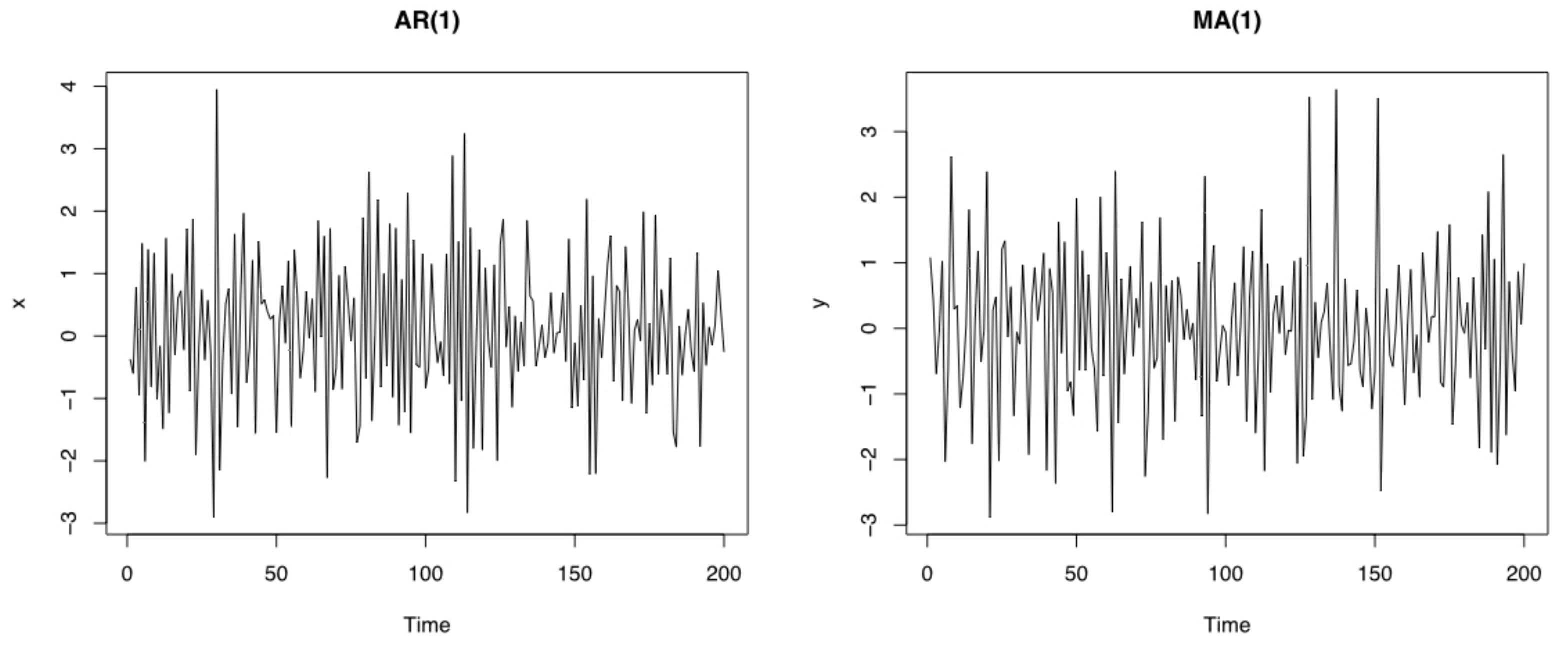
AR and MA models

ARIMA MODELS IN R

David Stoffer
Professor of Statistics at the University of Pittsburgh

AR and MA Models

```
x <- arima.sim(list(order = c(1, 0, 0), ar = -.7), n = 200)
y <- arima.sim(list(order = c(0, 0, 1), ma = -.7), n = 200)
par(mfrow = c(1, 2))
plot(x, main = "AR(1)")
plot(y, main = "MA(1)")
```

