# Ch 6. Model Specification

Time Series Analysis

We start to build ARIMA(p,d,q) models. The subjects include:

1. how to determine $p$, $d$, $q$ for a given series (Chapter 6);
2. how to estimate the parameters ($\phi$'s and $\theta$'s) of a specific ARIMA(p,d,q) model (Chapter 7);
3. how to evaluate the fitted model and improve it if needed (Chapter 8).

A standard model building process go through the above three procedures in order. We may iterate the process several times to improve the model.

## 6.1. Properties of the Sample ACF

Recall the sample autocorrelation function (sample ACF) for an observed series $\{Y_t\}_{t=1}^n$ is

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \overline{Y})(Y_{t-k} - \overline{Y})}{\sum_{t=1}^n (Y_t - \overline{Y})^2}, \qquad k = 1, 2, 3, \cdots$$

We use the sample ACF $\{r_k\}$ to estimate the ACF $\{\rho_k\}$, then use the pattern of ACF to decide the model. There are three basic questions about the sample ACF:

1. Is sample ACF a unbiased estimator of ACF ($E(r_k) = \rho_k$)?
2. How precise is our estimation ($\mathrm{Var}(\mathrm{r_k})$)?
3. What is the correlation $\mathrm{Corr}(\mathrm{r_k}, \mathrm{r_j})$ of $r_k$ and $r_j$?

## Theorem 1

*Suppose* $\{Y_t\}$ *is a stationary ARMA process. For any fixed* $m$*, the joint distribution of*

$$\sqrt{n}(r_1 - \rho_1),\ \sqrt{n}(r_2 - \rho_2),\ \cdots,\ \sqrt{n}(r_m - \rho_m)$$

*approaches, as* $n \to \infty$*, a joint normal distribution with zero means and certain covariance matrix* $\left[\, c_{ij} \,\right]_{m \times m}$ *(see (6.1.2) on pp110.)*

$$c_{ij} = \sum_{k=-\infty}^{\infty} (\rho_{k+i}\rho_{k+j} + \rho_{k-i}\rho_{k+j} - 2\rho_i\rho_k\rho_{k+j} - 2\rho_j\rho_k\rho_{k+i} + 2\rho_i\rho_j\rho_k^2) \qquad (6.1.2)$$

*In particular, for large* $n$*,*

1. *each* $r_k \approx N(\rho_k, \frac{c_{kk}}{n})$ *(so* $r_k$ *is an unbiased estimator of* $\rho_k$ *with the variance inversely proportional to* $n$*);*
2. $\mathrm{Corr}\,(\mathrm{r_k}, \mathrm{r_j}) \approx \frac{\mathrm{c_{kj}}}{\sqrt{\mathrm{c_{kk}c_{jj}}}}$ *is approximately constant for large* $n$*.*

The sample ACF behaves much simpler for some special cases:

- If $\{Y_t\}$ is white noise, then $\mathrm{Var}\left(\mathrm{r_k}\right) \approx \frac{1}{\mathrm{n}}$ and $\mathrm{Corr}\left(\mathrm{r_k}, \mathrm{r_j}\right) \approx 0$ for $k \neq j$.
- If $\{Y_t\} \sim AR(1)$, we get

$$\mathrm{Var}\left(\mathrm{r}_1\right) \approx \frac{1-\phi^2}{\mathrm{n}}.$$

The closer $\phi$ is to $\pm 1$, the more precise our estimate of $\rho_1 (= \phi)$ becomes.

$$\mathrm{Var}\left(\mathrm{r_k}\right) \approx \frac{1}{\mathrm{n}}\left[\frac{1+\phi^2}{1-\phi^2}\right] \quad \text{for large } \mathrm{k}.$$

The closer $\phi$ is to $\pm 1$, the more variance our estimate of large $\rho_k \approx \phi^k$ is.

$$\mathrm{Corr}\left(\mathrm{r}_1, \mathrm{r}_2\right) \approx 2\phi\sqrt{\frac{1-\phi^2}{1+2\phi^2-3\phi^4}}.$$

Exhibit 6.1 on pp111 gives approximate standard deviations and correlations for several lags and a few values of $\phi$ in AR(1) models.

**Exhibit 6.1 Large Sample Results for Selected $r_k$ from an AR(1) Model**

| $\phi$ | $\sqrt{Var(r_1)}$ | $\sqrt{Var(r_2)}$ | $Corr(r_1, r_2)$ | $\sqrt{Var(r_{10})}$ |
|---|---|---|---|---|
| ±0.9 | $0.44/\sqrt{n}$ | $0.807/\sqrt{n}$ | ±0.97 | $2.44/\sqrt{n}$ |
| ±0.7 | $0.71/\sqrt{n}$ | $1.12/\sqrt{n}$ | ±0.89 | $1.70/\sqrt{n}$ |
| ±0.4 | $0.92/\sqrt{n}$ | $1.11/\sqrt{n}$ | ±0.66 | $1.18/\sqrt{n}$ |
| ±0.2 | $0.98/\sqrt{n}$ | $1.04/\sqrt{n}$ | ±0.38 | $1.04/\sqrt{n}$ |

- For $\{Y_t\} \sim MA(1)$, we have $\mathrm{Var}\,(\mathrm{r_k}) \approx \mathrm{c_{kk}/n}$ and $\mathrm{Corr}\,(\mathrm{r_k}, \mathrm{r_j}) \approx \frac{\mathrm{c_{kj}}}{\sqrt{\mathrm{c_{kk}c_{jj}}}}$ where

$$c_{11} = 1-3\rho_1^2+4\rho_1^4, \quad c_{kk} = 1+2\rho_1^2 \text{ for } k > 1, \quad c_{12} = 2\rho_1(1-\rho_1^2).$$

Exhibit 6.2 lists large-sample standard deviations and correlations for the sample ACF for several lags and $\theta$-values.

**Exhibit 6.2 Large-Sample Results for Selected $r_k$ from an MA(1) Model**

| $\theta$ | $\sqrt{Var(r_1)}$ | $\sqrt{Var(r_k)}$ for $k>1$ | $Corr(r_1, r_2)$ |
|---|---|---|---|
| $\pm 0.9$ | $0.71/\sqrt{n}$ | $1.22/\sqrt{n}$ | $\mp 0.86$ |
| $\pm 0.7$ | $0.73/\sqrt{n}$ | $1.20/\sqrt{n}$ | $\mp 0.84$ |
| $\pm 0.5$ | $0.79/\sqrt{n}$ | $1.15/\sqrt{n}$ | $\mp 0.74$ |
| $\pm 0.4$ | $0.89/\sqrt{n}$ | $1.11/\sqrt{n}$ | $\mp 0.53$ |

The sample ACF can be highly correlated and that the standard deviation of $r_k$ is larger for $k > 1$ than for $k = 1$.

