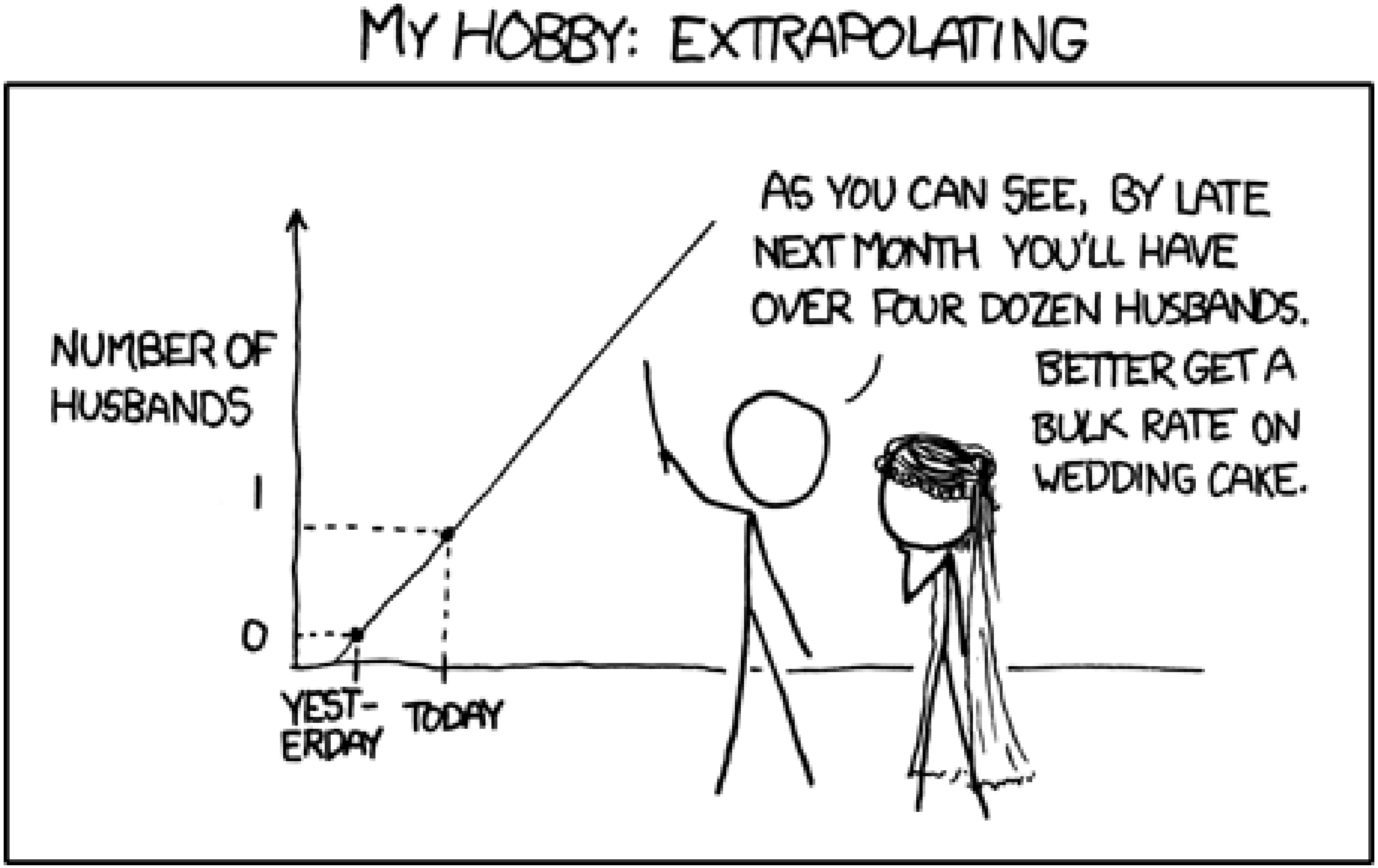
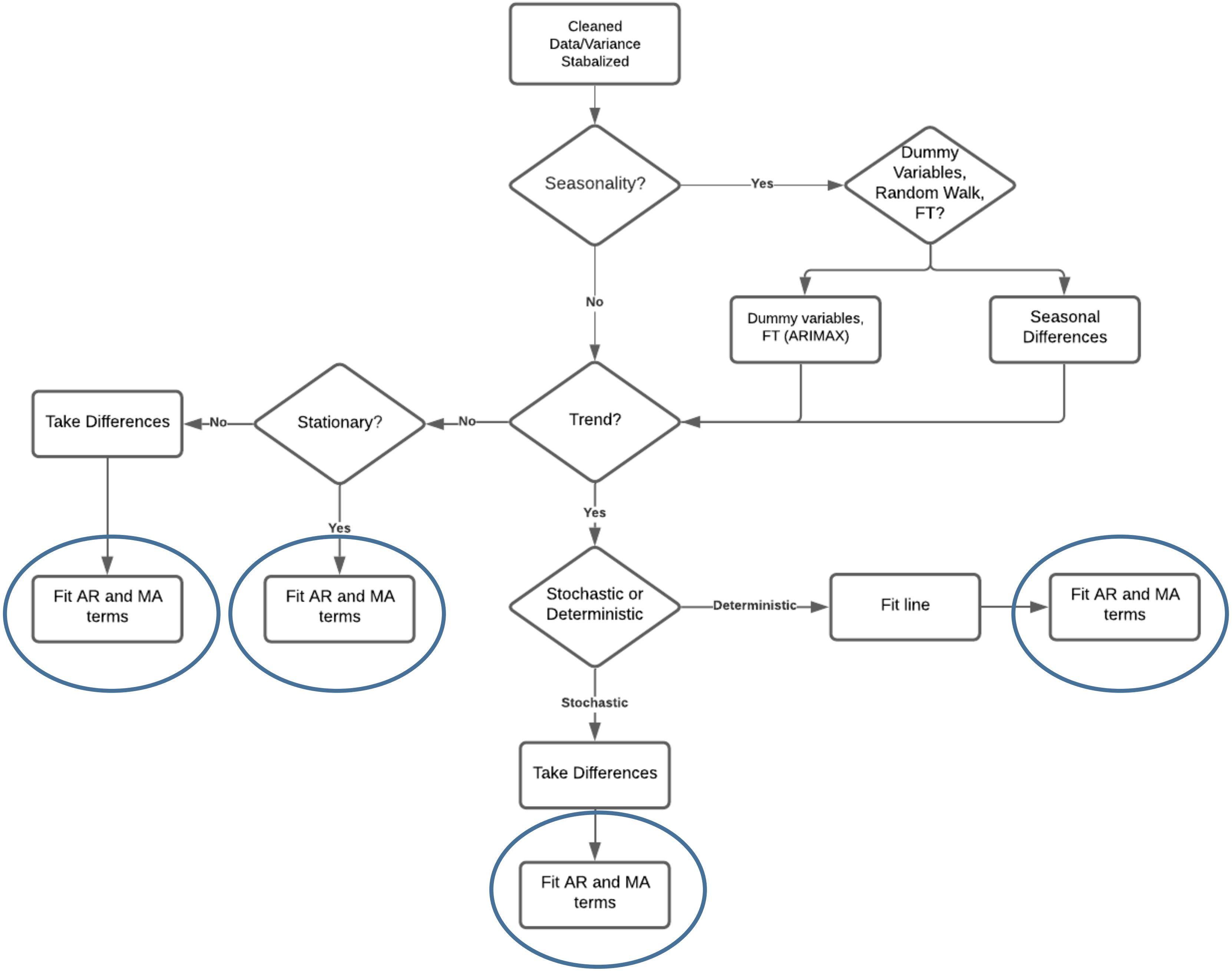