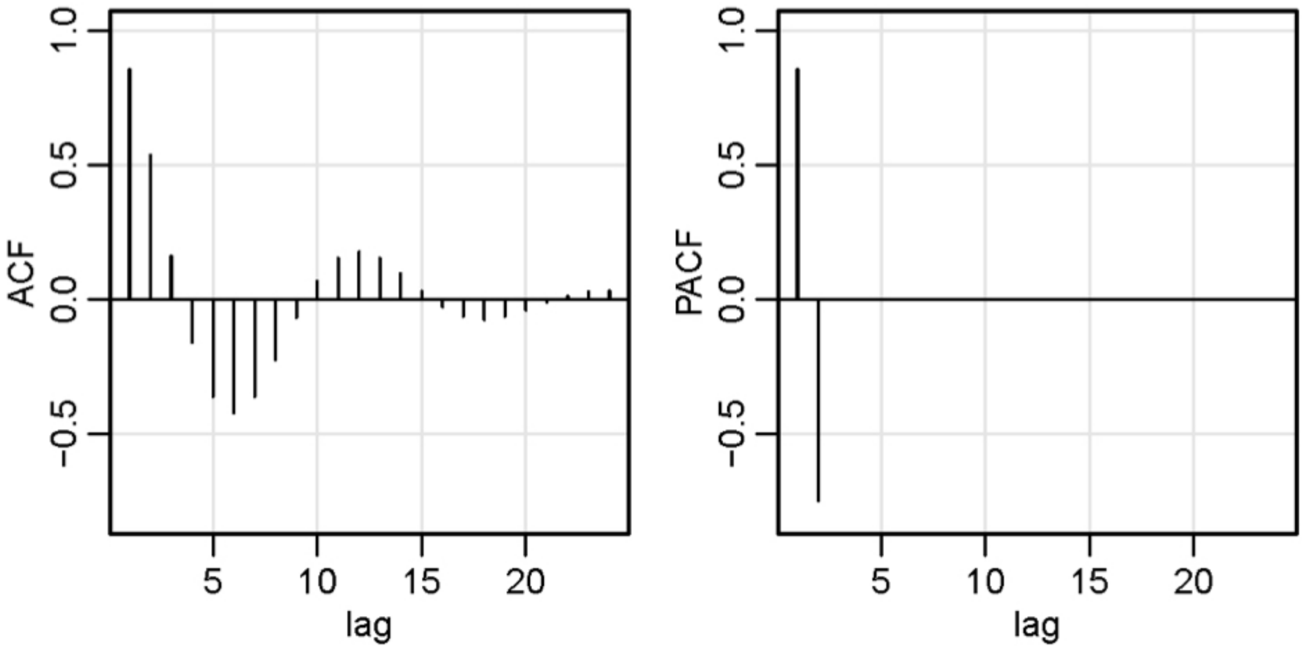


ACF and PACF

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off lag q	Tails off
PACF	Cuts off lag p	Tails off	Tails off

ACF and PACF

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off lag q	Tails off
PACF	Cuts off lag p	Tails off	Tails off



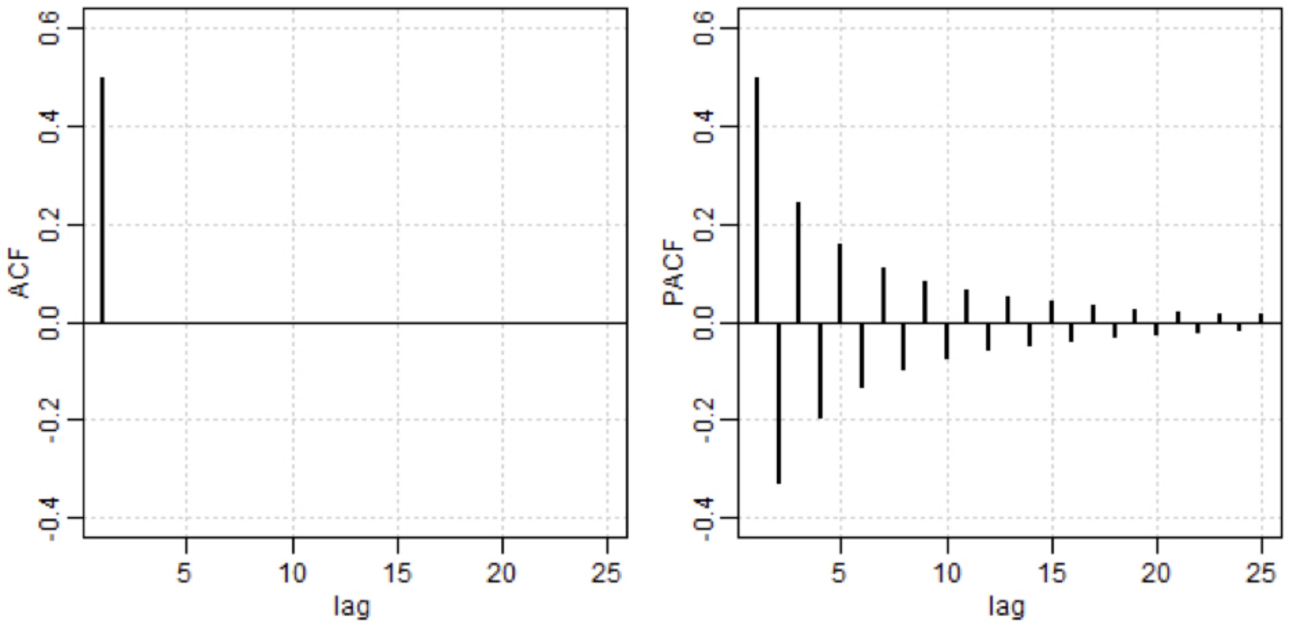
AR(2)

ACF and PACF

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off lag q	Tails off
PACF	Cuts off lag p	Tails off	Tails off