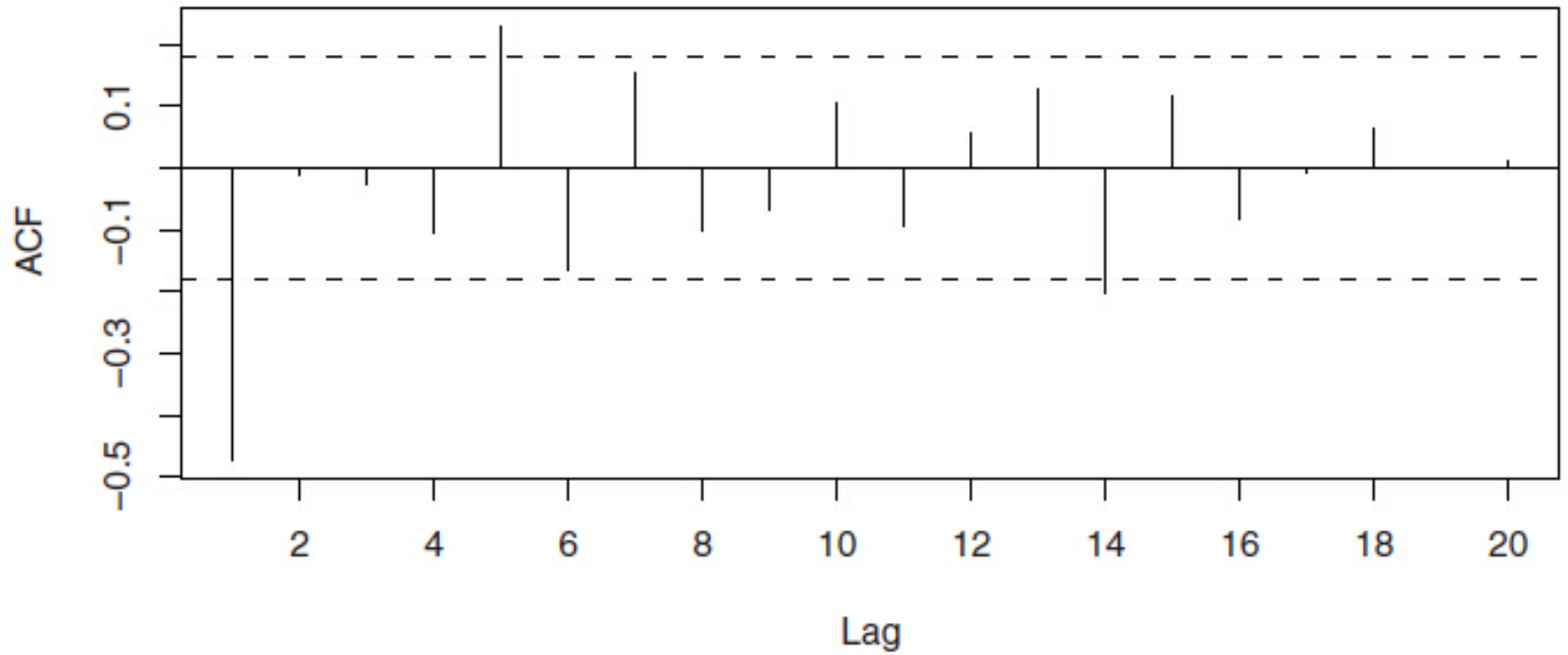
- For $\{Y_t\} \sim MA(q)$, we get $\mathrm{Er_k} = \rho_\mathrm{k} = 0$ for $k > q$,

$$\mathrm{Var}\left(\mathrm{r_k}\right) = \frac{1}{\mathrm{n}}\left[1 + 2\sum_{\mathrm{j}=1}^{\mathrm{q}} \rho_\mathrm{j}^2\right] \quad \text{for } \mathrm{k} > \mathrm{q}. \tag{1}$$

**Theorem 2**

*A test of the hypothesis that an observed series is MA(q) could be done by comparing $r_k$ $(k > q)$ to $\pm 2$ standard errors, where the standard error is estimated by $\sqrt{\mathrm{Var}\left(\mathrm{r_k}\right)}$ using the above formula and replacing $\rho$'s by r's. We would reject the null hypothesis iff $r_k$ lies outside these bounds.*

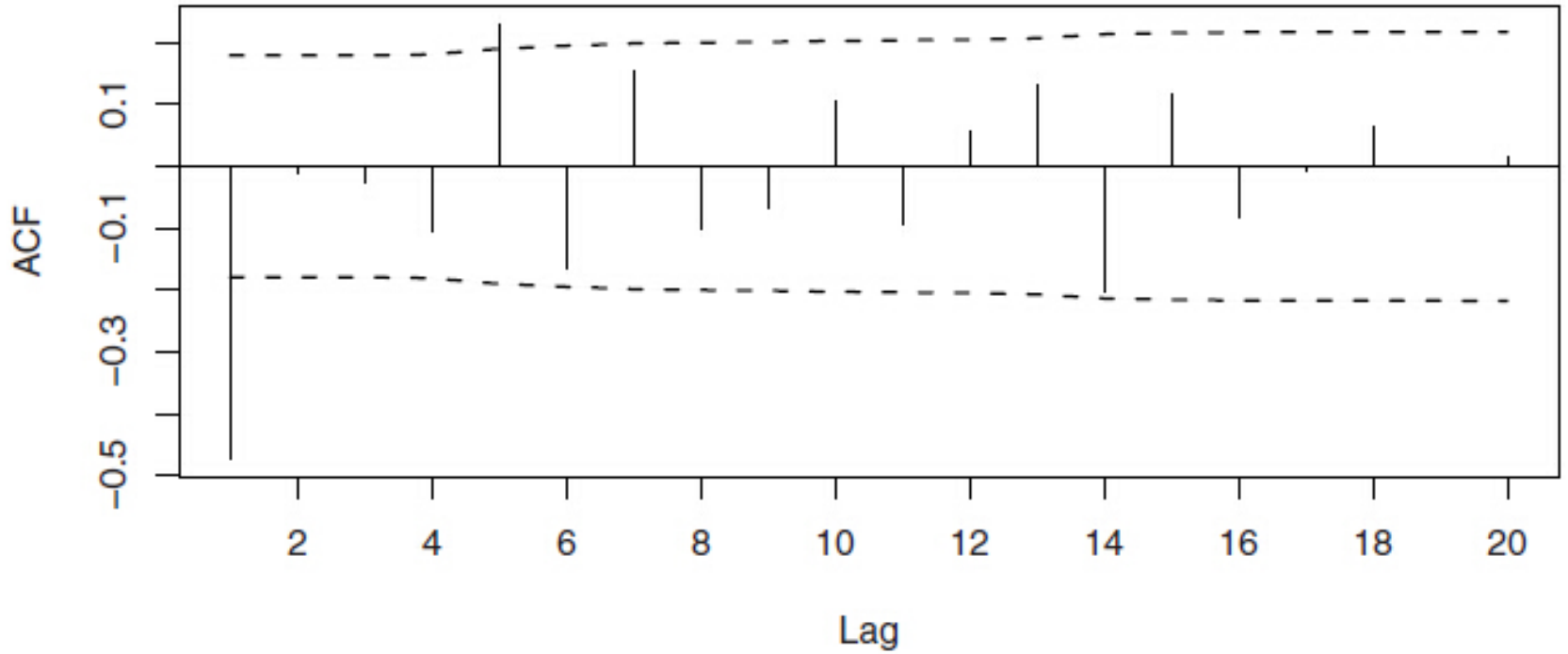
## Exhibit 6.5 Sample Autocorrelation of an MA(1) Process with $\theta = 0.9$




```
> data(ma1.1.s)
> win.graph(width=4.875,height=3,pointsize=8)
> acf(ma1.1.s,xaxp=c(0,20,10))
```

The dashed horizontal lines plotted at $\pm 2/\sqrt{n}$ are used to test whether or not the ACFs are significantly different from 0. The limits are based on the approximate large sample standard error that applies to a white noise process, namely $1/\sqrt{n}$. The sample ACF values exceed these rough critical values at lags 1, 5, and 14.