Source: xkcd: Extrapolating

AR AND MA MODELS AND WHITE NOISE

Introduction to ARIMA models



Cleaned Data/Variance Stabalized
Seasonality?
Yes
Dummy Variables, Random Walk, FT?
No
Dummy variables, FT (ARIMAX)
Seasonal Differences
Trend?
No
Stationary?
No
Take Differences
Yes
Fit AR and MA terms
Yes
Fit AR and MA terms
Stochastic or Deterministic
Deterministic
Fit line
Fit AR and MA terms
Stochastic
Take Differences
Fit AR and MA terms

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 (on a stationary time series):

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

Autoregressive (AR) Models

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

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

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

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

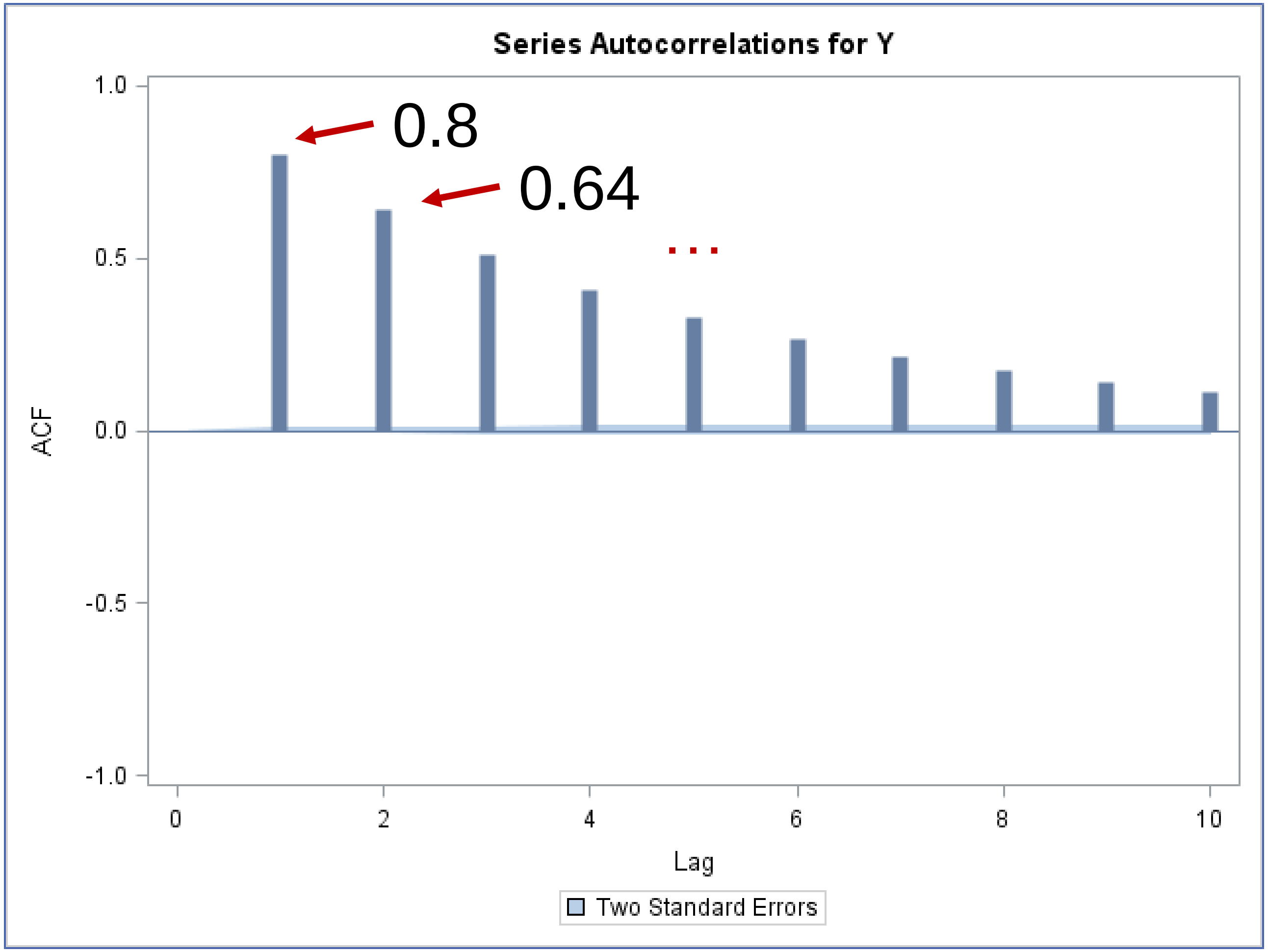
Correlation Functions for AR(1)

- 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:

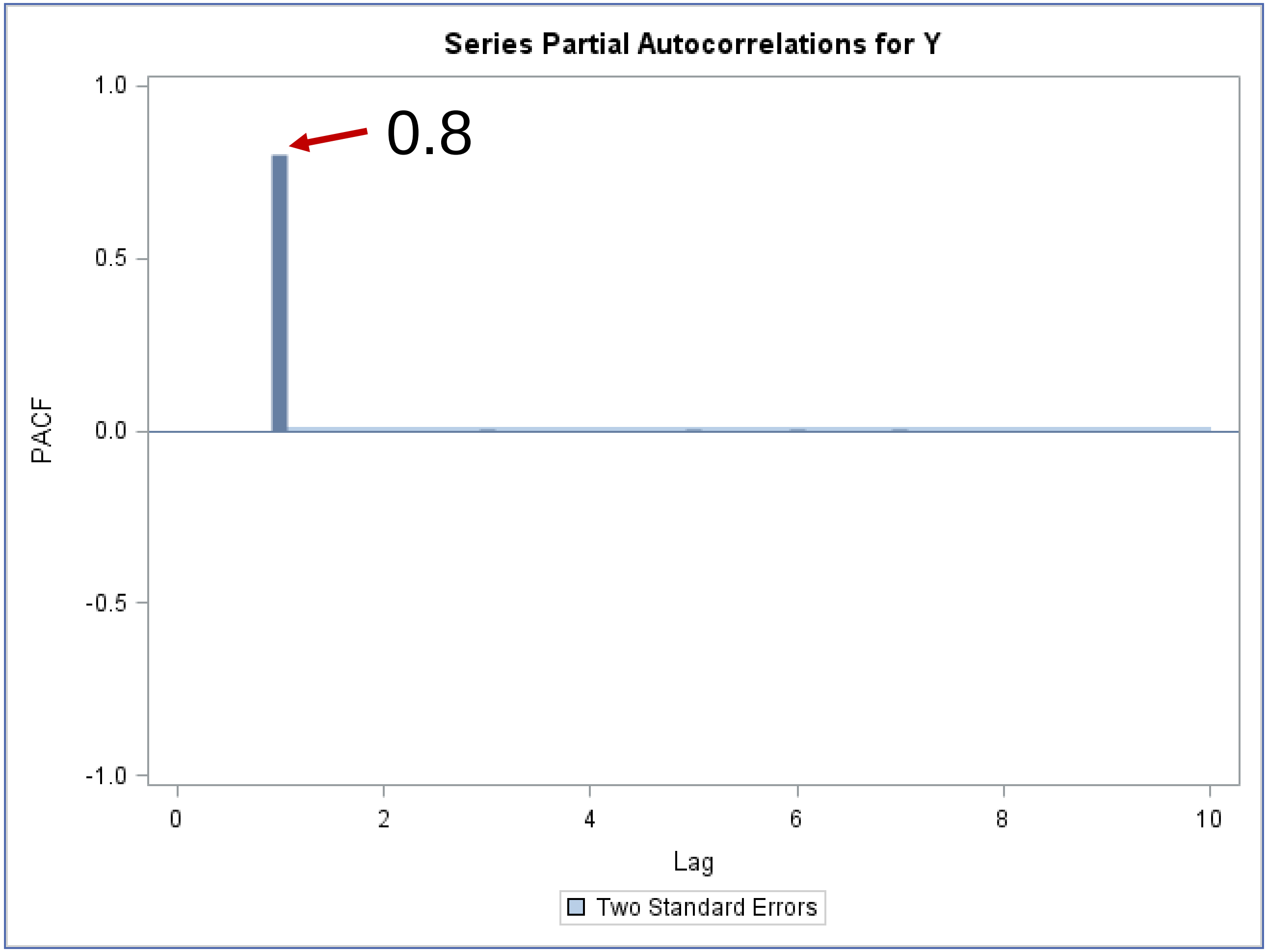
$$Y_t = 0 + 0.8Y_{t-1} + e_t$$

AR(1) – ACF



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

AR(1) – PACF

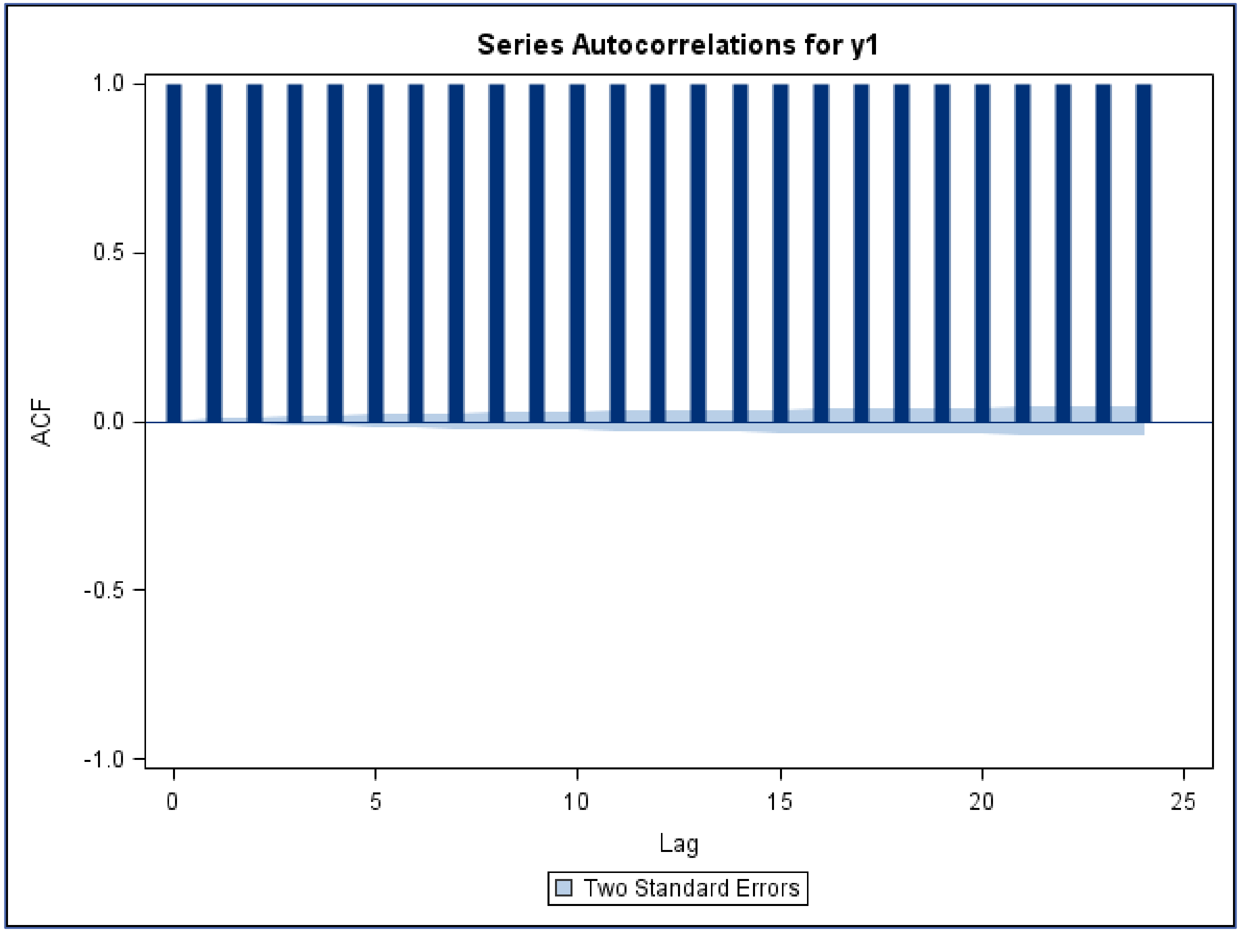


Series Partial Autocorrelations for Y
0.8
1.0
0.5
0.0
-0.5
-1.0
PACF
0
2
4
6
8
10
Lag
Two Standard Errors