ACF and PACF

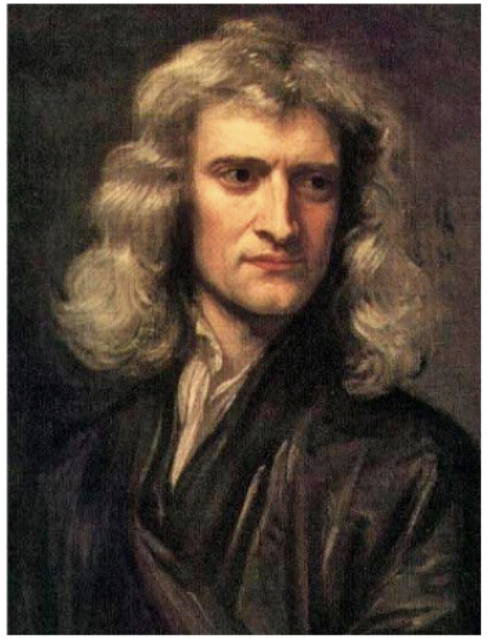
	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off lag q	Tails off
PACF	Cuts off lag p	Tails off	Tails off



MA(1)

Estimation

- Estimation for time series is similar to using least squares for regression
- Estimates are obtained numerically using ideas of Gauss and Newton



Estimation with astsa

- AR(2) with mean 50:

$$W_t = 50 + 1.5(X_{t-1} - 50) - .75(X_{t-2} - 50) + W_t$$

```
x <- arima.sim(list(order = c(2, 0, 0),
                    ar = c(1.5, -.75)),
                    n = 200) + 50
x_fit <- sarima(x, p = 2, d = 0, q = 0)
x_fit$ttable
```

```
      Estimate     SE  t.value p.value
ar1     1.5429 0.0435  35.4417       0
ar2    -0.7752 0.0434 -17.8650       0
xmean  49.6984 0.3057 162.5788       0
```

Estimation with astsa

- MA(1) with mean 0:

$$X_t = W_t - .7W_{t-1}$$

```
y <- arima.sim(list(order = c(0, 0, 1), ma = -.7), n = 200)
y_fit <- sarima(y, p = 0, d = 0, q = 1)
y_fit$ttable
```

```
      Estimate     SE   t.value  p.value
ma1    -0.7459 0.0513  -14.5470   0.0000
xmean   0.0324 0.0191    1.6946   0.0917
```

Let's practice!

ARIMA MODELS IN R

AR and MA together

ARIMA MODELS IN R

David Stoffer
Professor of Statistics at the University of Pittsburgh

AR and MA Together: ARMA

$$X_t = \phi X_{t-1} + W_t + \theta W_{t-1}$$

AR and MA Together: ARMA

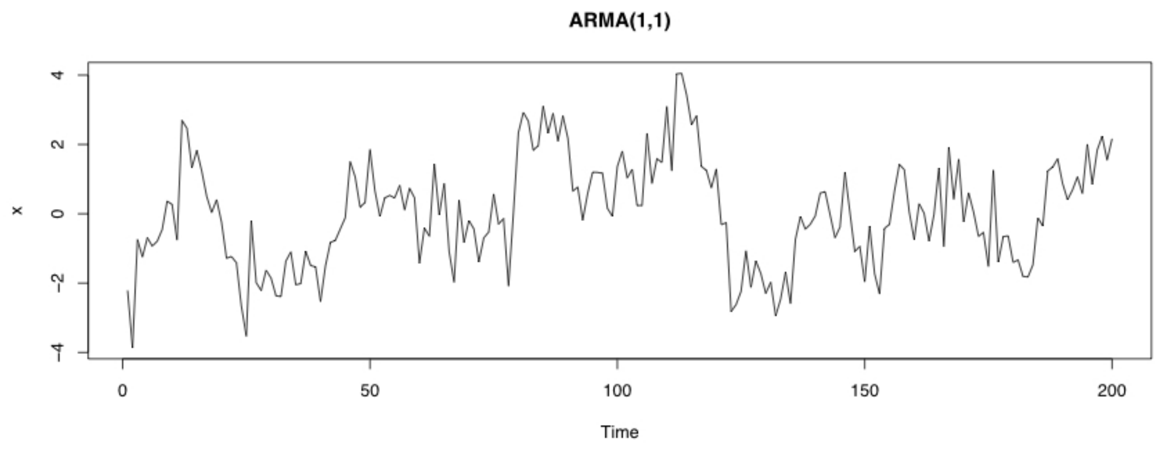
$$X_t = \phi X_{t-1} + W_t + \theta W_{t-1}$$

auto-regression with correlated errors

AR and MA Together: ARMA

$$X_t = \phi X_{t-1} + W_t + \theta W_{t-1}$$

auto-regression with correlated errors




```r
x <- arima.sim(list(order = c(1, 0, 1),
                    ar = .9,
                    ma = -.4),
                    n = 200)
plot(x,  main = "ARMA(1, 1)")
```