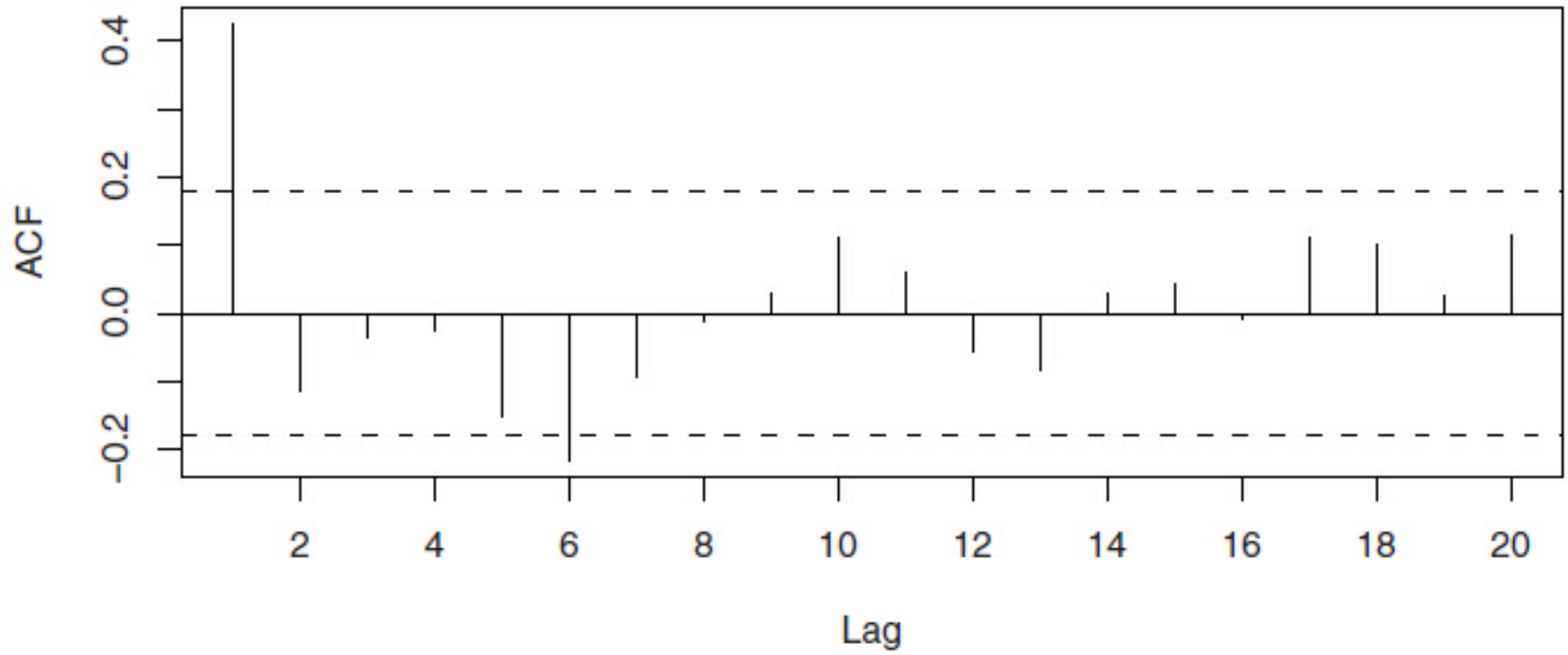
**Exhibit 6.6 Alternative Bounds for the Sample ACF for the MA(1) Process**




```
> acf(ma1.1.s,ci.type='ma',xaxp=c(0,20,10))
```

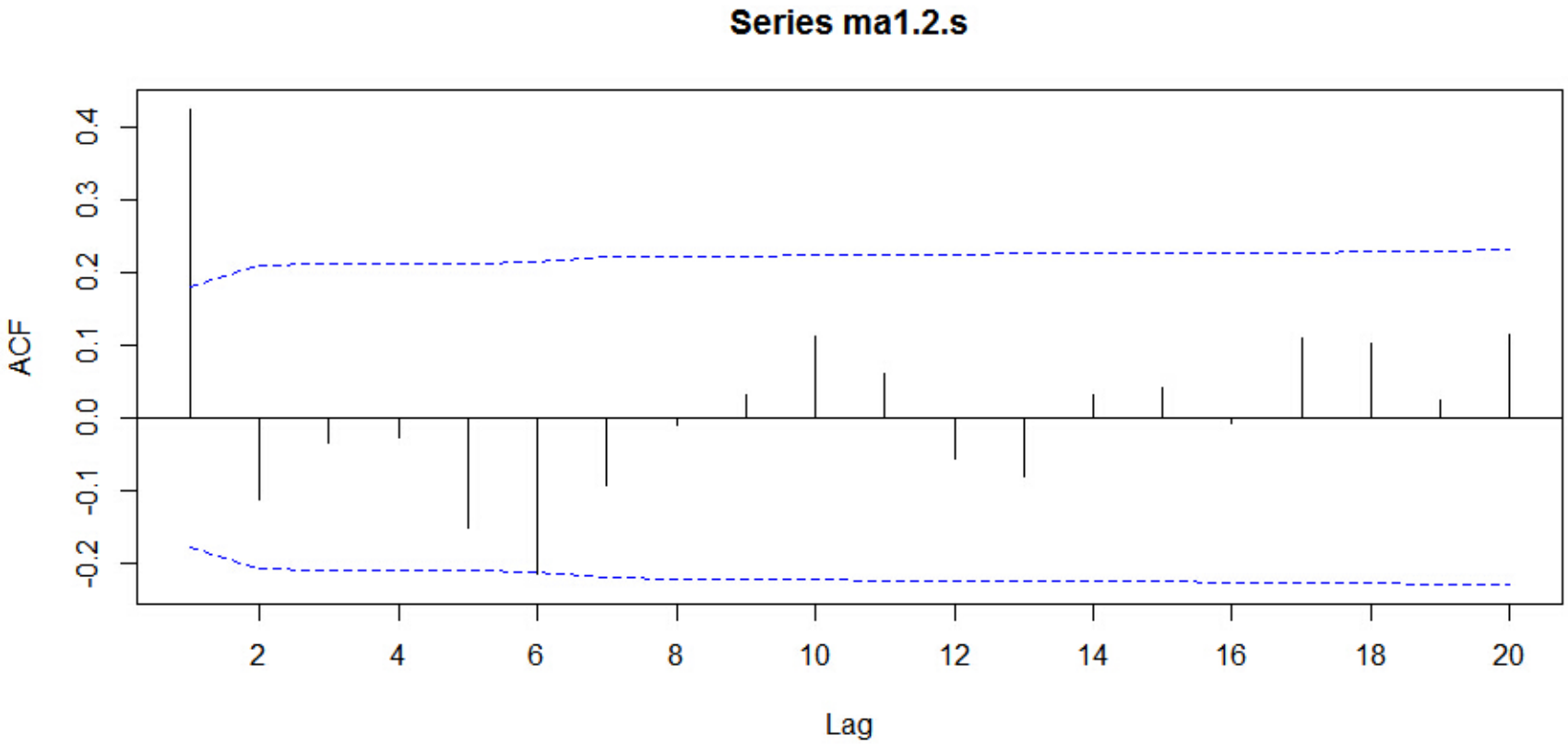
The same sample ACF with critical bounds calculated by (1). We replace $\rho$'s by $r$'s, let $q$ equal $1, 2, 3, \ldots$ successively, and take the square root. The sample ACF at lag 5 is barely significant. This plot suggests us to consider a MA(1) model.

