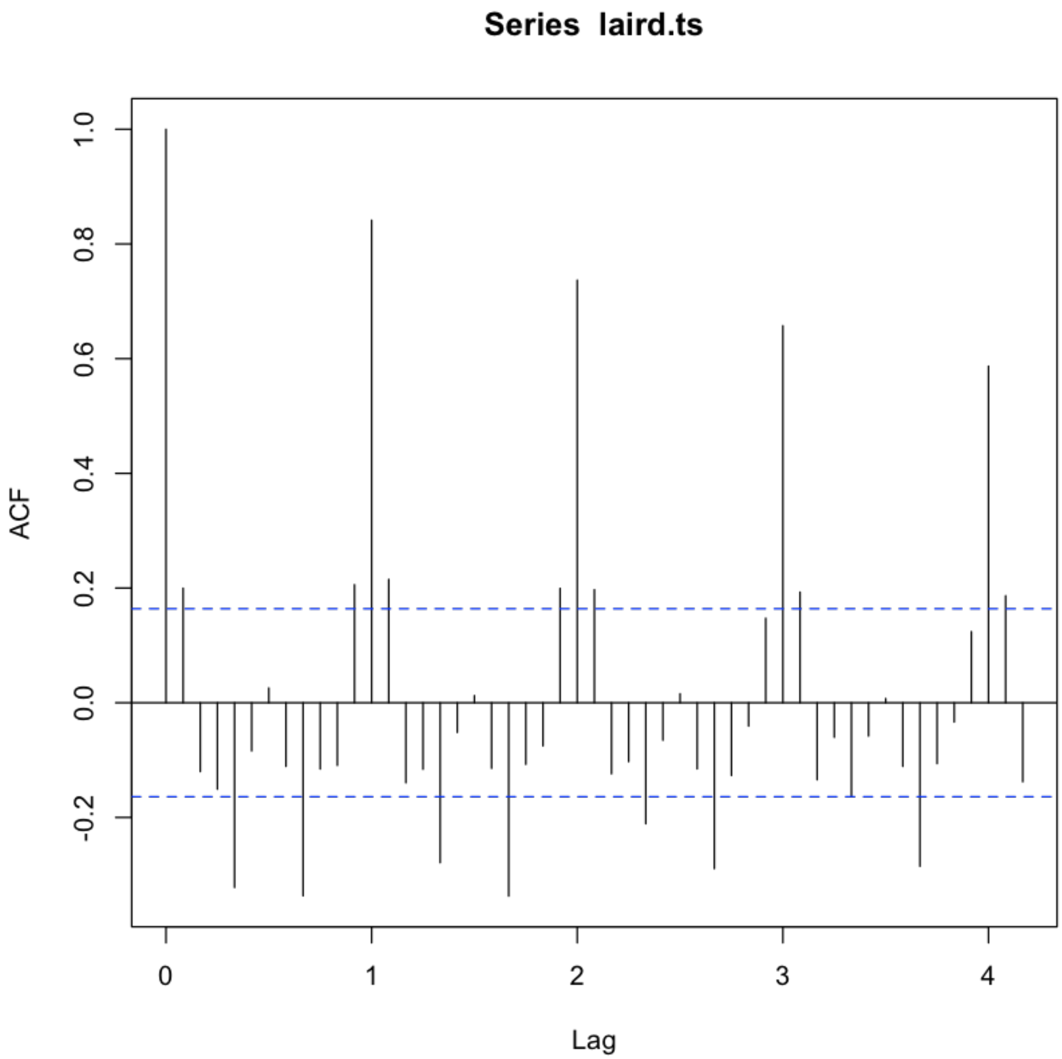
  $W_t = \Delta \ln(Y_t) = (1 - L)\ln(Y_t)$

```
In [45]: laird.ts <- diff(lair.ts, 1)

         plot(laird.ts)
```




```
In [48]: acf(laird.ts, 50)
```