Autoregressive (AR(1)) Models

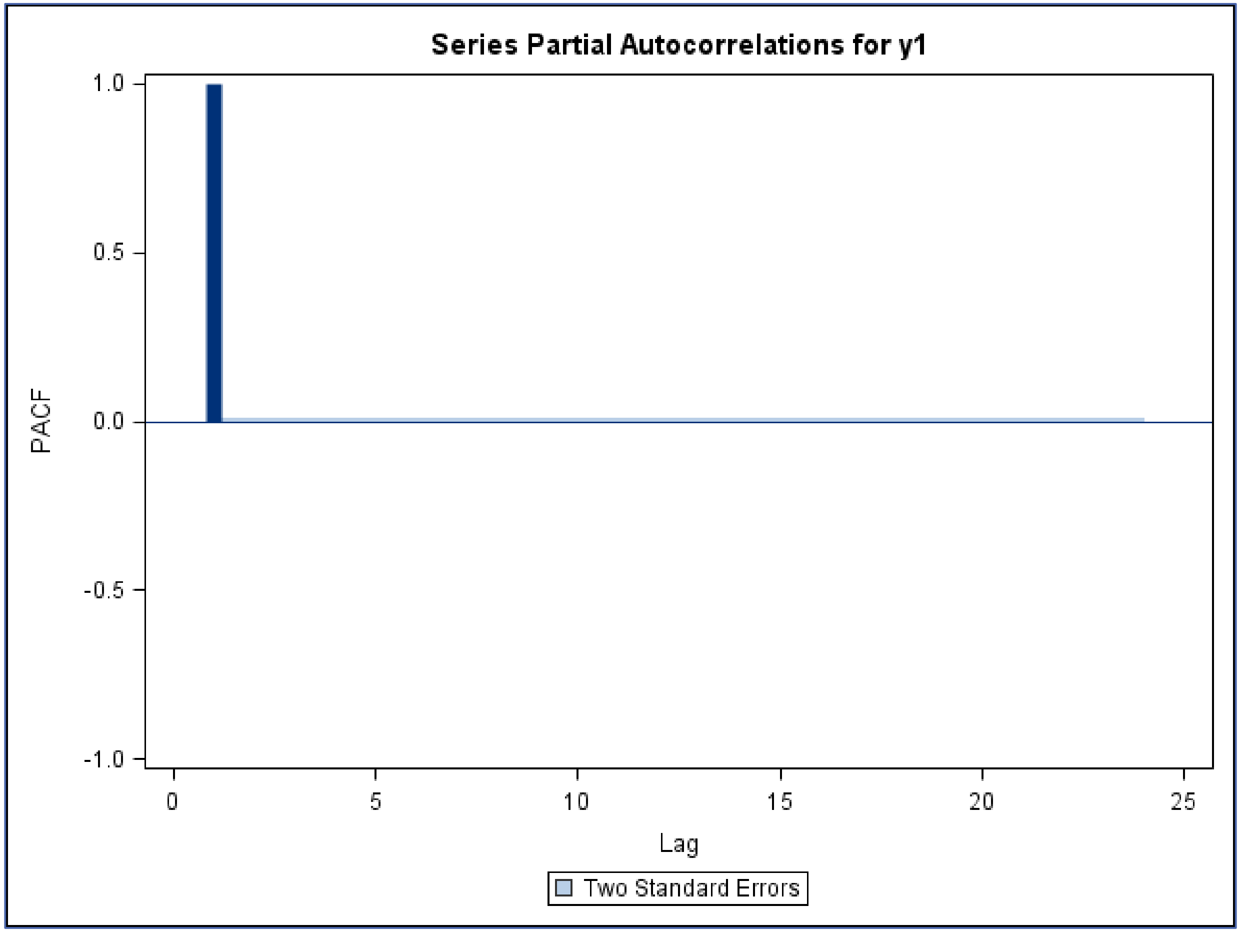
- We see the diminishing impact of observations from earlier in time (the ACF plot keeps diminishing).
- This identifies a pattern for AR(1) model….ACF is exponentially decreasing and PACF has one spike at lag 1.
- If ϕ =1, then Random Walk and NOT Autoregressive model (recall…$Y_t = Y_{t-1} + \varepsilon_t$…..or ϕ =1 in RW)…can see this in ACF plot!
- 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) and generally see the ACF exponentially decreasing.
- For these models to be stationary, we need $|\phi_1 + \phi_2| < 1$.

AR(p) Model

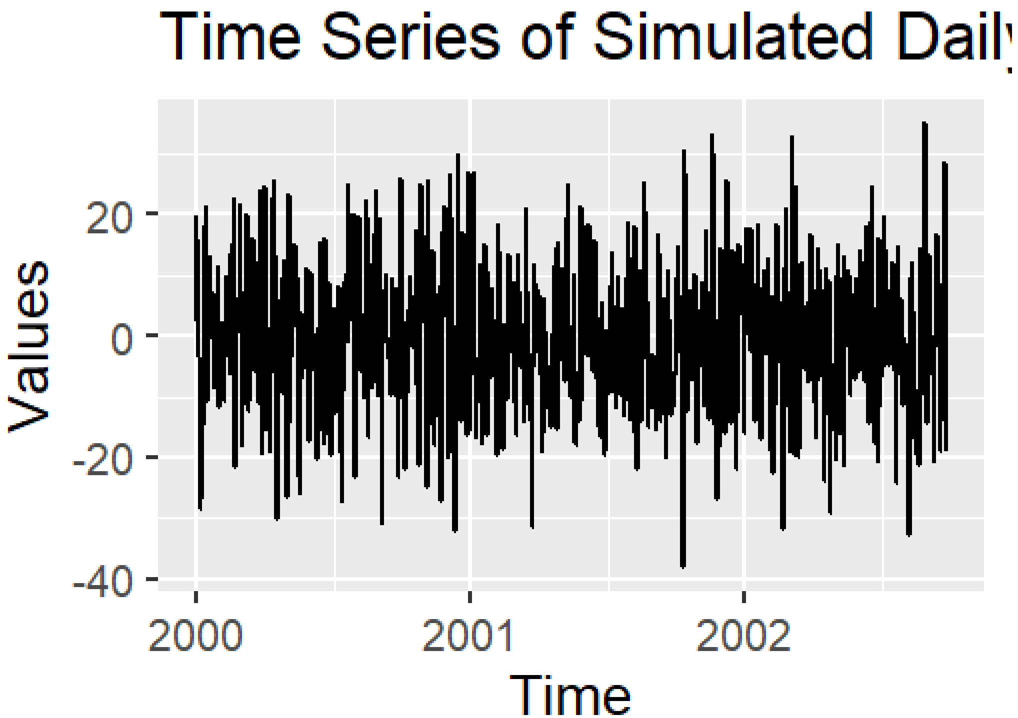
- A time series that is a linear function of *p* past values plus error is called an autoregressive process of order p – AR(p).

