AR AND MA MODELS

Introduction to ARIMA models

Notation

- We will first discuss each model individually
  - Autoregressive (AR)
  - Moving Averages (MA)
- Then moved into the combined ARMA and ARIMA models
- Notation:
  - AR(p)
  - MA(q)
  - ARMA(p,q)
  - ARIMA(p,d,q)

AUTOREGRESSIVE MODELS

Dr. Susan Simmons

Institute for Advanced Analytics

Autoregressive (AR) Models

- Often you can forecast a series based solely on the past values of Y_t.
- We are going to focus on the most basic case – only one lag value of Y_t – called an AR(1) model:

$$Y_t = \omega + \phi Y_{t-1} + e_t$$

Autoregressive (AR) Models

- This relationship between *t* and *t-1* exists for all one time period differences across the data set.

$$Y_t = \omega + \phi Y_{t-1} + e_t$$

$$Y_{t-1} = \omega + \phi Y_{t-2} + e_{t-1}$$

$$Y_{t-2} = \omega + \phi Y_{t-3} + e_{t-2}$$

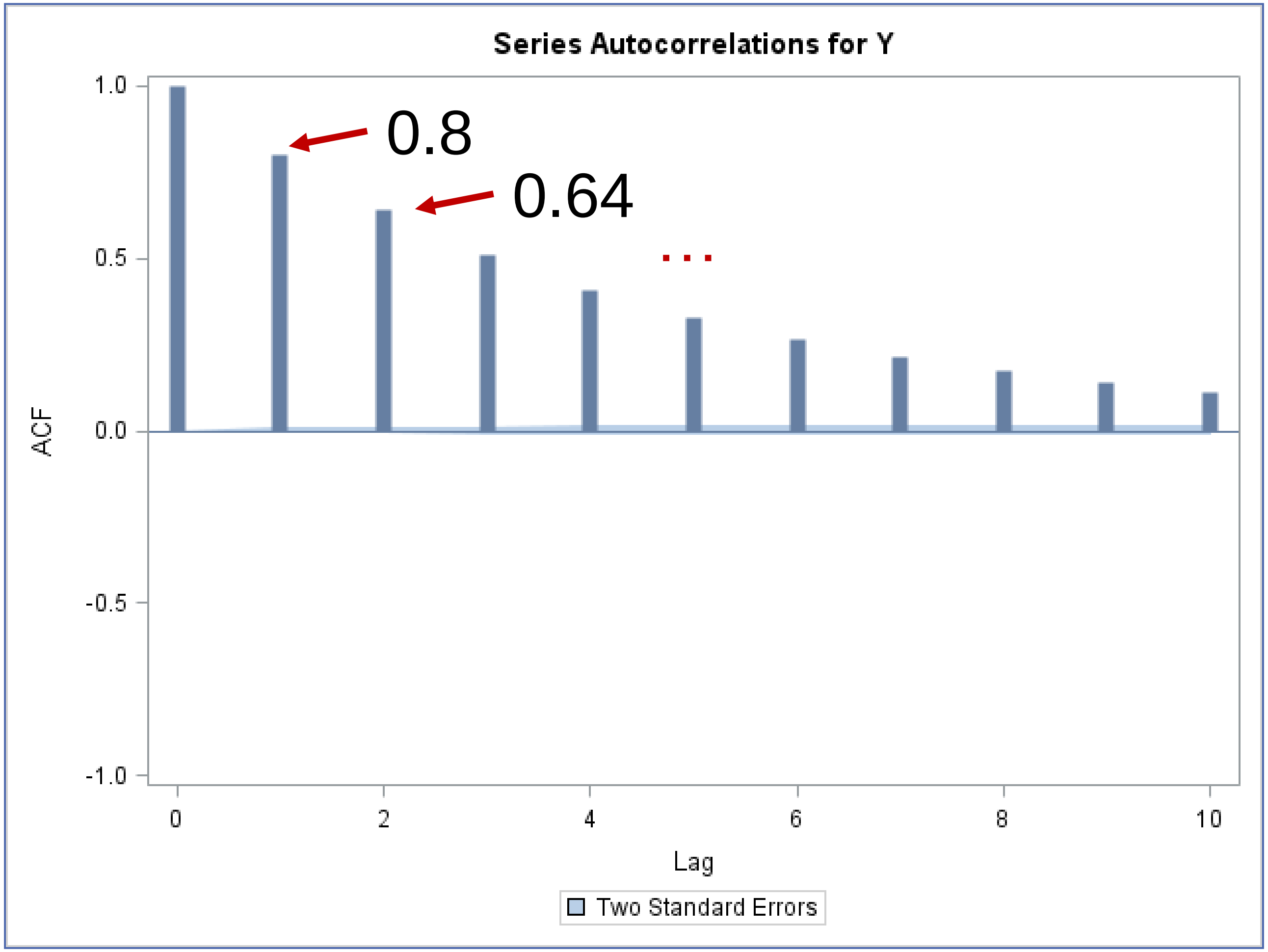
Correlation Functions for AR(1)

- The ACF decreases exponentially as the number of lags increases.
- The PACF has a significant spike at the first lag, followed by nothing after.

- Let's examine the following AR(1) model:

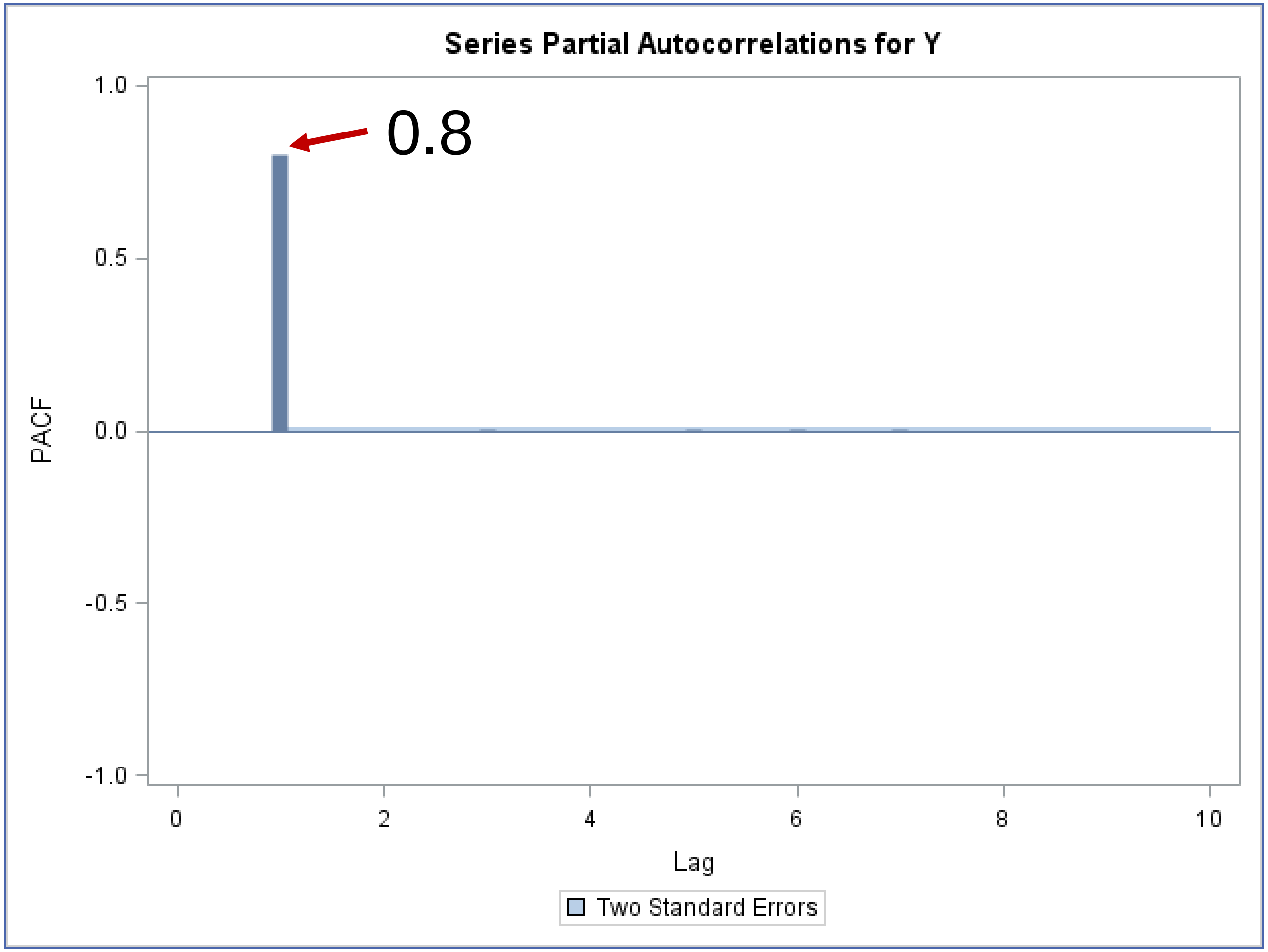
$$Y_t = 0 + 0.8Y_{t-1} + e_t$$

AR(1) – ACF



Series Autocorrelations for Y
0.8
0.64
...
1.0
0.5
0.0
-0.5
-1.0
ACF
0
2
4
6
8
10
Lag
Two Standard Errors

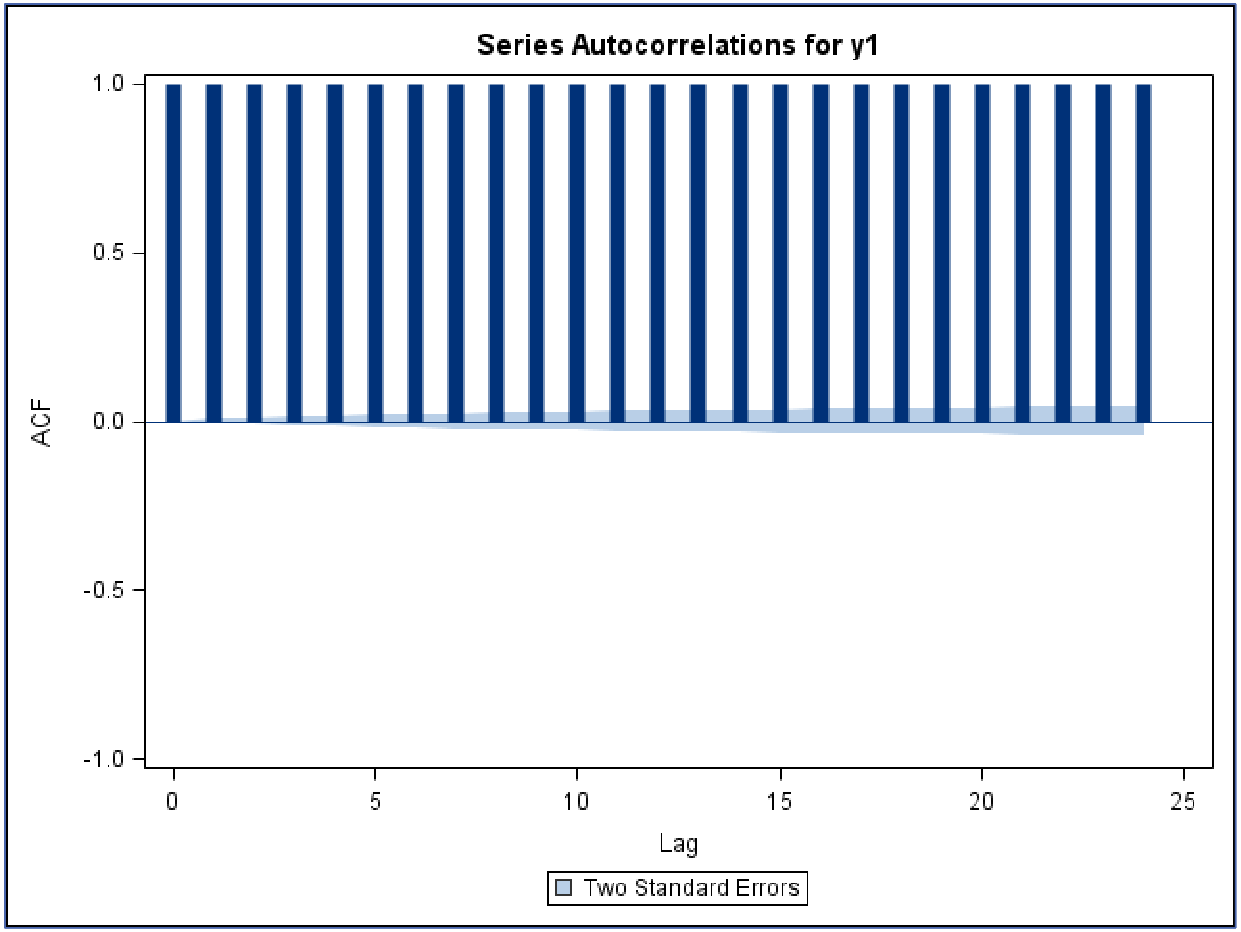
AR(1) – PACF


Series Partial Autocorrelations for Y
0.8
1.0
0.5
0.0
-0.5
-1.0
PACF
0
2
4
6
8
10
Lag
Two Standard Errors