**Exhibit 6.7 Sample Autocorrelation for an MA(1) Process with $\theta = -0.9$**



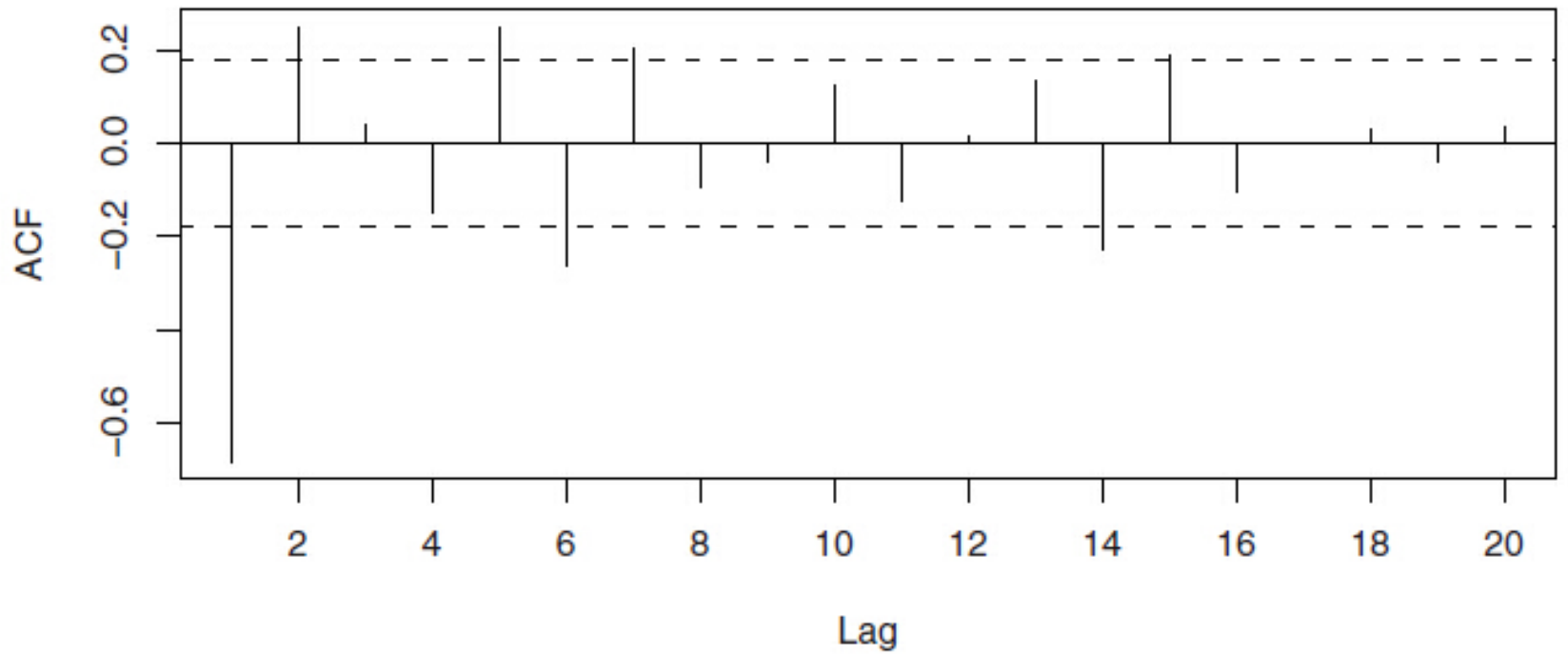

```
> data(ma1.2.s); acf(ma1.2.s,xaxp=c(0,20,10))
```

As the second example, Exhibit 6.7 displays the sample ACF for an MA(1) Process `ma1.2.s` with $\theta = -0.9$. The critical values based on the approximately white noise limits point to a MA(1) model. However, the lag 6 sample ACF looks really significant.




If we use the new limits for the sample ACF of the dataset `ma1.2.s`, the lag 6 sample ACF barely pass the line.

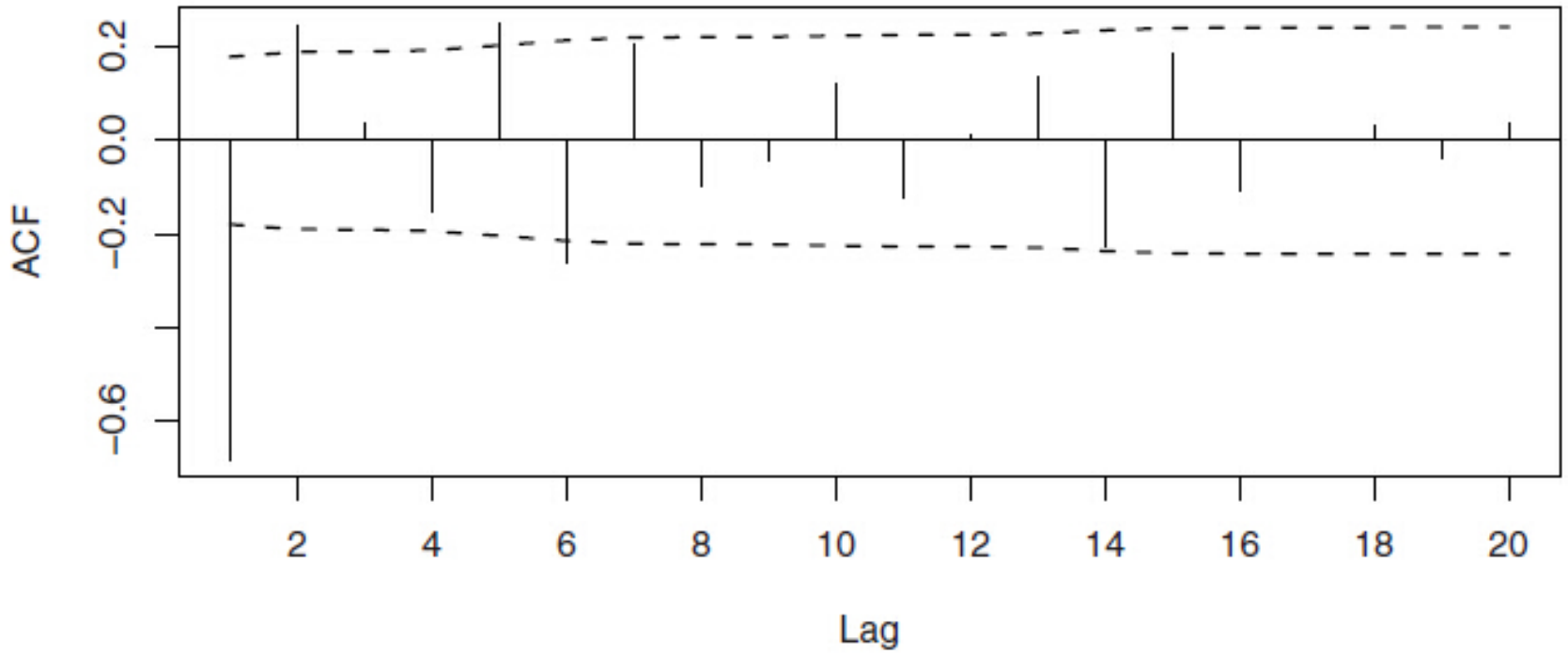
**Exhibit 6.8 Sample ACF for an MA(2) Process with $\theta_1 = 1$ and $\theta_2 = -0.6$**




```
> data(ma2.s); acf(ma2.s,xaxp=c(0,20,10))
```

For dataset `ma2.s`, the sample ACF displays significance at lags 1, 2, 5, 6, 7, 14 and 15 based on the simple standard error bounds.