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

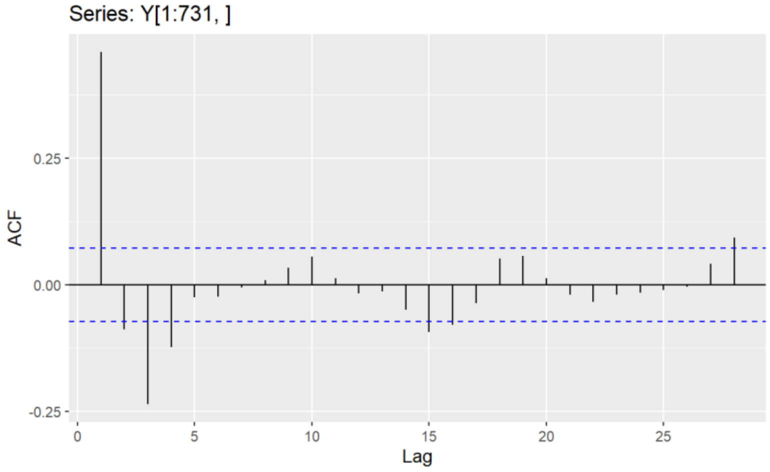
- No consistent pattern (potentially see spikes up to lag p in the PACF plot).
- More complicated restrictions on ϕ_i's (software will warn you when this becomes an issue)

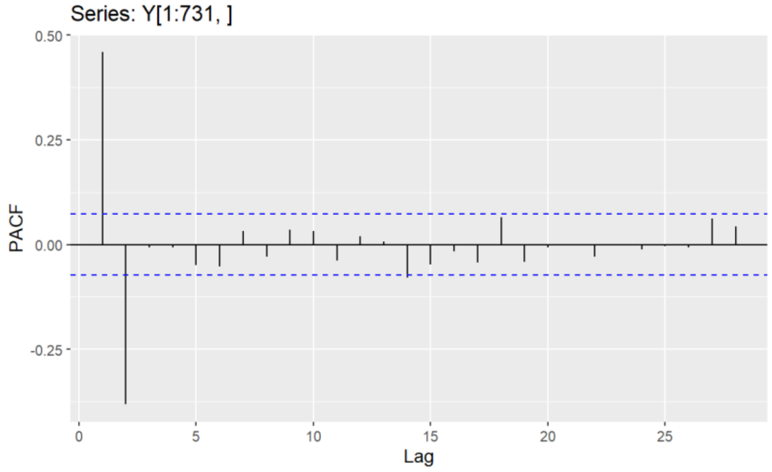
Simulated data set

- The data set Y was simulated to be an AR(2) model
- Made it into a daily time series (it is stationary):



```
ggAcf(Y[1:731,])
ggPacf(Y[1:731,])
```





Exponentially decreasing in the ACF plot

Two big spikes in the PACF plot

Autoregressive Models – R

```
Y.ARIMA <- Y_train %>%
model(ARIMA(Y~pdq(2,0,0)+PDQ(0,0,0)))

report(Y.ARIMA)
```

```
## Series: Y
## Model: ARIMA(2,0,0)
##
## Coefficients:
##          ar1      ar2
##       0.6399  -0.3838
## s.e.  0.0342   0.0342
##
## sigma^2 estimated as 93.75:  log likelihood=-2696.14
## AIC=5398.28   AICc=5398.32   BIC=5412.07
```

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 (however, they need to be invertible).

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

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

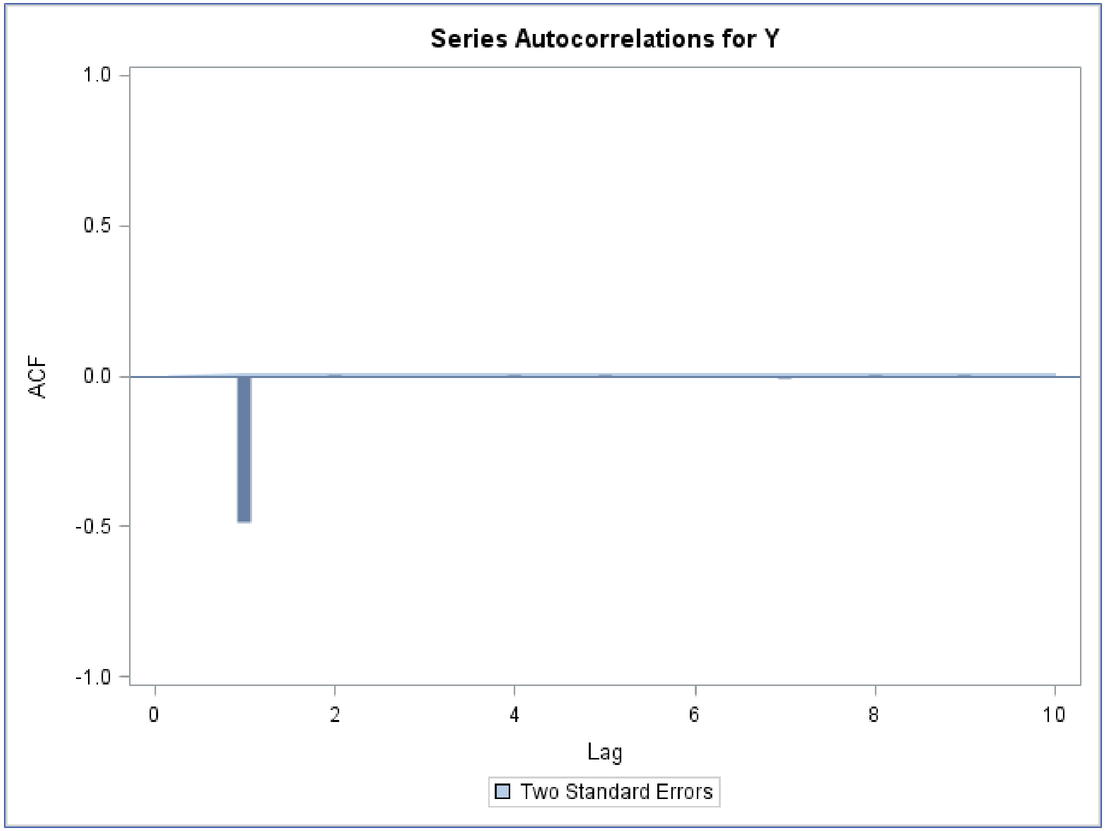
Correlation Functions for MA(1)