Autoregressive (AR(1)) Models

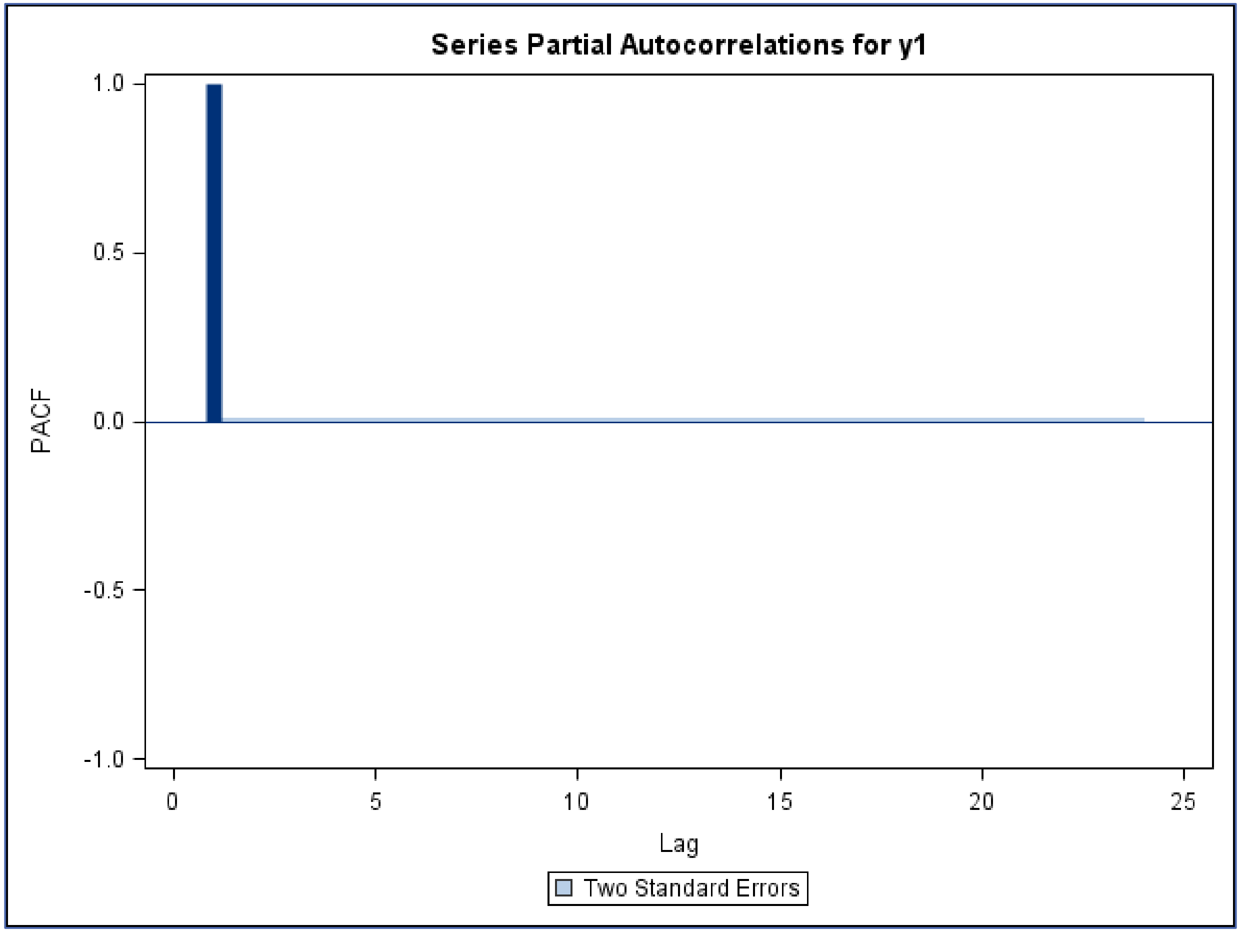
- So the effect of shocks that happened long ago has little effect on the present *IF* the value for $|\phi| < 1$.
- This goes back to our idea of stationarity – the dependence of previous observations declines over time.
- There is a pattern for AR(1) models when it comes to stationarity.
- If ϕ =1, then Random Walk and NOT Autoregressive model
- If ϕ >1, then today depends on tomorrow (doesn't really make sense)

RW – Autocorrelation Function



Notice that RW affect the correlation plots

RW– Partial Autocorrelation Function



Only dependent on previous observation. Perfect correlation

AR(2) Model

- A time series that is a linear function of *2* past values plus error is called an autoregressive process of order 2 – AR(2).

$$Y_t = \omega + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$$

AR(2) Model

- There is a pattern in PACF plots for AR(2) models when it comes to stationarity (2 spikes in PACF).
- The effect of shocks that happened long ago has little effect on the present *IF* the value for $|\phi_1 + \phi_2| < 1$.

AR(p) Model

- A time series that is a linear function of *p* past values plus error is called an autoregressive process of order p – AR(p).

$$Y_t = \omega + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

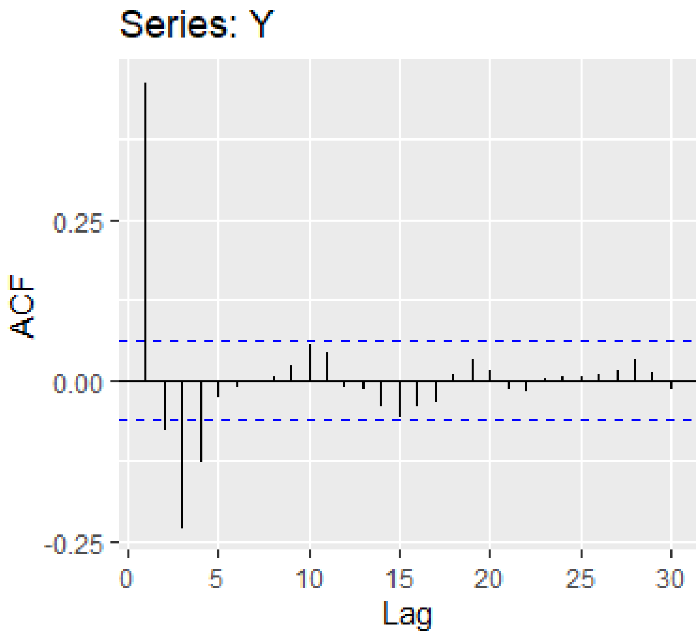
- More complicated restrictions on ϕ_i's (software will warn you when this becomes an issue)

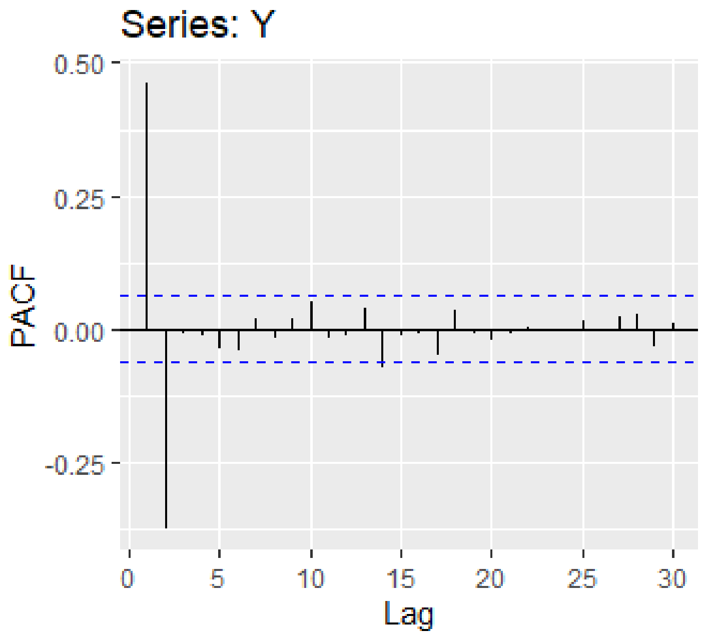
Correlation Functions for AR(p)

- The PACF has significant spikes at the lags up to *p* lags, followed by nothing after.