**Exhibit 6.9 Alternative Bounds for the Sample ACF for the MA(2) Process**




```
> acf(ma2.s,ci.type='ma',xaxp=c(0,20,10))
```

With the more sophisticated standard error bounds, the lag 7,14, 15 sample ACF of `ma2.s` is no longer significant. Later we will see that MA(2) is the most appropriate model for these data.

**Ex.** [HW6.12] From a time series of 100 observations, we calculate

$$r_1 = -0.49, \quad r_2 = 0.31, \quad r_3 = -0.21, \quad r_4 = 0.11;$$
$$|r_k| < 0.09 \text{ for } k > 4.$$

On this basis alone, what ARIMA model would we tentatively specify for the series?

## 6.2. The Partial and Extended Autocorrelation Functions

For MA(q) models, $\rho_k = 0$ for $k > q$. So the ACF/sample ACF is a good indicator for the order of the process. However, the ACF of AR(p) models does not cut off. We need new approaches./new functions to determine the type and the order of AR(p) models.

**Ex.** Consider the AR(1) model $Y_t = \phi Y_{t-1} + e_t$, in which $\gamma_2 = \phi^2 \gamma_0 \neq 0$ because $Y_t$ is dependent on $Y_{t-2}$ through $Y_{t-1}$. Suppose we break this chain of dependence by removing (or partial out) the effect of $Y_{t-1}$. The least square regression of $Y_t$ on $Y_{t-1}$ is $\rho_1 Y_{t-1} = \phi Y_{t-1}$. We have

$$\mathrm{Cov}\left(\mathrm{Y_t} - \phi \mathrm{Y_{t-1}}, \mathrm{Y_{t-2}} - \phi \mathrm{Y_{t-1}}\right) = \mathrm{Cov}\left(\mathrm{e_t}, \mathrm{Y_{t-2}} - \phi \mathrm{Y_{t-1}}\right) = 0.$$

## 6.2.1. PACF

**Def.** The **partial autocorrelation at lag** $k$ is the correlation between $Y_t$ and $Y_{t-k}$ after removing the effect of the intervening variables $Y_{t-1}, Y_{t-2}, \cdots, Y_{t-k+1}$.

1. If $\{Y_{tt}\}$ is normally distributed, we can let

$$\phi_{kk} = \mathrm{Corr}\left(\mathrm{Y_t}, \mathrm{Y_{t-k}} \mid \mathrm{Y_{t-1}}, \mathrm{Y_{t-2}}, \cdots, \mathrm{Y_{t-k+1}}\right).$$

2. If we predict $Y_t$ and $Y_{t-k}$ based on a least square estimation of a linear function of $Y_{t-1}, \cdots, Y_{t-k+1}$, we get the predictors

$$\begin{aligned} \hat{Y}_t &= \beta_1 Y_{t-1} + \cdots + \beta_{k-1} Y_{t-k+1}, \\ \hat{Y}_{t-k} &= \beta_1 Y_{t-k+1} + \cdots + \beta_{k-1} Y_{t-1}, \end{aligned}$$

(where the coefficients are coincide due to stationarity.) The **partial autocorrelation function (PACF)** at lag $k$ is then defined to be the correlation between the prediction errors:

$$\begin{aligned} \phi_{kk} = \mathrm{Corr}( & Y_t - \beta_1 Y_{t-1} - \cdots - \beta_{k-1} Y_{t-k+1}, \\ & Y_{t-k} - \beta_1 Y_{t-k+1} - \cdots - \beta_{k-1} Y_{t-1}). \end{aligned}$$

We now study the PACF for AR(p) and MA(q) models:

- (AR(p):) For $k > p$, the best linear predictor of $Y_t$ w.r.t. $Y_{t-1}, Y_{t-2}, \cdots, Y_p, \cdots, Y_{t-k+1}$ is $\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p}$ (Chapter 9). Let $\hat{Y}_{t-p}$ be the linear predictor of $Y_{t-p}$ w.r.t. $Y_{t-1}, Y_{t-2}, \cdots, Y_{t-k+1}$. Then

$$\begin{aligned} \phi_{kk} &= \mathrm{Corr}\left(\mathrm{Y_t} - \phi_1 \mathrm{Y_{t-1}} - \cdots - \phi_\mathrm{p} \mathrm{Y_{t-p}}, \mathrm{Y_{t-p}} - \hat{\mathrm{Y}}_{\mathrm{t-p}}\right) \\ &= \mathrm{Corr}\left(\mathrm{e_t}, \mathrm{Y_{t-p}} - \hat{\mathrm{Y}}_{\mathrm{t-p}}\right) = 0. \end{aligned}$$

**Theorem 3**

*The PACF of an AR(p) model satisfies that $\phi_{kk} = 0$ for $k > p$.*

- (MA(1) & MA(q):) We will get

$$\phi_{kk} = -\frac{\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}}, \qquad \text{for } k \geq 1.$$

The PACF of an MA(1) model decays to zero exponentially. In general, the PACF of an MA(q) model behaves like the ACF of an AR(q) model.

## Theorem 4

*For any stationary process with given ACF $\{\rho_i\}$, the PACF $\{\phi_{kk}\}$ could be solved by Yule-Walker equations:*

$$\rho_{j-1}\phi_{k1} + \rho_{j-2}\phi_{k2} + \rho_{j-3}\phi_{k3} + \cdots + \rho_{j-k}\phi_{kk} = \rho_j \qquad \text{for } j = 1, 2, \cdots, k. \tag{2}$$

*Explicitly,*

$$\begin{aligned}
\phi_{k1} + \quad & \rho_1\phi_{k2} + \quad \rho_2\phi_{k3} + \quad \cdots \quad + \quad \rho_{k-1}\phi_{kk} = \quad \rho_1 \\
\rho_1\phi_{k1} + \quad & \phi_{k2} + \quad \rho_1\phi_{k3} + \quad \cdots \quad + \quad \rho_{k-2}\phi_{kk} = \quad \rho_2 \\
& \vdots \\
\rho_{k-1}\phi_{k1} + \quad & \rho_{k-2}\phi_{k2} + \quad \rho_{k-3}\phi_{k3} + \quad \cdots \quad + \quad \phi_{kk} = \quad \rho_k
\end{aligned}$$

*We solve $\phi_{k1}, \phi_{k2}, \cdots, \phi_{kk}$ and discard all but $\phi_{kk}$.*

For $AR(p)$ model and $k = p$, the parameters $\phi_1, \cdots, \phi_p$ also satisfies the Yule-Walker equations. We have $\phi_{pp} = \phi_p \neq 0$ and $\phi_{kk} = 0$ for $k > p$. Thus PACF effectively display the correct order of an AR(p) process.

## 6.2.2 Sample PACF

**Definition 5**

In the Yule-Walker equations (2), we estimate the ACF $\rho$'s by the sample ACF $r$'s, and solve the resulting linear equations for $k = 1, 2, 3, \ldots$ to get estimates of $\phi_{kk}$. The estimated function is called the **sample partial autocorrelation function (sample PACF)**, denoted by $\hat{\phi}_{kk}$.