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

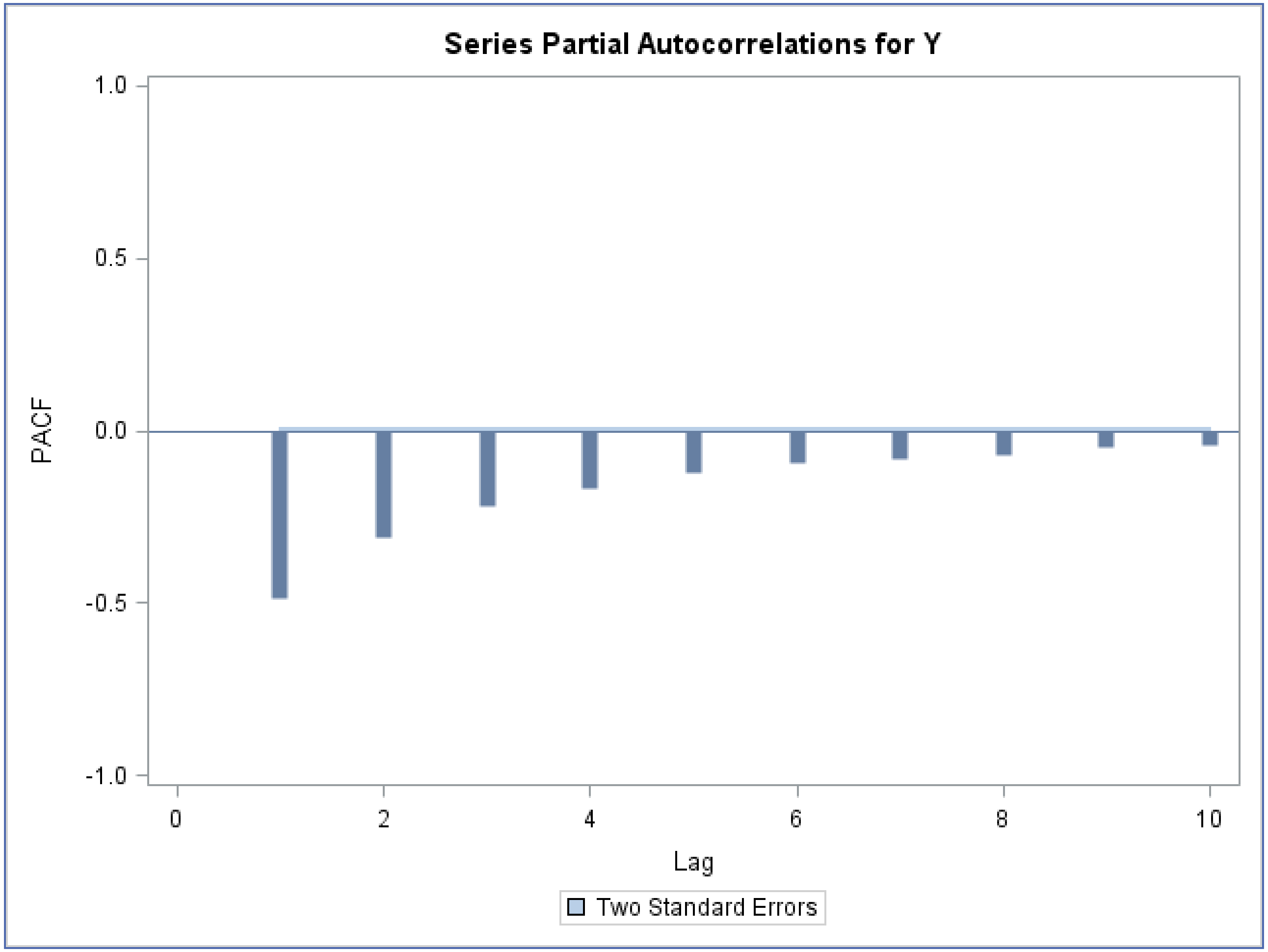
- Let's examine the following MA(1) model:

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

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

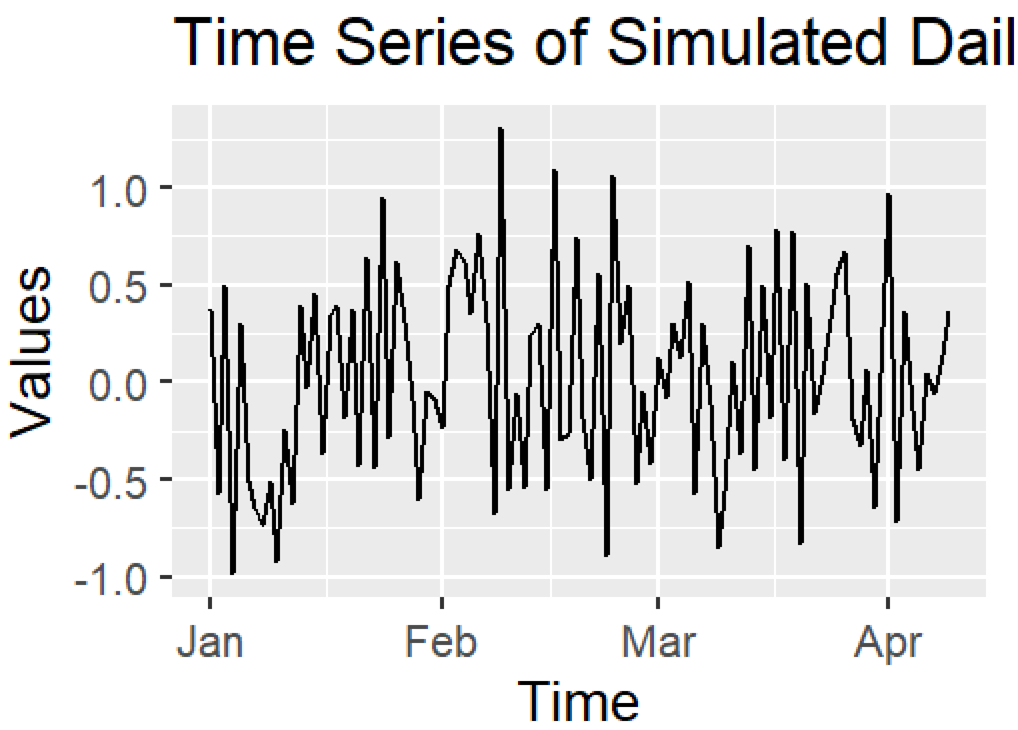
- A time series that is a linear function of *q* past errors is called a moving average process of order q – called an MA(q).

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

- The ACF for an MA(q) has significant spikes at lags up to lag *q*, followed by nothing after (kinda exponentially decreasing for PACF).

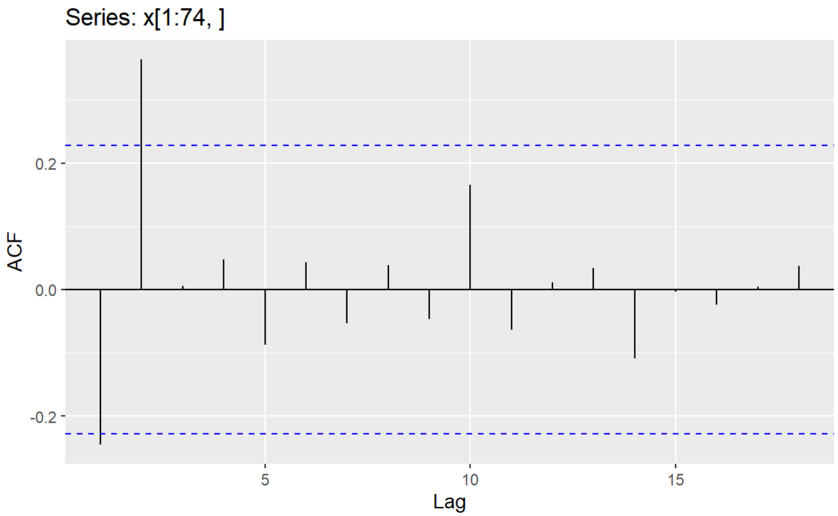
Simulated data set

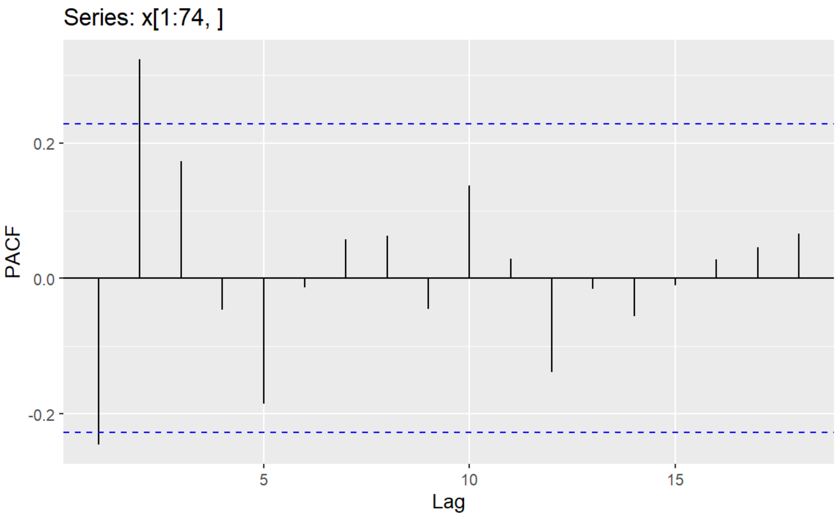
- The data set imported as “x” is simulated as an MA(2) model