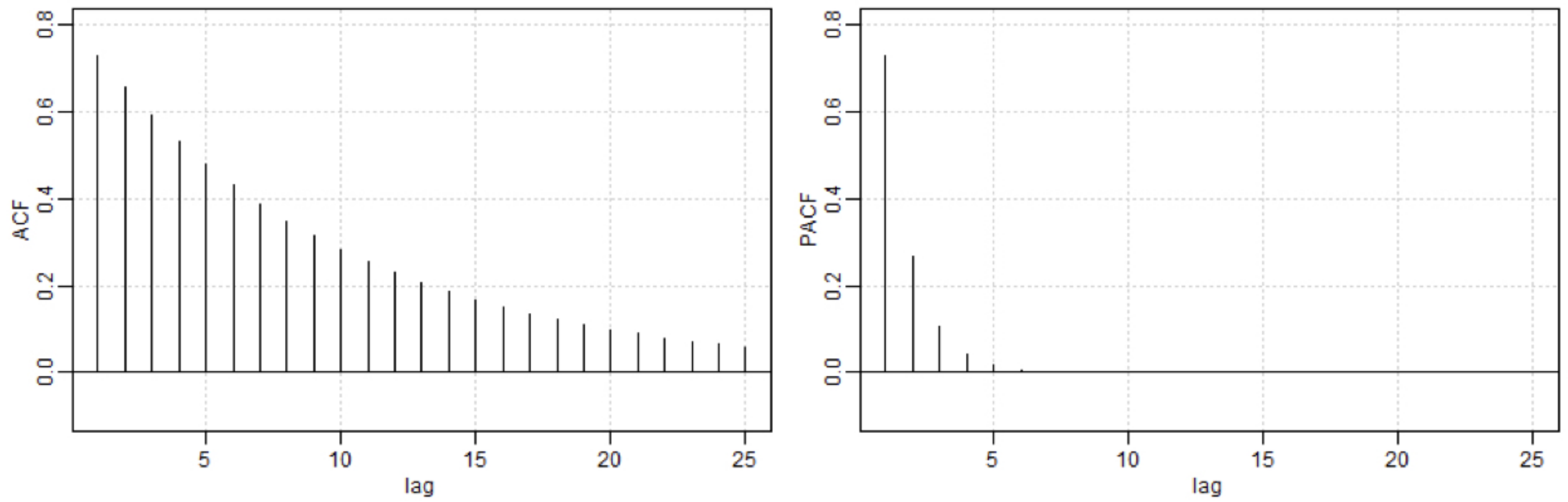
ACF and PACF of ARMA Models

	AR(p)	**MA(q)**	**ARMA(p, q)**
ACF	Tails off	Cuts off lag q	Tails off
PACF	Cuts off lag p	Tails off	Tails off

ACF and PACF of ARMA Models

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off lag q	Tails off
PACF	Cuts off lag p	Tails off	Tails off

$$X_t = .9X_{t-1} + W_t - .4W_{t-1}$$



Estimation

$$X_t = .9X_{t-1} + W_t - .4W_{t-1}$$

```
x <- arima.sim(list(order = c(1, 0, 1),
                    ar = .9,
                    ma = -.4),
                    n = 200)
x_fit <- sarima(x, p = 1, d = 0, q = 1)
x_fit$ttable
```

```
      Estimate     SE  t.value p.value
ar1     0.9083 0.0424  21.4036       0
ma1    -0.4458 0.0879  -5.0716       0
xmean  49.5647 0.4079 121.5026       0
```

Let's practice!

ARIMA MODELS IN R

Model choice and residual analysis

ARIMA MODELS IN R

David Stoffer
Professor of Statistics at the University of Pittsburgh

AIC and BIC

Error Number of Parameters

$$average(observed - predicted)^2 \quad + \quad k(p+q)$$

↓ ↑

- AIC and BIC measure the error and penalize (differently) for adding parameters
- For example, AIC has $k = 2$ and BIC has $k = log(n)$
- Goal: find the model with the smallest AIC or BIC

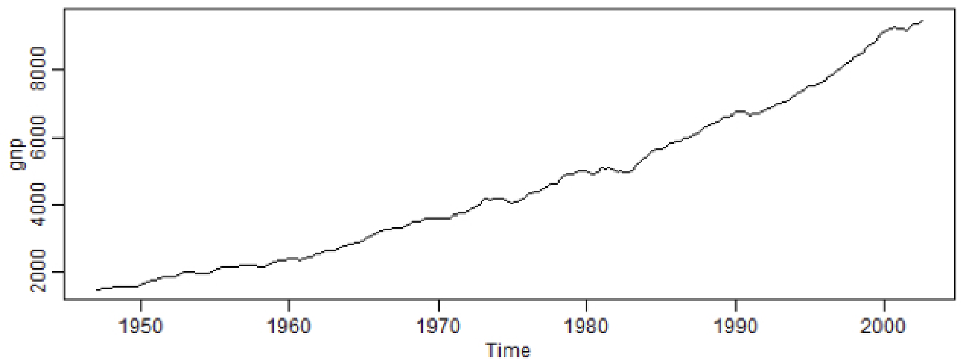
Model Choice: AR(1) vs. MA(2)