The Yule-Walker equations (2) can be solved recursively:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j} \qquad (6.2.9)$$

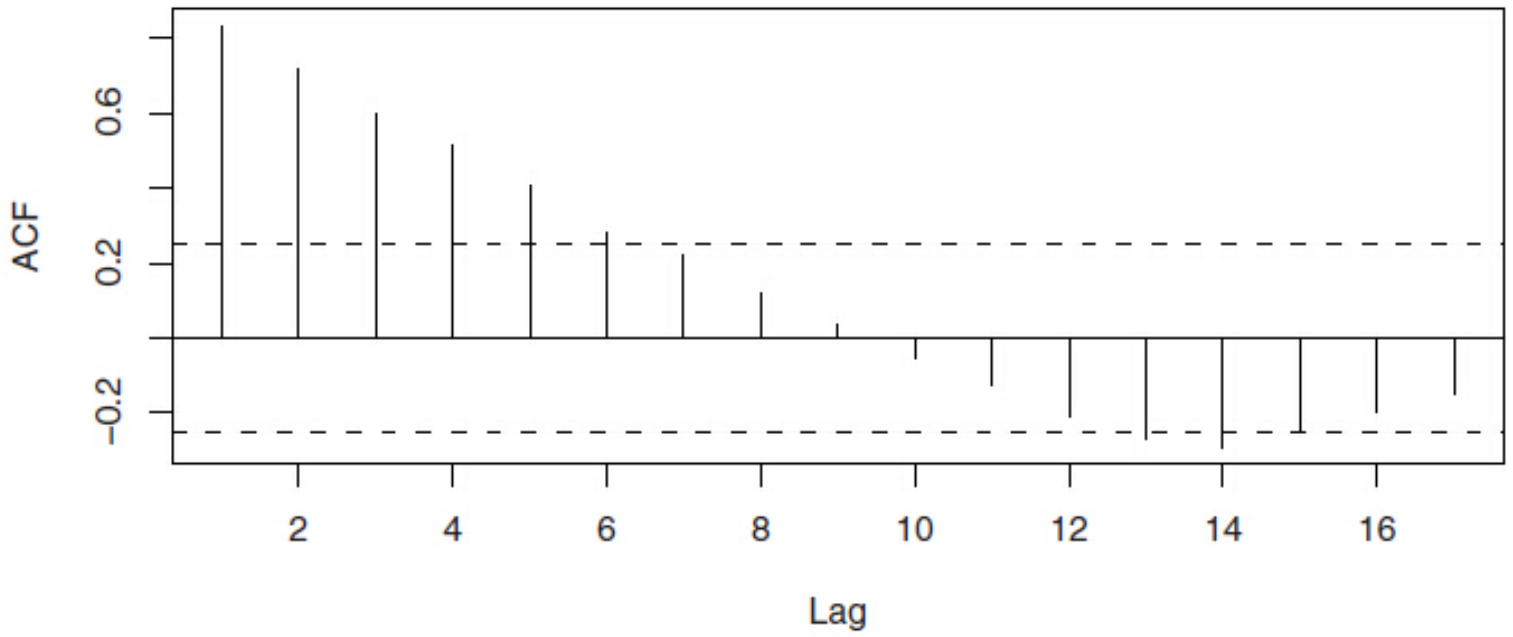
where

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j} \quad \text{for } j = 1, 2, \ldots, k-1$$

For an observed series, the sample PACF $\{\hat{\phi}_{kk}\}$ can be estimated by replacing $\rho$'s by $r$'s in the above formula.

Under the hypothesis that an AR(p) model is correct, for $k > p$, $\hat{\phi}_{kk}$ is approximately $\sim N(0, 1/n)$. Thus for $k > p$, $\pm 2/\sqrt{n}$ can be used as critical limits on $\hat{\phi}_{kk}$.

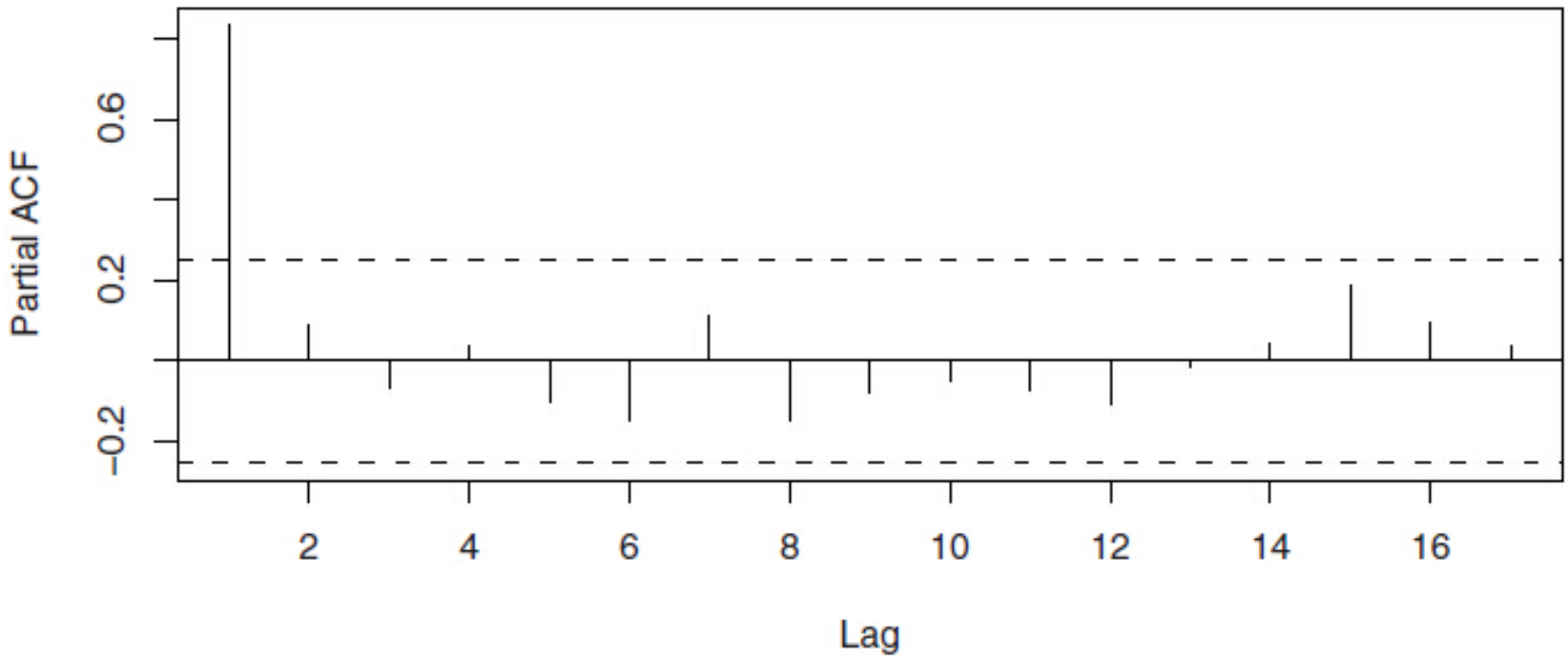
**Exhibit 6.10 Sample ACF for an AR(1) Process with $\phi = 0.9$**




```
> data(ar1.s); acf(ar1.s,xaxp=c(0,20,10))
```

The ACF shows strong positive correlations at lags 1, 2, 3, as seen early. However, the sample ACF decreases more linearly than exponentially as theory suggests. Also contrary to theory, the sample ACF goes negative at lag 10 and remains so for many lags.