AR(2)

```
ggAcf(Y)
ggPacf(Y)
```





Autoregressive Models – R

```
Y.ts <- ts(Y)
Y.ARIMA <- Arima(Y.ts, order=c(2,0,0))

ggAcf(Y.ARIMA$residuals)
ggPacf(Y.ARIMA$residuals)
```

ARIMA(2,0,0) with non-zero mean

Coefficients:

	ar1	ar2	mean
	0.6406	-0.3760	-0.1371
s.e.	0.0294	0.0294	0.4187

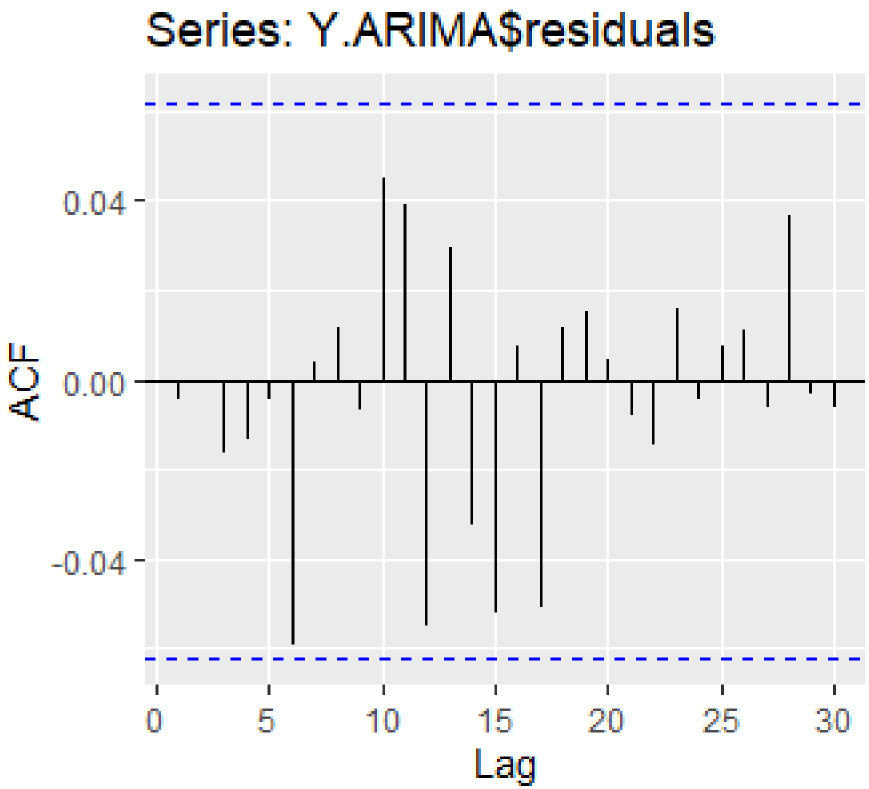
sigma^2 estimated as 95.07: log likelihood=-3695.01

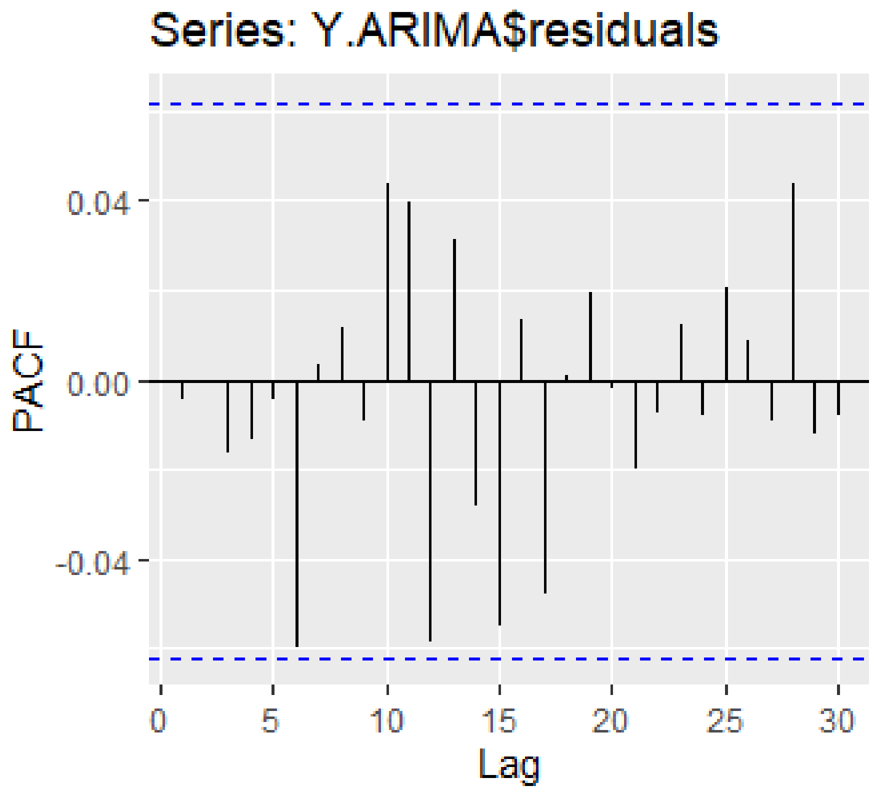
AIC=7398.03 AICc=7398.07 BIC=7417.66

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE
Training	-0.003995096	9.735715	7.710788	48.02355	313.2086

	MASE	ACF1
Training set	0.7905136	-0.004541939


Series: Y.ARIMA$residuals
ACF
0.04
0.00
-0.04
0
5
10
15
20
25
30
Lag


Series: Y.ARIMA$residuals
PACF
0.04
0.00
-0.04
0
5
10
15
20
25
30
Lag

MOVING AVERAGE MODELS

Dr. Susan Simmons

Institute for Advanced Analytics

MOVING AVERAGE MODELS

Moving Average (MA) Models

- You can also forecast a series based solely on the past *error* values.
- We are going to focus on the most basic case – only one error lag value of e_t, called an MA(1) model:

$$Y_t = \omega + e_t - \theta e_{t-1}$$

MA(1) Model

- This is true for all observations (each observation is dependent on the error from the previous observation).
- Therefore, for an MA(1) model, individual "shocks" only last for a short time.
- In the MA model, we do not have the restrictions that we did on the AR models (but do want them to be invertible).