MA(2)

```
ggAcf(x[1:74,])
ggPacf(x[1:74,])
```





Two spikes in the ACF

Exponentially decreasing in the PACF

Moving Average models

```
x.ARIMA <- x_train %>%
model(ARIMA(x~pdq(0,0,2)+PDQ(0,0,0)))

report(x.ARIMA)
```

Output

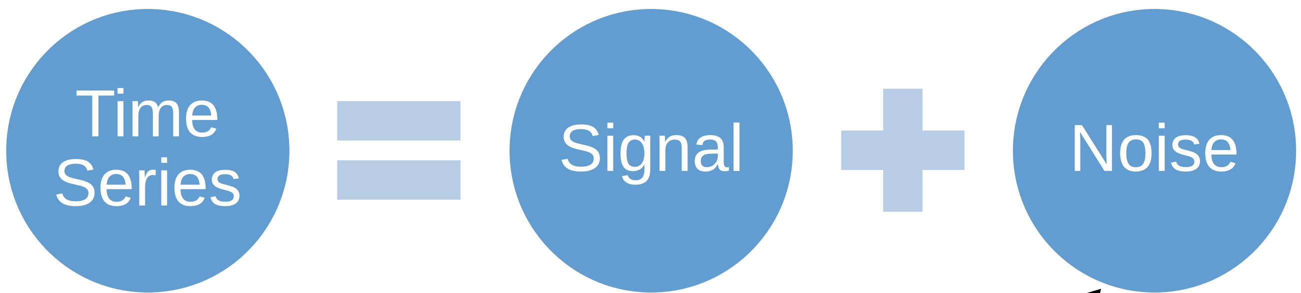
```
Series: x
Model: ARIMA(0,0,2)

Coefficients:
          ma1     ma2
      -0.2585  0.4874
s.e.   0.1031  0.1063

sigma^2 estimated as 0.2299:  log likelihood=-49.88
AIC=105.77   AICc=106.11   BIC=112.68
```

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 for MA models.

WHITE NOISE

Statistical Forecasting



If we are successful in removing all “correlation” 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.
- The autocorrelation and partial autocorrelation functions of the residuals from these models have a value close to zero at every time point (should NOT see any significant spikes).

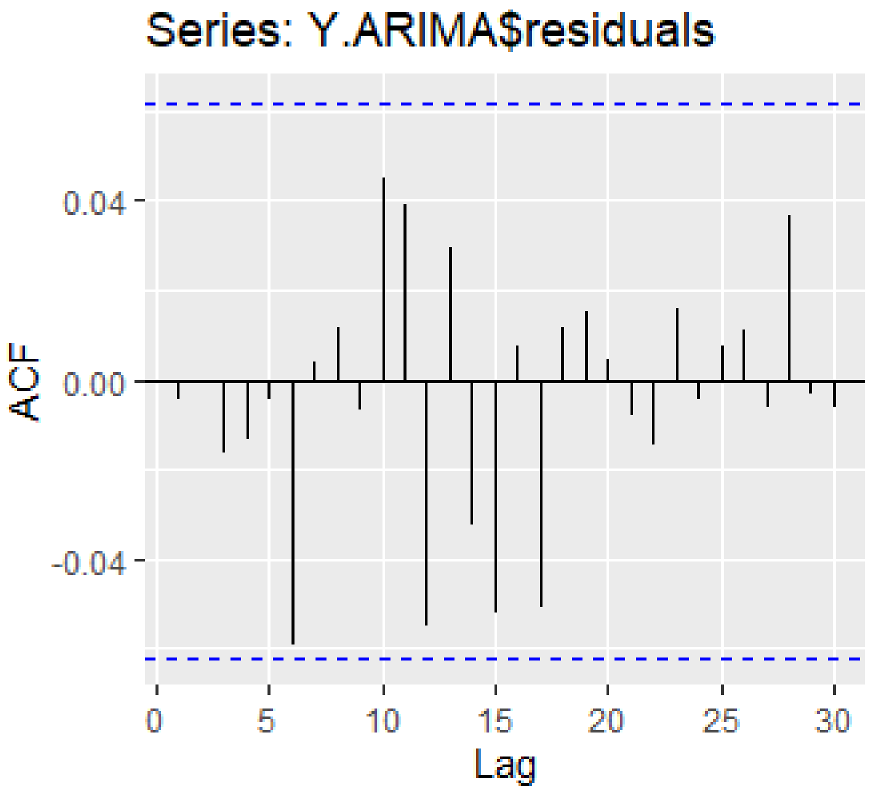
White Noise

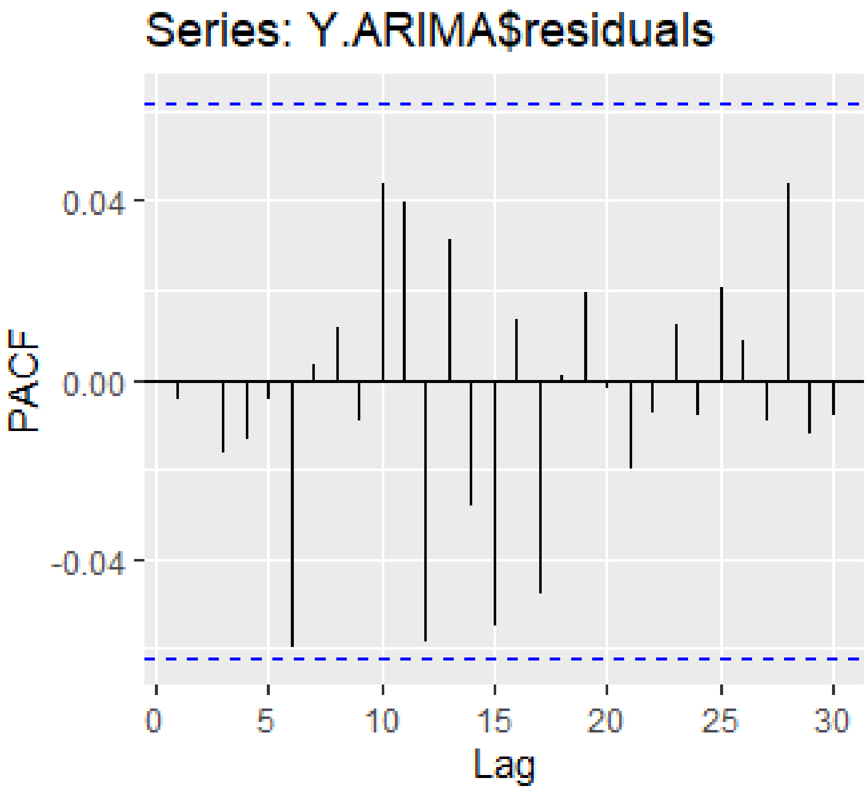
- The goal of modeling time series is to be left with white noise residuals in the time series.
- If the residuals still have a “significant” dependence structure, then more modeling can typically be done.
- How do we know when we are left with white noise at the end of the model? (you already know how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