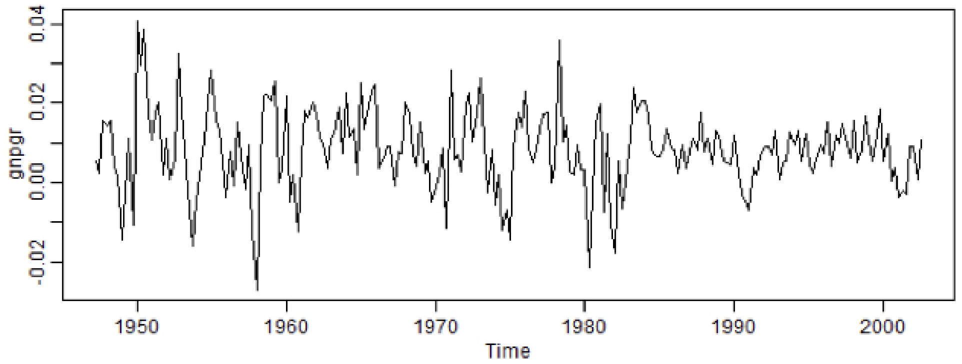
```r
gnpgr <- diff(log(gnp))
sarima(gnpgr, p = 1, d = 0, q = 0)
```

```
$AIC                          $BIC
-8.294403                     -9.263748
```

```r
sarima(gnpgr, p = 0, d = 0, q = 2)
```

```
$AIC                          $BIC
-8.297695                     -9.251712
```

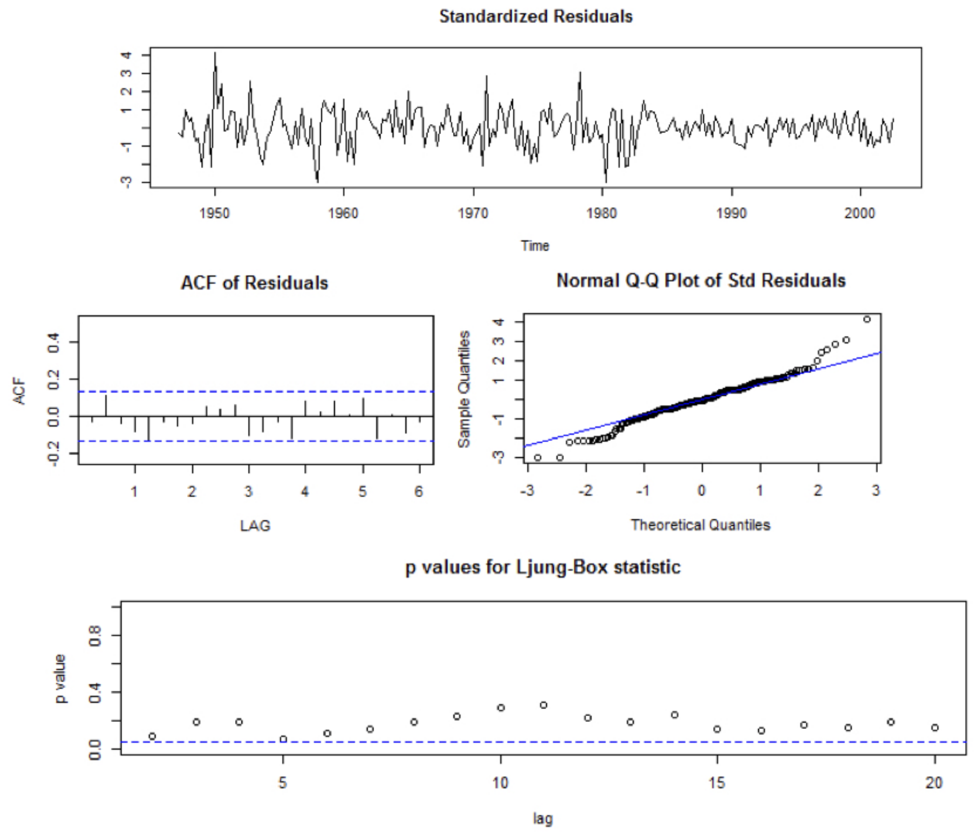




Residual Analysis

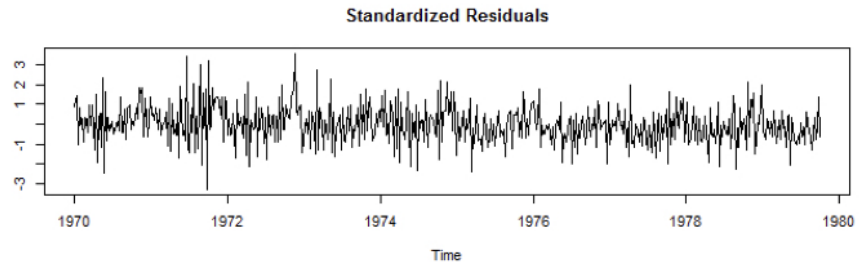
`sarima()` includes residual analysis graphic showing:

1. Standardized residuals
2. Sample ACF of residuals
3. Normal Q-Q plot
4. Q-statistic p-values



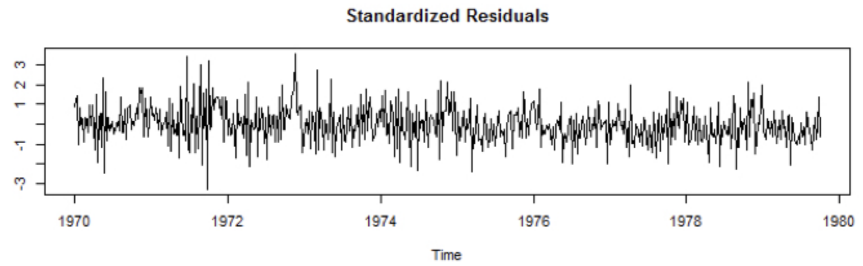
Bad Residuals

- Pattern in the residuals



Bad Residuals

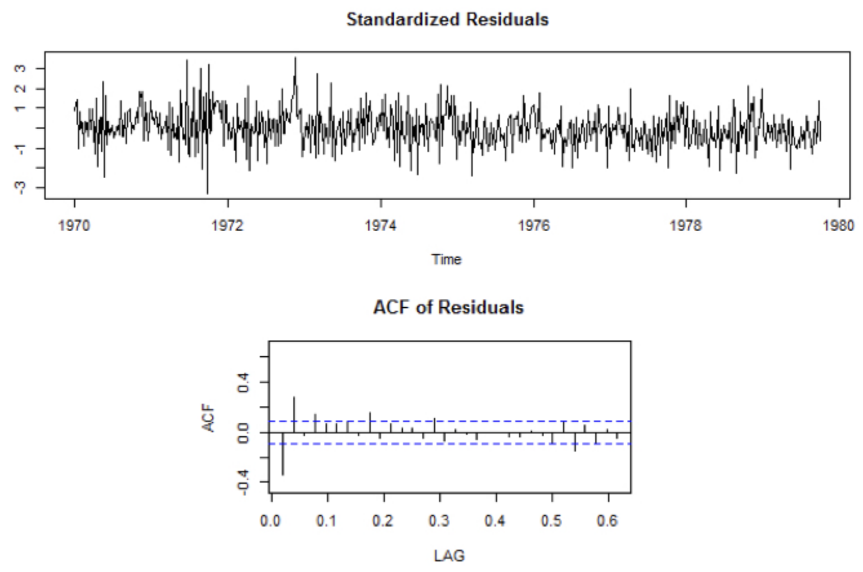
❌ Pattern in the residuals



Bad Residuals

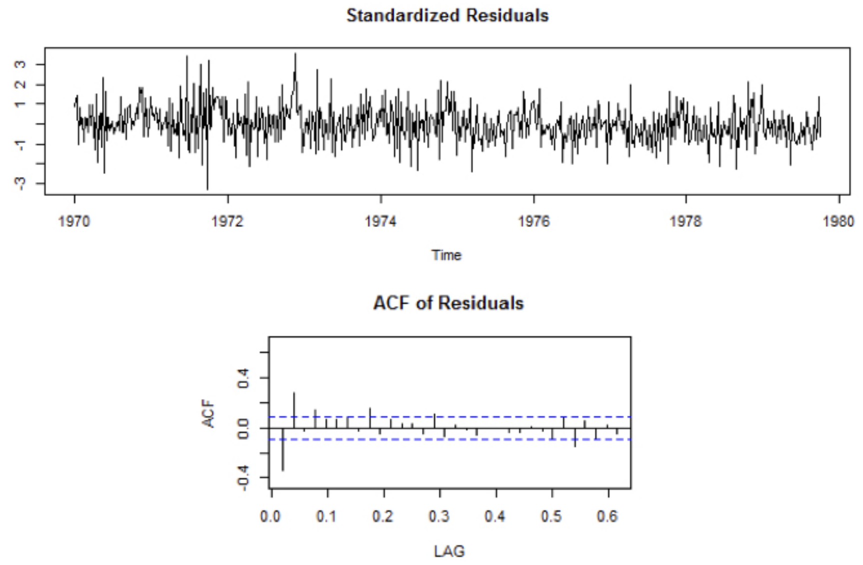