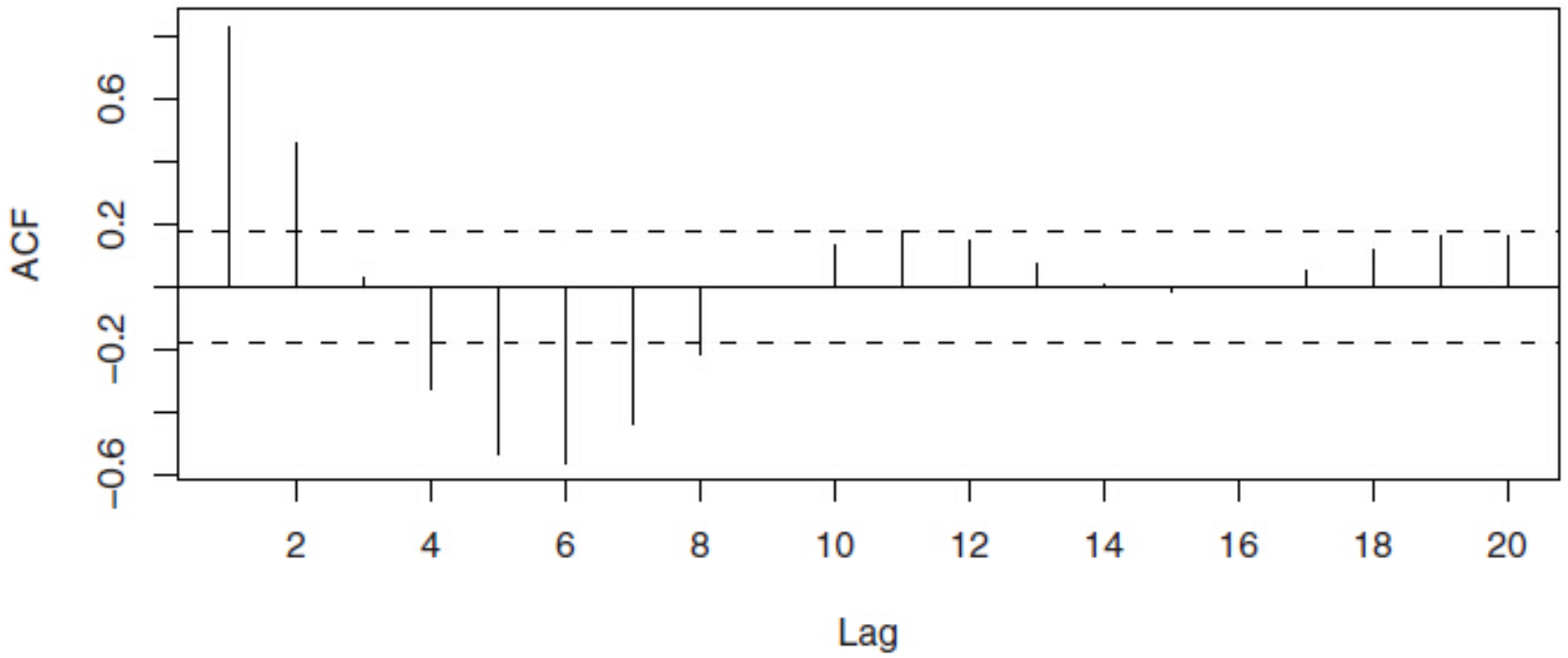
**Exhibit 6.11 Sample Partial ACF for an AR(1) Process with $\phi = 0.9$**




```
> pacf(ar1.s,xaxp=c(0,20,10))
```

The sample PACF gives a much clearer picture and strongly suggests an AR(1) model.

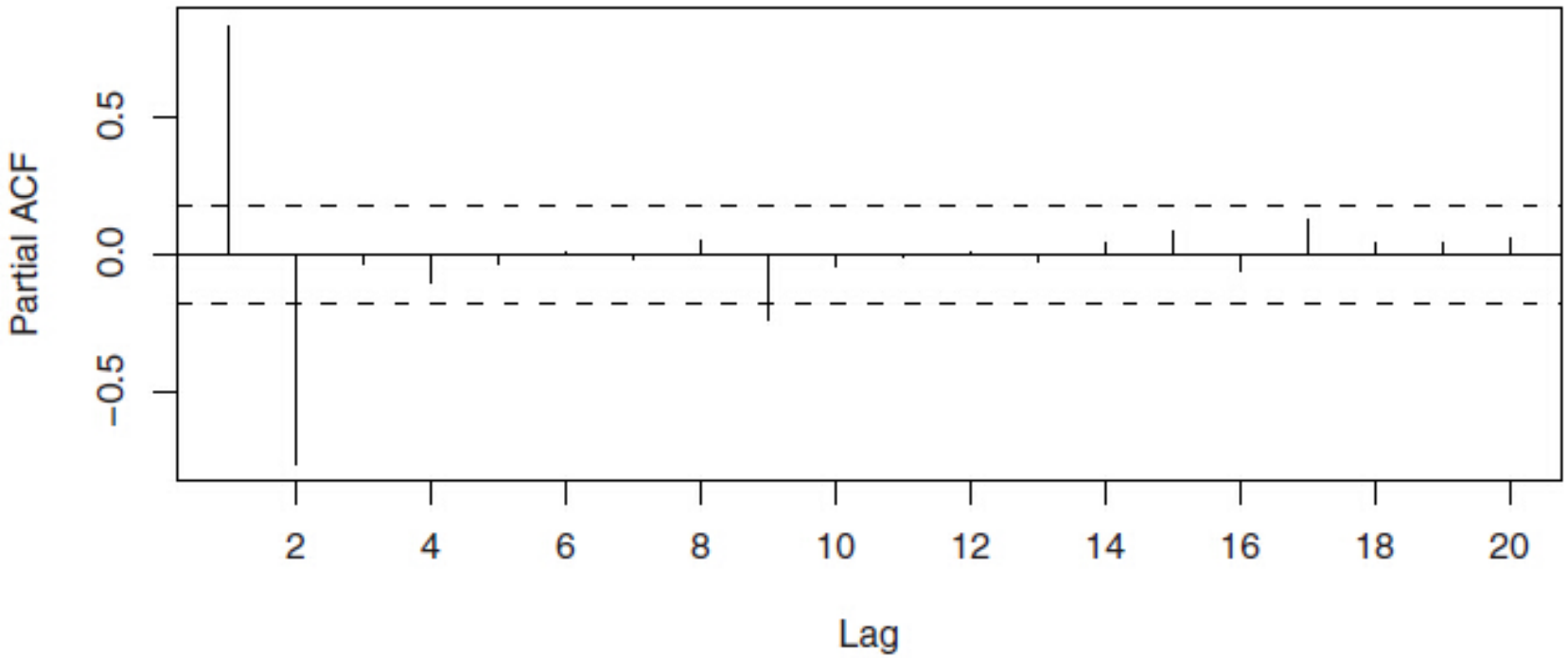
**Exhibit 6.12 Sample ACF for an AR(2) Process with $\phi_1 = 1.5$ and $\phi_2 = -0.75$**




```
> acf(ar2.s,xaxp=c(0,20,10))
```

The sample ACF looks like the damped wave. However, it does not damp down nearly as quickly as theory predicts.

**Exhibit 6.13 Sample PACF for an AR(2) Process with $\phi_1 = 1.5$ and $\phi_2 = -0.75$**




```
> pacf(ar2.s,xaxp=c(0,20,10))
```

The sample PACF gives a strong indication that we should consider an AR(2) model for these data. The sample PACF at lag 9 exceeds critical line and it would need to be investigated further during model diagnostics.