White Noise

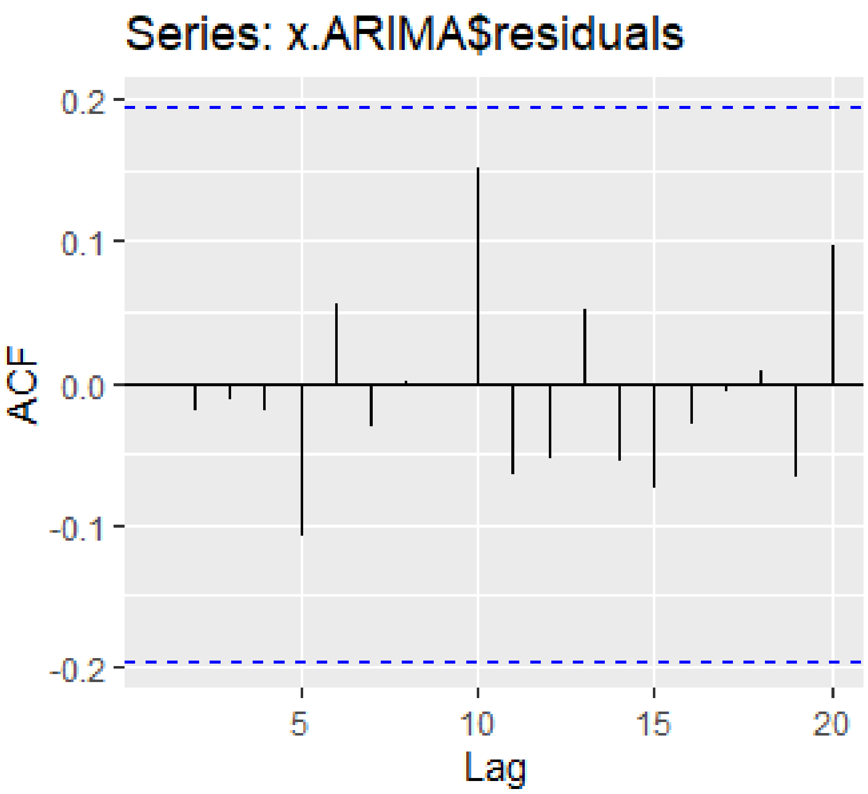
- There are two ways to investigate if there is not significant correlation left in the model
  - Graphical: Look at ACF and PACF plots of the residuals. Do you see any significant spikes (if you see “spikes”, this could indicate that you still have significant autocorrelation left to try to model).
  - Formal: Peform a Ljung-Box test on the residuals (H_0: No significant autocorrelation versus H_A: significant autocorrelation)

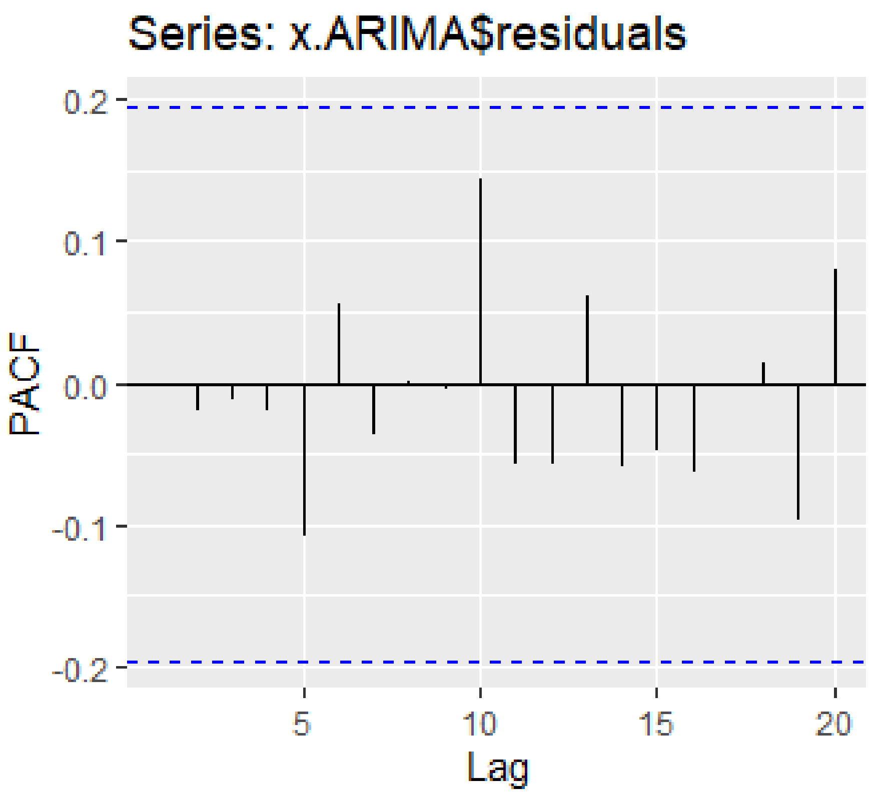
Residuals of AR(2) model





Residuals of MA(2) model





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

```
###Looking at original series first (BEFORE any
model….you don't really need to do this…..)
ljung_box(Y[1:731,], lag = 10, dof=0)

### After fitting model
augment(Y.ARIMA) %>%
  features(.innov,ljung_box, lag=10, dof = 2)
```

lb_stat	lb_pvalue
217.3408	0.0000

A tibble: 1 × 3

.model <chr>	lb_stat <dbl>	lb_pvalue <dbl>
ARIMA(Y ~ pdq(2, 0, 0) + PDQ(0, 0, 0))	6.306862	0.6129007

Testing for White Noise – R

###Looking at original series first (BEFORE any model….you don’t really need to do this…..)
ljung_box(Y[1:731,], lag = 10, dof=0)

After fitting model
augment(Y.ARIMA) %>%
  features(.innov,ljung_box, lag=10, dof = 2)

IF residuals, this is **p + q**

lb_stat	lb_pvalue
217.3408	0.0000

A tibble: 1 × 3

.model <chr>	lb_stat <dbl>	lb_pvalue <dbl>
ARIMA(Y ~ pdq(2, 0, 0) + PDQ(0, 0, 0))	6.306862	0.6129007