$$Y_t = \omega + e_t - \theta e_{t-1}$$

$$Y_{t-1} = \omega + e_{t-1} - \theta e_{t-2}$$

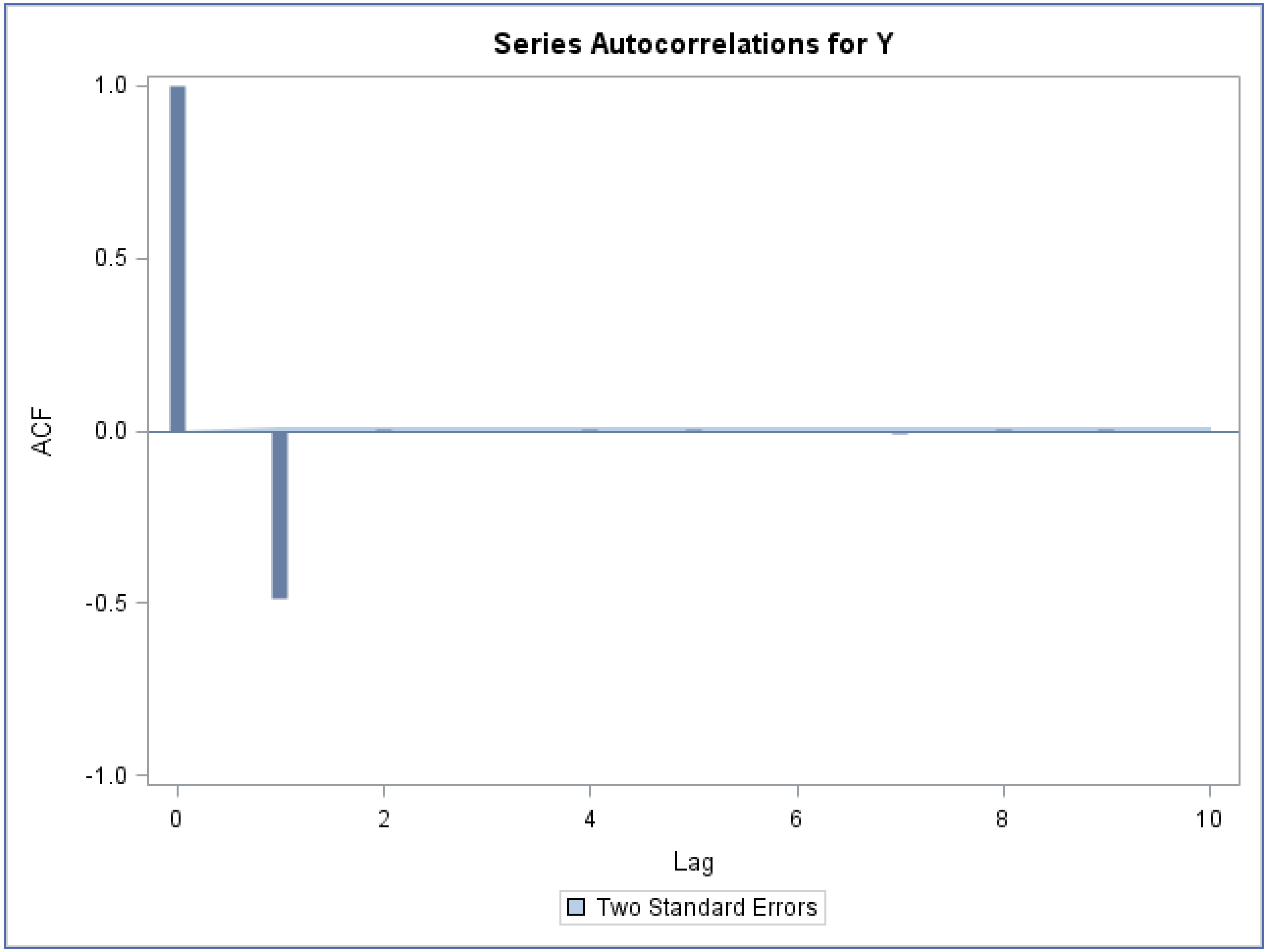
Correlation Functions for MA(1)

- The ACF has a significant spike at the first lag, followed by nothing after.
- The PACF decreases exponentially as the number of lags increases.

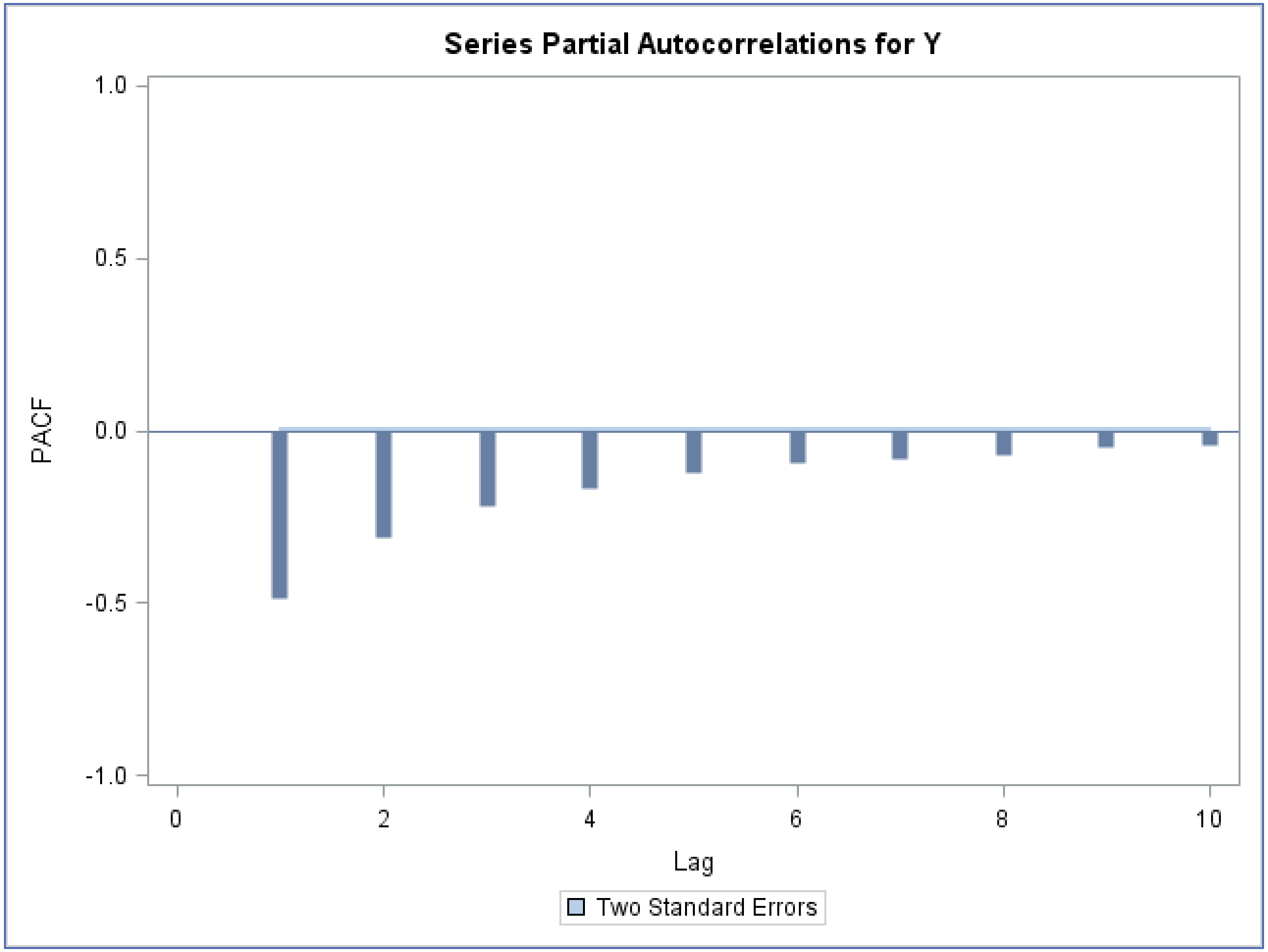
- Let's examine the following MA(1) model:

$$Y_t = 0 + e_t - 0.8e_{t-1}$$

MA(1) – ACF


Series Autocorrelations for Y
1.0
0.5
0.0
-0.5
-1.0
ACF
0
2
4
6
8
10
Lag
Two Standard Errors

MA(1) – PACF


Series Partial Autocorrelations for Y
1.0
0.5
0.0
-0.5
-1.0
PACF
0
2
4
6
8
10
Lag
Two Standard Errors

MA(q) Model

- A time series that is a linear function of *q* past errors is called a moving average process of order q – called an MA(q).