4.4 Second round of differencing:

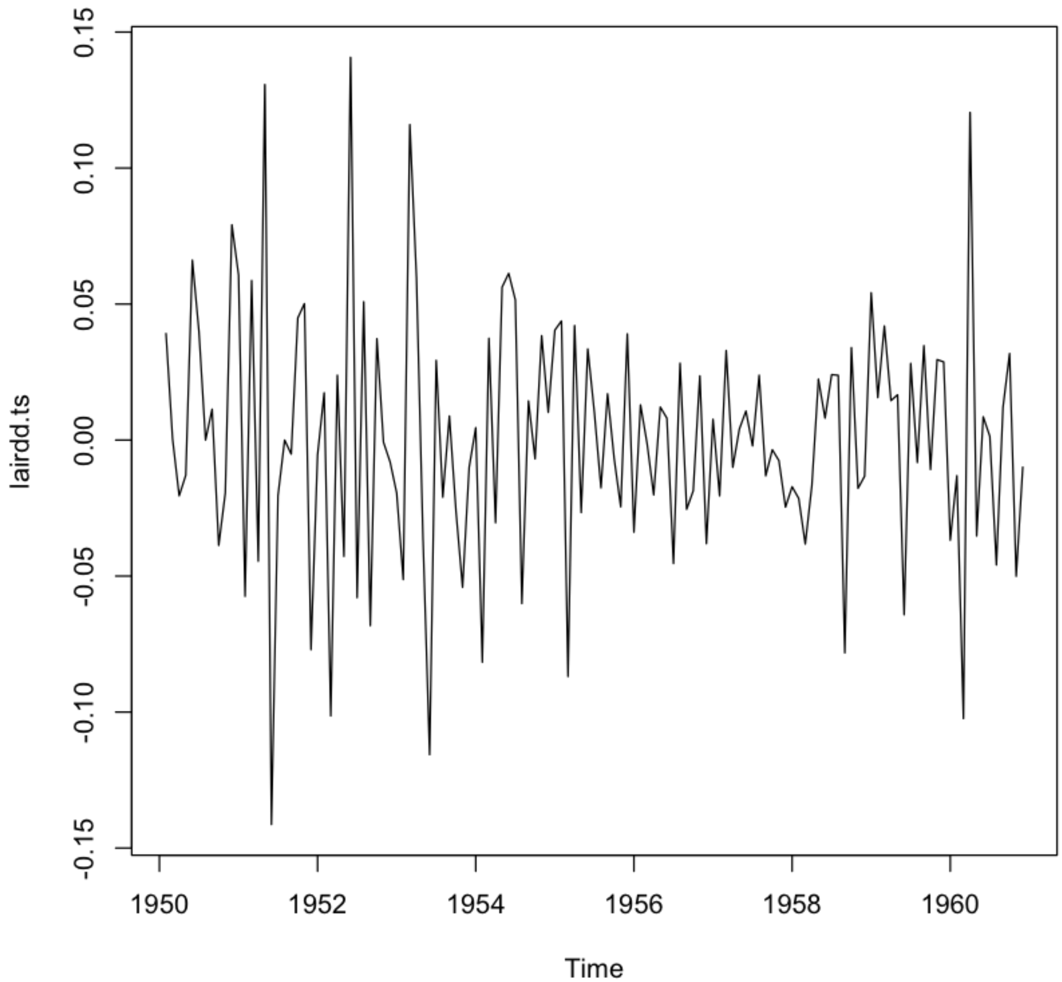
- Following the first round, we can clearly see seasonal correlation in the ACF
- I.e. there is a strong annual correlation
- Will apply seasonal difference operator:

  $\Delta_s = (1 - L^s)$

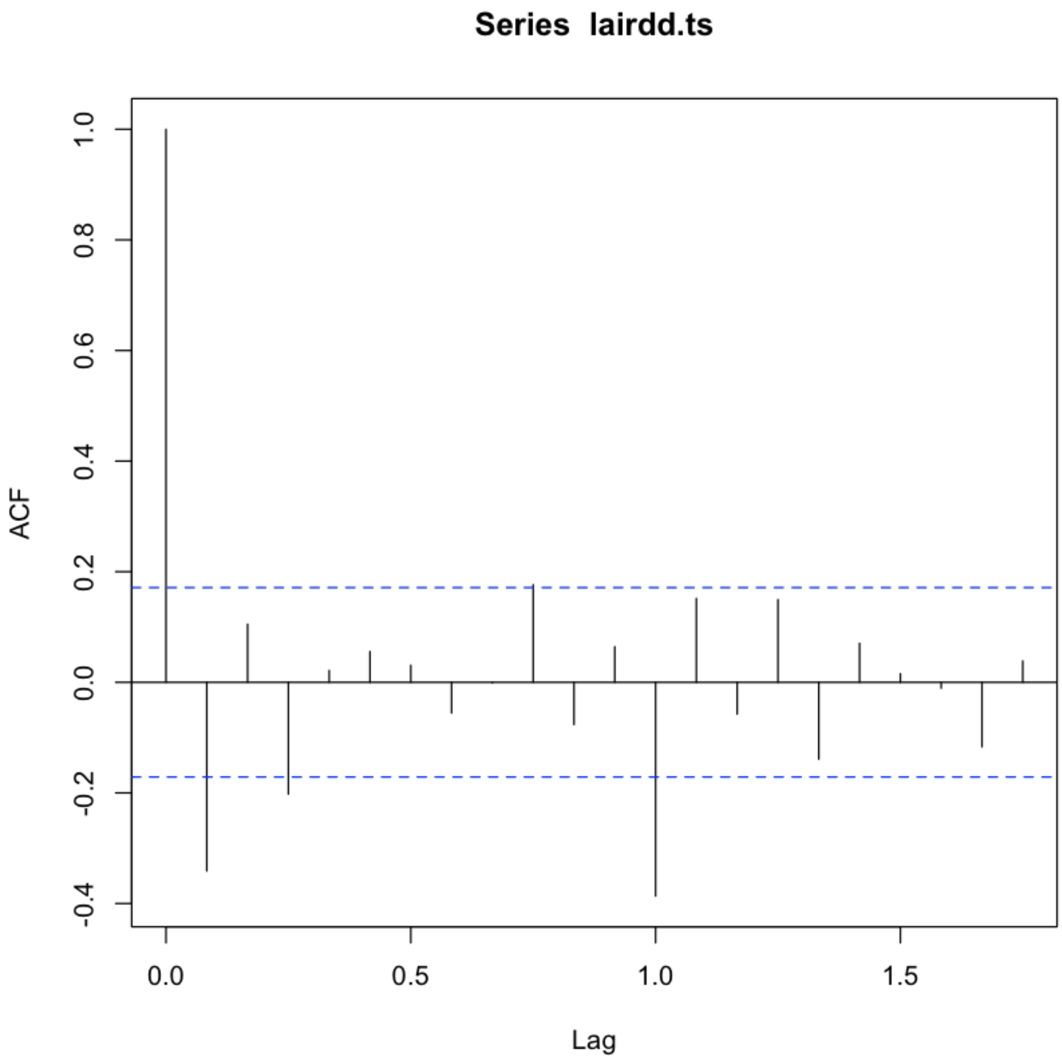
  And apply it with a period of $s = 12$.

```
In [49]: lairdd.ts <-  diff(laird.ts, 12)
```

```
In [50]: plot(lairdd.ts)
```




```
In [51]: acf(lairdd.ts)
```



Looks like the transformed process:

$$W_t = (1 - L)(1 - L^{12}) \ln Y_t$$

Is an MA(1) or an MA(2) process.
The multiplicative seasonal **ARIMA** model will therefore be:

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

```
In [59]: lair.arima.1 <- arima(lair.ts, order=c(0,1,1), seasonal=list(order=c(0,1,1), period=12

         lair.arima.1
```

```
Call:
arima(x = lair.ts, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1), period = 12))

Coefficients:
          ma1     sma1
      -0.4018  -0.5569
s.e.   0.0896   0.0731

sigma^2 estimated as 0.001348:  log likelihood = 244.7,  aic = -483.4
```

4.4.2 What model have we fit here?

We've switched $Y_t \to \ln Y_t$ and fit the ARIMA model to $\ln Y_t$.

Let's just refer to $\ln Y_t$ as Y_t here.

We've fit an MA(1) model to $\{W_t\}$, where:

$$W_t = (1 - L)(1 - L^{12})Y_t$$

Let's first write out the FULL seasonal **ARMA** model for W_t:

$$\phi(L)\Phi(L^s)(W_t - \mu) = \theta(L)\Theta(L^s)\epsilon_t$$

where:

AR characteristic polynomial: > $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p$

MA characteristic polynomial: > $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q$

AR *seasonal* characteristic polynomial: > $\Phi(L^2) = 1 - \Phi_1 L^s - \Phi_2 L^{2s} - ... - \Phi_P L^{Ps}$

MA *seasonal* characteristic polynomial > $\Theta(L^s) = 1 + \Theta_1 L^s + \Theta_2 L^{2s} + ... + \Theta_Q L^{Qs}$.

4.5 Simplification:

When you're dealing with an MA(q) process - i.e. $p = 0$ - then $\phi(L) = 1$ and $\Phi(L^2) = 1$.

Likewise, when you're dealing with an AR(p) process - i.e. when $q = 0$ - then $\theta(L) = 1$ and $\Theta(L^s) = 1$.