✗ Pattern in the residuals

- ACF has large values



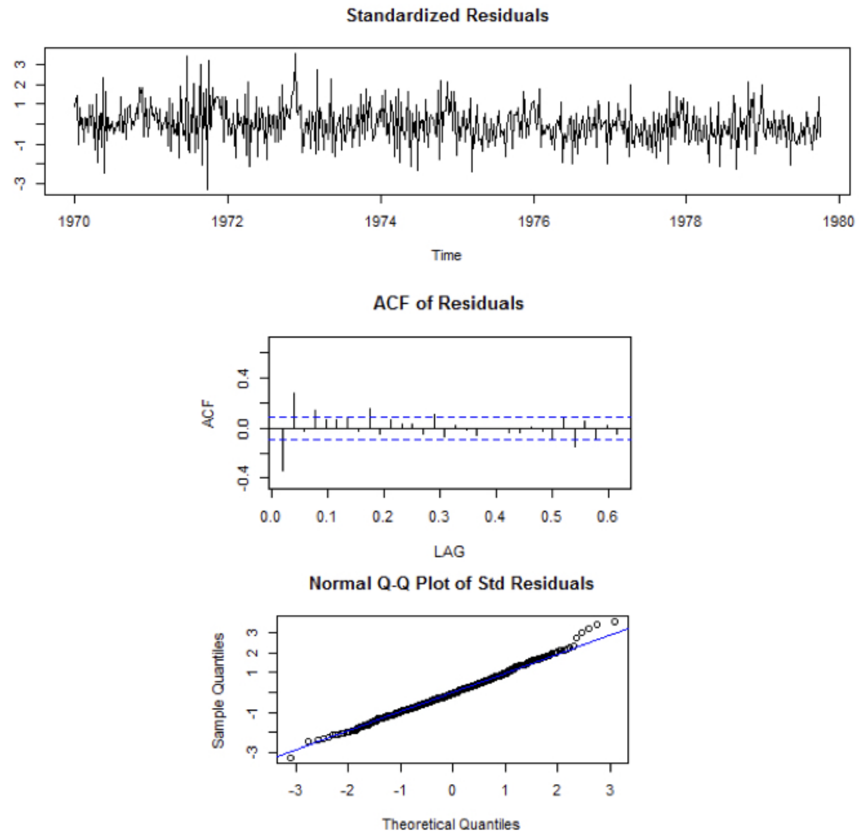
Bad Residuals

- ✗ Pattern in the residuals
- ✗ ACF has large values



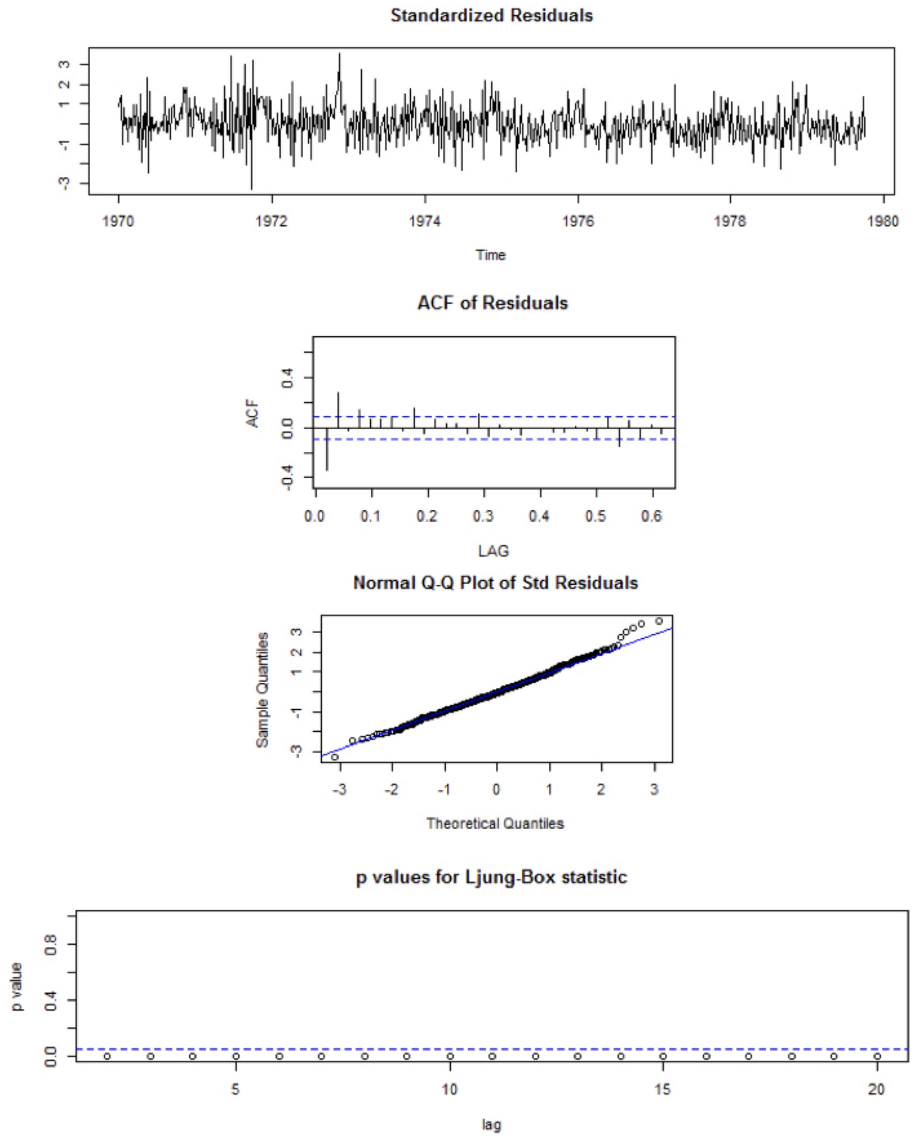
Bad Residuals

- ✕ Pattern in the residuals
- ✕ ACF has large values
- Q-Q plot suggests normality



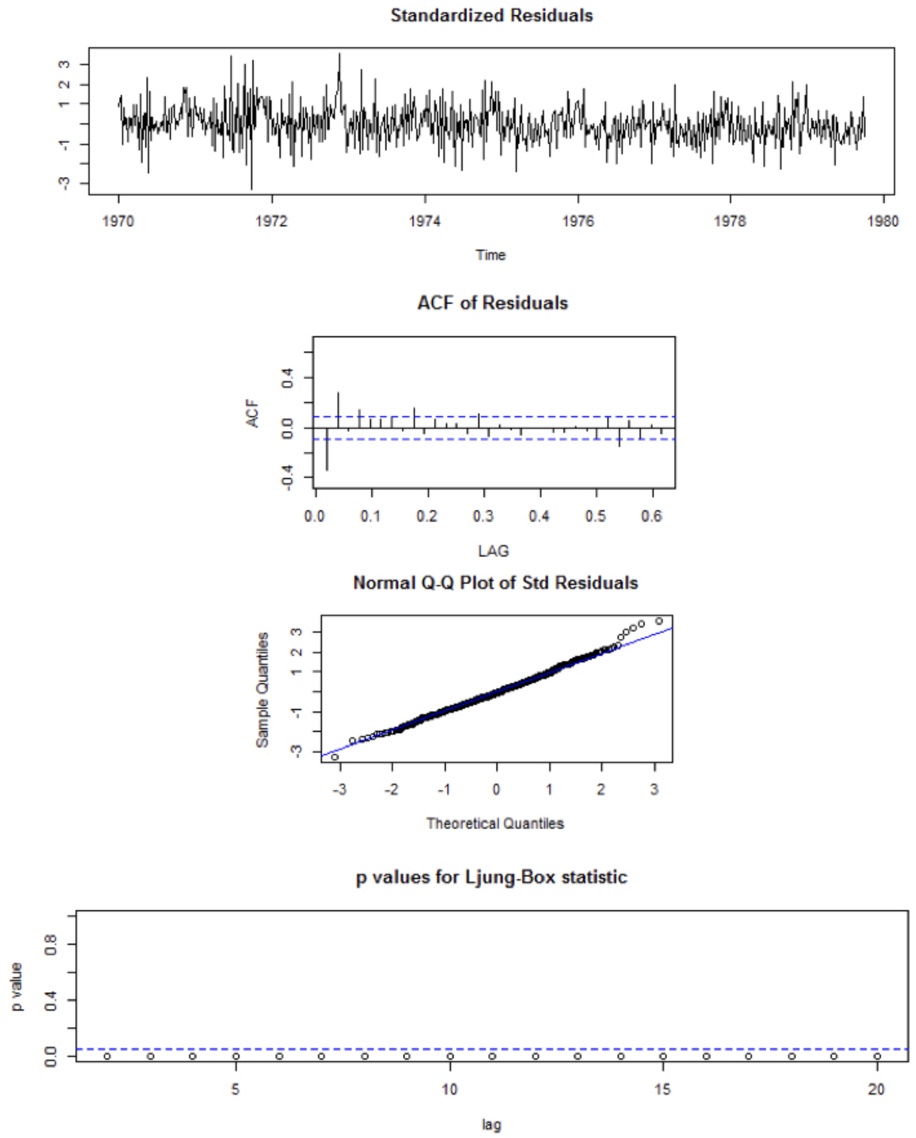
Bad Residuals

- ✗ Pattern in the residuals
- ✗ ACF has large values
- Q-Q plot suggests normality
- Q-statistic - all points below line



Bad Residuals

- ✗ Pattern in the residuals
- ✗ ACF has large values
- Q-Q plot suggests normality
- ✗ Q-statistic - all points below line



Let's practice!

ARIMA MODELS IN R