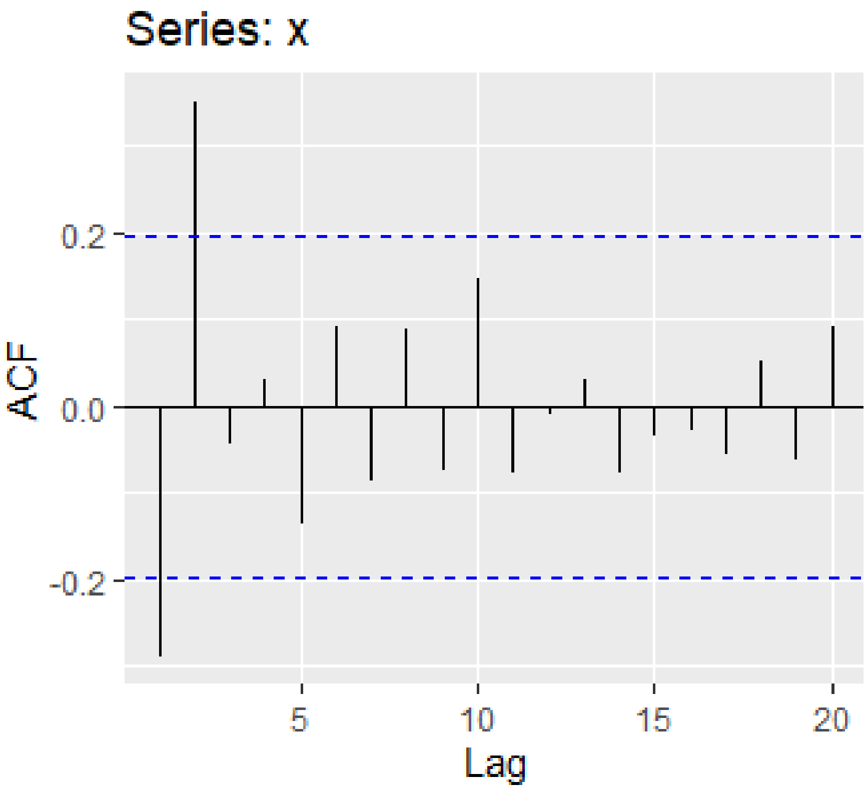
$$Y_t = \omega + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

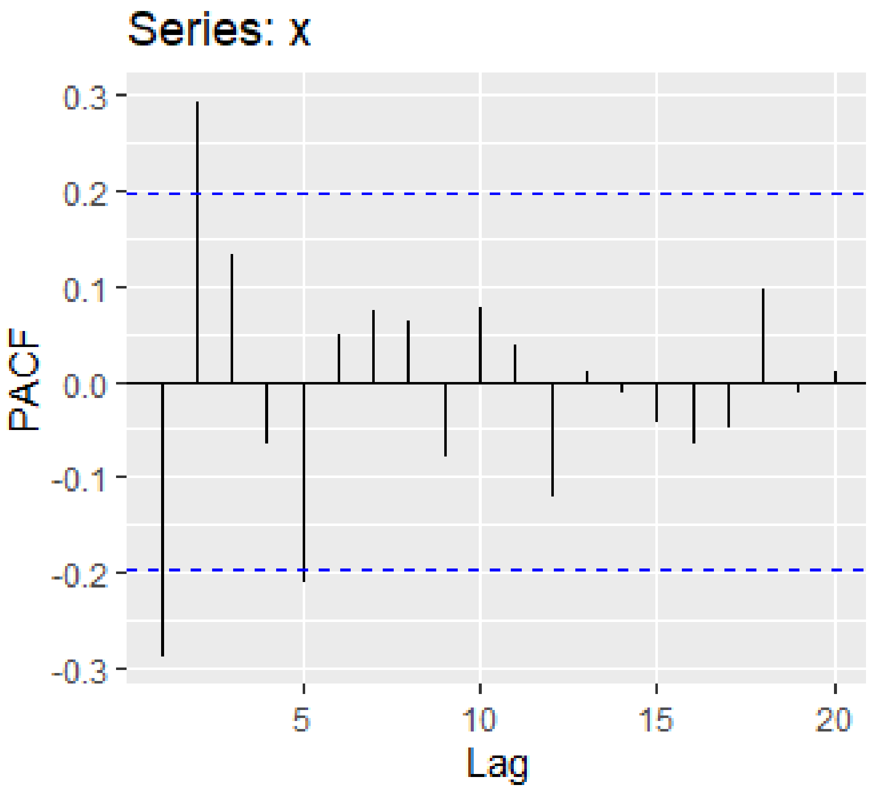
Correlation Functions for MA(q)

- The ACF has significant spikes at lags up to lag q, followed by nothing after.