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

Starting with the W_t as an MA(1) process:

$$W_t = \theta(L)\Theta(L^s)\epsilon_t = \theta(L)\Theta(L^{12})\epsilon_t$$

Subbing $W_t = (1 - L)(1 - L^{12})Y_t$:

$$(1 - L)(1 - L^{12})Y_t = \theta(L)\Theta(L^{12})\epsilon_t$$

For MA(1):

$$\theta(L) = 1 + \theta_1 L$$

$$\Theta(L^{12}) = 1 + \Theta_1 L^{12}$$

Therefore the full expression for the model:

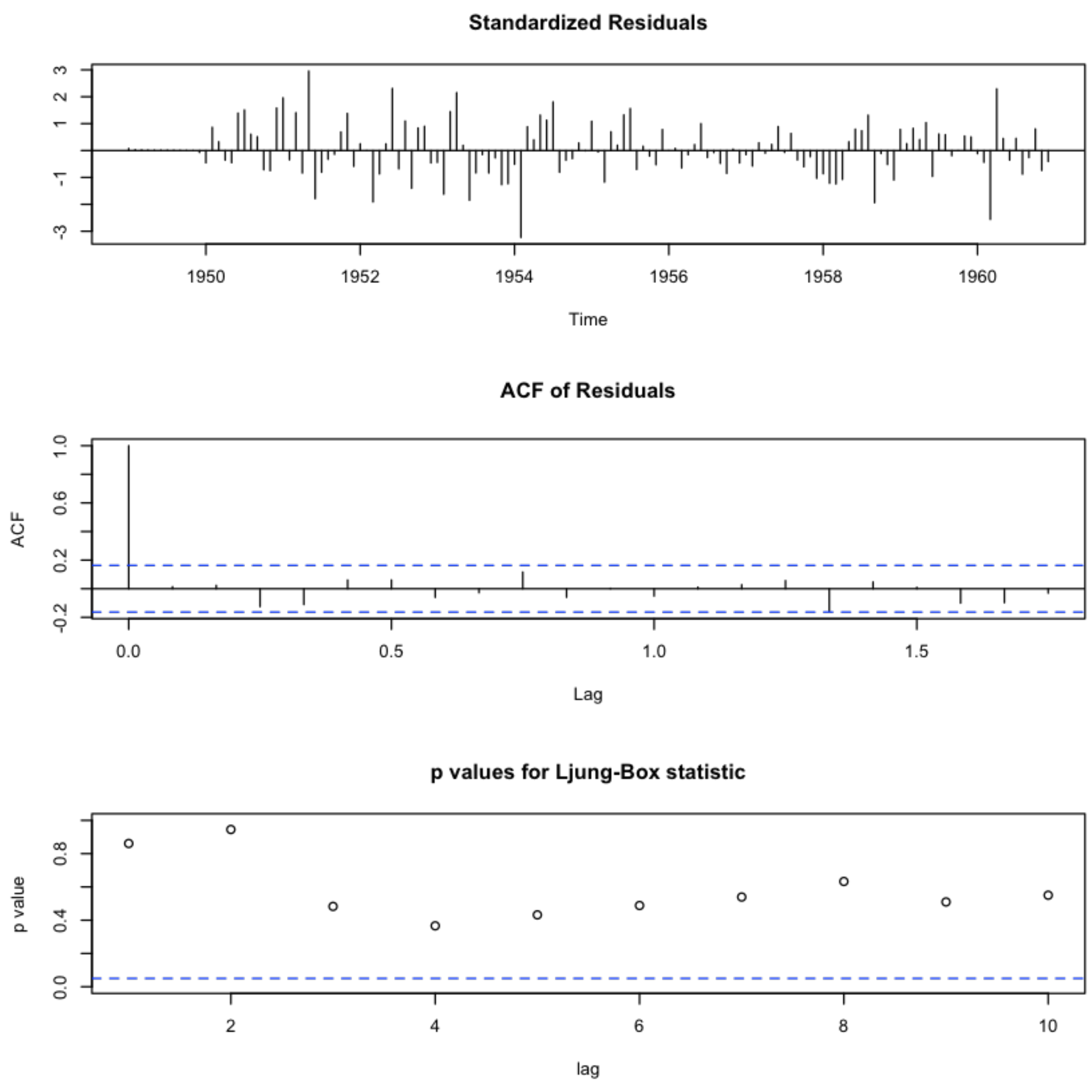
$$(1 - L)(1 - L^{12})Y_t = (1 + \theta_1 L)(1 + \Theta_1 L^{12})\epsilon_t$$

And you can multiply all of this out to produce:

$$Y_t = Y_{t-12} + Y_{t-1} - Y_{t-13} + \epsilon_t + \Theta_1 \epsilon_{t-12} + \theta_1 \epsilon_{t-1} + \theta_1 \Theta_1 \epsilon_{t-13}$$

And you can calculate the parameters from the R output.

```
In [57]: tsdiag(lair.arima.1)
```




```
In [ ]:
```