MA(2)

```
ggAcf(x)
ggPacf(x)
```



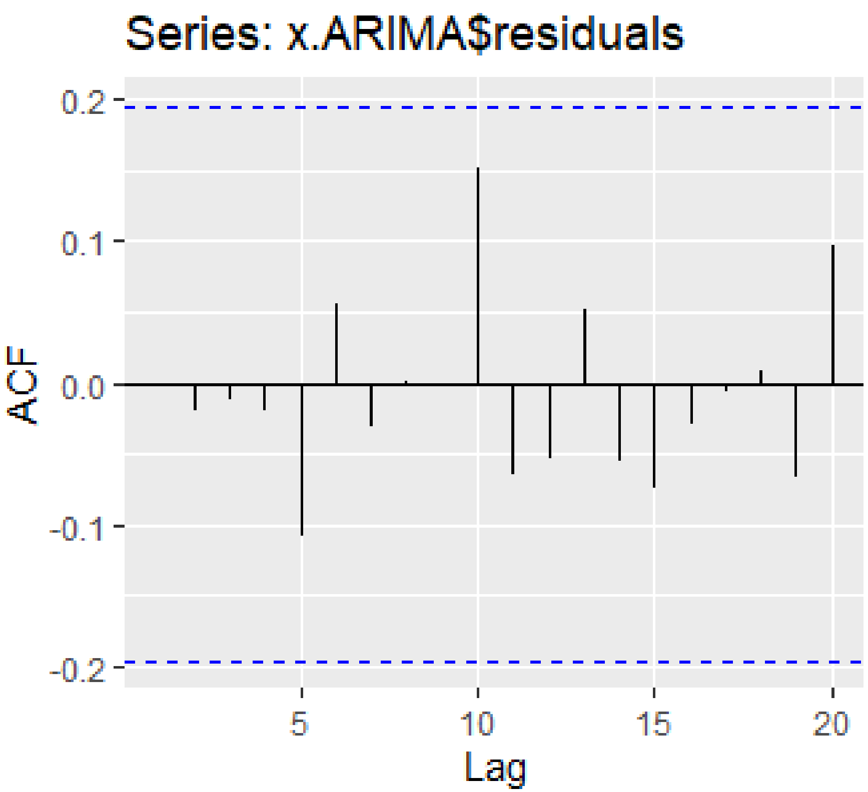


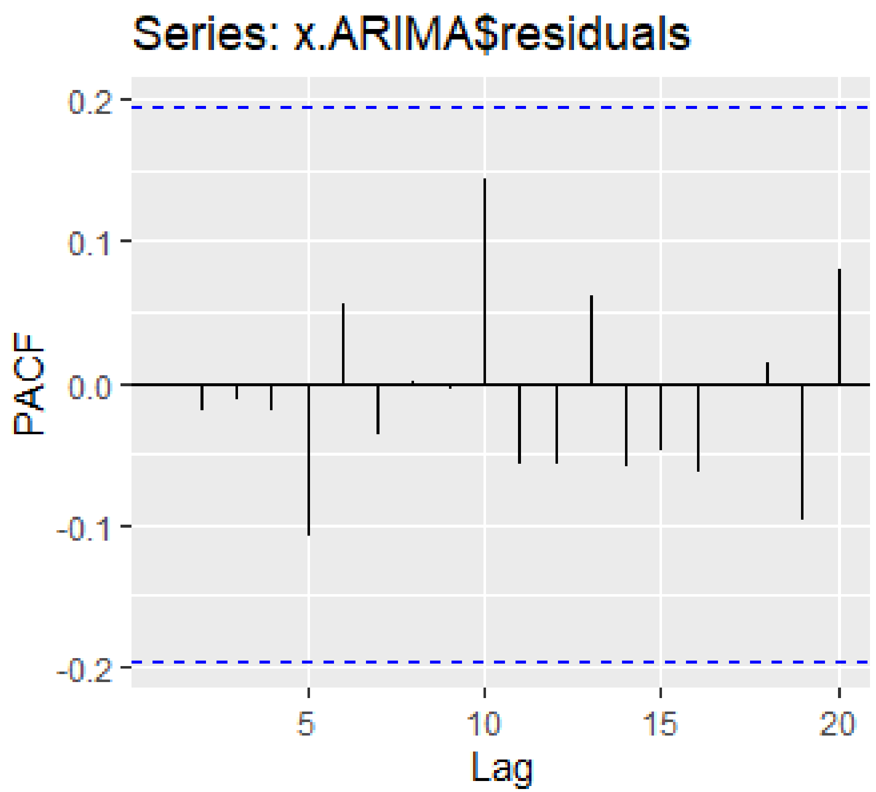
MA(2)

```
x.ts <- ts(x)
x.ARIMA <- Arima(x.ts, order=c(0,0,2))
summary(x.ARIMA)
ggAcf(x.ARIMA$residuals)
ggPacf(x.ARIMA$residuals)
```

Output

```
coefficients:
            ma1       ma2    mean
        -0.2460  0.4772  0.0250
s.e.     0.0857  0.0923   0.0567

sigma^2 estimated as 0.2207:  log likelihood=-65.1
AIC=138.2   AICc=138.63   BIC=148.62

Training set error measures:
                          ME              RMSE       MAE       MPE
Training set 0.0008828966 0.4627151 0.3808289 74.99115
MAPE         MASE          ACF1
114.4434  0.5453401  -0.002299708
```


Series: x.ARIMA$residuals
ACF
0.2
0.1
0.0
-0.1
-0.2
5
10
15
20
Lag


Series: x.ARIMA$residuals
PACF
0.2
0.1
0.0
-0.1
-0.2
5
10
15
20
Lag

Some notes about AR and MA models

- Any AR(p) model can be rewritten as an MA(∞).
- If the MA(q) model is invertible, then this MA(q) model can be rewritten as an AR(∞).
- Software should warn you if model is not invertible, if there is no convergence or any other issues….pay attention to the log and any warnings that you encounter when fitting these models.
- Depending on how software parameterizes equations, parameters can have different signs.

WHITE NOISE

Statistical Forecasting

Time Series = Signal + Noise

If we are successful in removing all signals, we are left with independent errors.

White Noise

White Noise

- A white noise time series have errors that follow a Normal distribution (or bell-shaped) with mean zero and positive, *constant* variance in which all observations are independent of each other.
- Autocorrelation and partial autocorrelation functions have a value close to zero at every time point (except for lag of 0).

White Noise

- The goal of modeling time series is to be left with white noise residuals in the time series.
- If the residuals still have a “significant” dependence structure, then more modeling can typically be done.
- How do we know when we are left with white noise at the end of the model? (you already know how to check for normality and constant variance, so we will focus on the dependence structure).

Ljung-Box χ^2 Test for White Noise

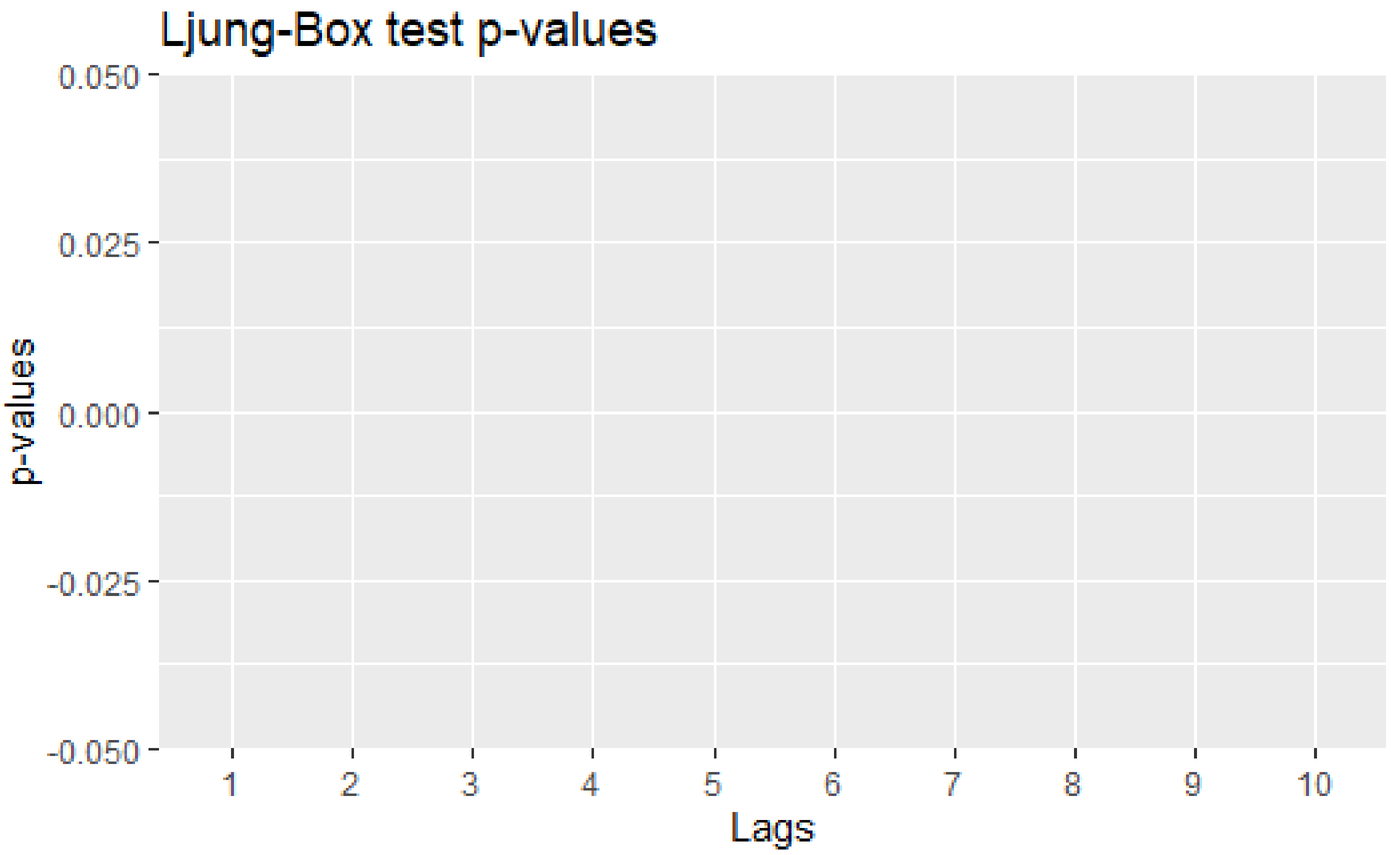
- The Ljung-Box test may be applied to the original data or to the residuals after fitting a model.
- The null hypothesis is that the series has NO autocorrelation, and the alternative hypothesis is that one or more autocorrelations up to lag *m* are not zero.

$$\chi_m^2 = n(n+2)\sum_{k=1}^{m}\frac{\hat{\rho}_k^2}{n-k}$$

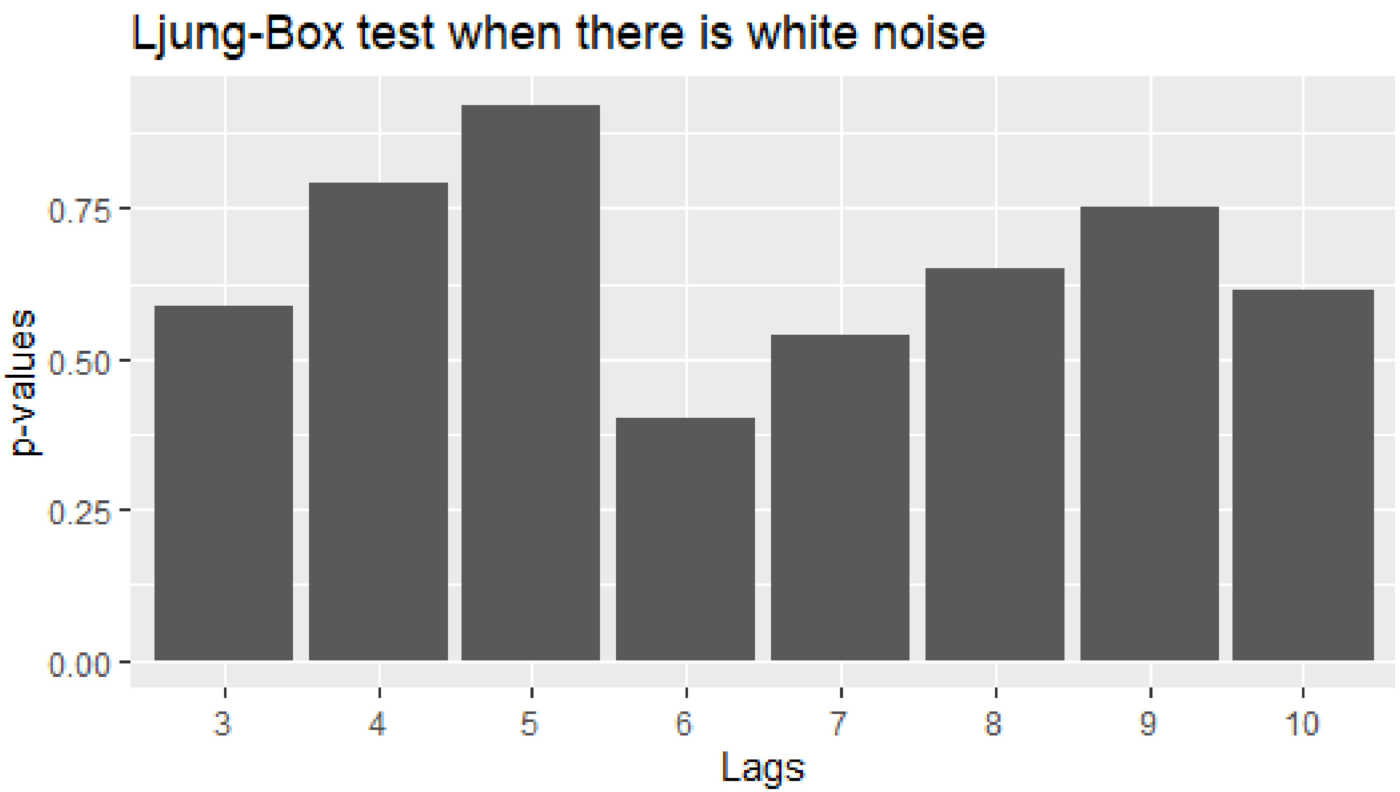
Testing for White Noise – R

```
White.LB <- rep(NA, 10)
for(i in 1:10){
  White.LB[i] <- Box.test(Y, lag=i, type="Ljung-Box", fitdf =
0)$p.value
}

white.dat=data.frame(cbind(White.LB,index1))
colnames(white.dat)=c("pvalues","Lag")

ggplot(white.dat,aes(x=factor(Lag),y=pvalues))+geom_col()
+labs(title="Ljung-Box test p-values",x="Lags",y="p-
values")
```

Small p-values!!



After fitting appropriate model



Code

```
Y.ARIMA=Arima(Y,order=c(2,0,0))
White.LB <- rep(NA, 10)
for(i in 3:10){
  White.LB[i] <- Box.test(Y.ARIMA$residuals, lag=i,
type="Ljung-Box", fitdf = 2)$p.value
}
white.dat=data.frame(cbind(White.LB[3:10],index1[3:10]))
colnames(white.dat)=c("pvalues","Lag")

ggplot(white.dat,aes(x=factor(Lag),y=pvalues))+geom_col()+l
abs(title="Ljung-Box test when there is white
noise",x="Lags",y="p-values")
```