## 6.2.3 Mixed Models and the Extended ACF

The behavior of ACF and PACF are summarized in Exhibit 6.3.

**Exhibit 6.3 General Behavior of the ACF and PACF for ARMA Models**

| | AR($p$) | MA($q$) | ARMA($p,q$), $p>0$, and $q>0$ |
|---|---|---|---|
| **ACF** | Tails off | Cuts off after lag $q$ | Tails off |
| **PACF** | Cuts off after lag $p$ | Tails off | Tails off |

The sample ACF and PACF provide effective tools for identifying pure AR(p) or MA(q) models.

*For a mixed ARMA model, many graphical tools have been proposed to effectively identify the ARMA orders: the corner method (Becuin et al., 1980), the extended autocorrelation (EACF) method (Tsay and Tiao, 1984), and the smallest canonical correlation (SCAN) method (Tsay and Tiao, 1985), among others. The EACF method seems to have good sampling properties for moderately large sample sizes (W. S. Chan, 1999).*

The EACF method uses the fact that if the AR part of a mixed ARMA model is known, "filtering out" the autoregression from the observed time series results in a pure MA process that enjoys the cutoff property in its ACF. The AR coefficients may be estimated by a finite sequence of regressions.

**Ex.** Consider the ARMA(1,1) model: $Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1}$. Recall

$$\rho_1 = \frac{(\phi - \theta)(1 - \phi\theta)}{1 - 2\phi\theta + \theta^2}, \qquad \rho_2 = \phi\rho_1.$$

1. Do a linear regression of $Y_t$ on $Y_{t-1}$ gets $\rho_1 Y_{t-1}$. The residual

$$R_t^{(1)} := Y_t - \rho_1 Y_{t-1} = (\phi - \rho_1) Y_{t-1} + e_t - \theta e_{t-1}$$

is not a MA(1) process.

2. Do a second multiple regression of $Y_t$ on $Y_{t-1}$ and on $R_{t-1}^{(1)} = Y_{t-1} - \rho_1 Y_{t-2}$, the lag 1 of the residuals from the first regression: $Y_t = \widetilde{\phi} Y_{t-1} + \widetilde{\mu}(Y_{t-1} - \rho_1 Y_{t-2})$. By least squares property, the correlations of $Y_t - \widetilde{\phi} Y_{t-1} - \widetilde{\mu}(Y_{t-1} - \rho_1 Y_{t-2})$ with $Y_{t-1}$ and $Y_{t-1} - \rho_1 Y_{t-2}$ (equivalently, with $Y_{t-2}$) should be zero. We solve that

$$\widetilde{\phi} = \phi, \qquad \widetilde{\mu} = \frac{\rho_1 - \phi}{1 - \rho_1^2}.$$

The coefficient $\widetilde{\phi}$ of $Y_{t-1}$ in the second regression is a consistent estimator of $\phi$.

3. The process $W_t = Y_t - \widetilde{\phi} Y_{t-1}$ is approximately an MA(1) process.

Similarly, the AR coefficients of an ARMA(p,q) model can be consistently estimated via a sequence of q regressions.

The AR and MA orders $p$ and $q$ can be determined by an iterative procedure. Let

$$W_{t,k,j} = Y_t - \widetilde{\phi}_1 Y_{t-1} - \cdots - \widetilde{\phi}_k Y_{t-k}$$

be the autoregressive residuals defined with the AR coefficients estimated iteratively assuming the AR order is k and the MA order is j. The sample autocorrelations of $W_{t,k,j}$ for $t$ are called **the sample EACFs**. For $k = p$ and $j \geq q$, $\{W_{t,k,j}\}_t \approx MA(q)$, so that its autocorrelations of lag $q+1$ or higher are equal to zero. For $k > p$, an overfitting problem occurs, and this increases the MA order for the $W$ process by the minimum of $k - p$ and $j - q$.

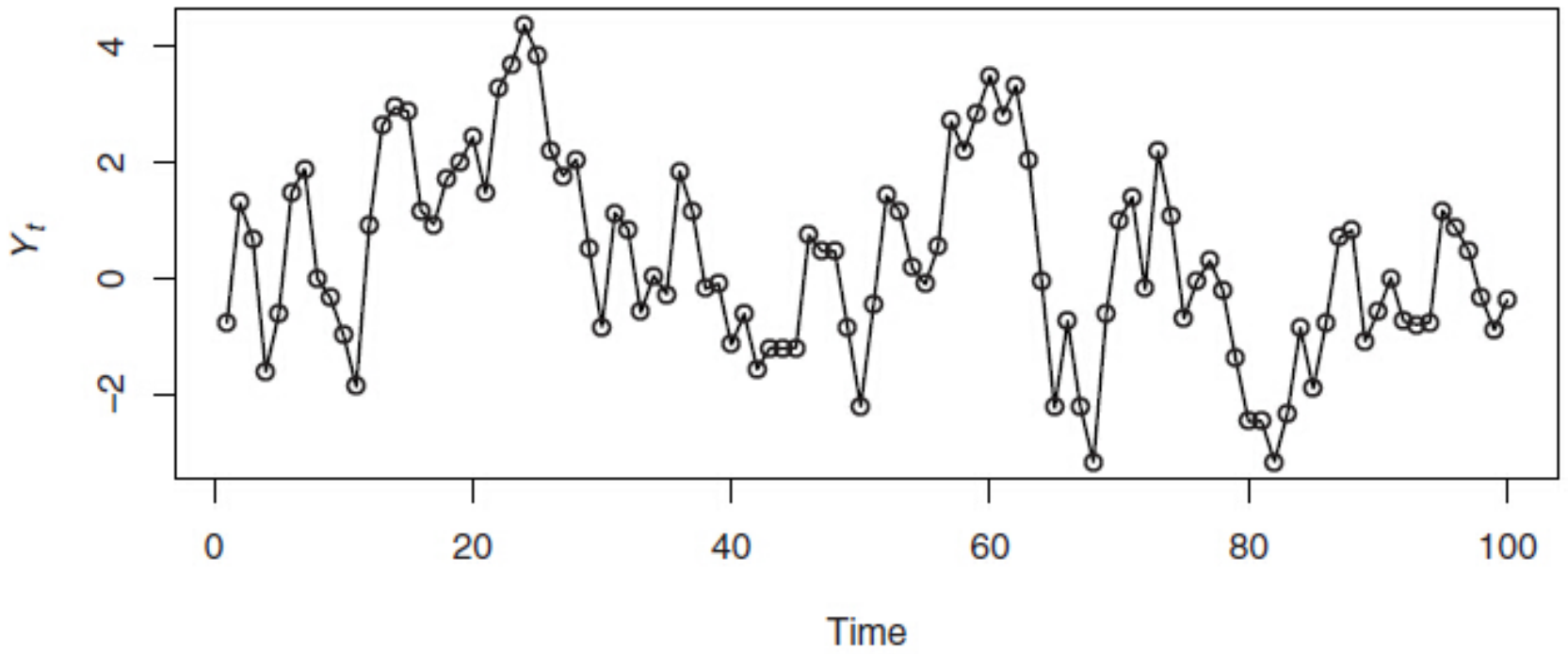
We can summarize the information in a table to determine $p$ and $q$.

**Exhibit 6.4 Theoretical Extended ACF (EACF) for an ARMA(1,1) Model**

| *AR/MA* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| 1 | x | 0* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | x | x | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | x | x | x | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | x | x | x | x | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | x | x | x | x | x | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | x | x | x | x | x | x | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | x | x | x | x | x | x | x | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

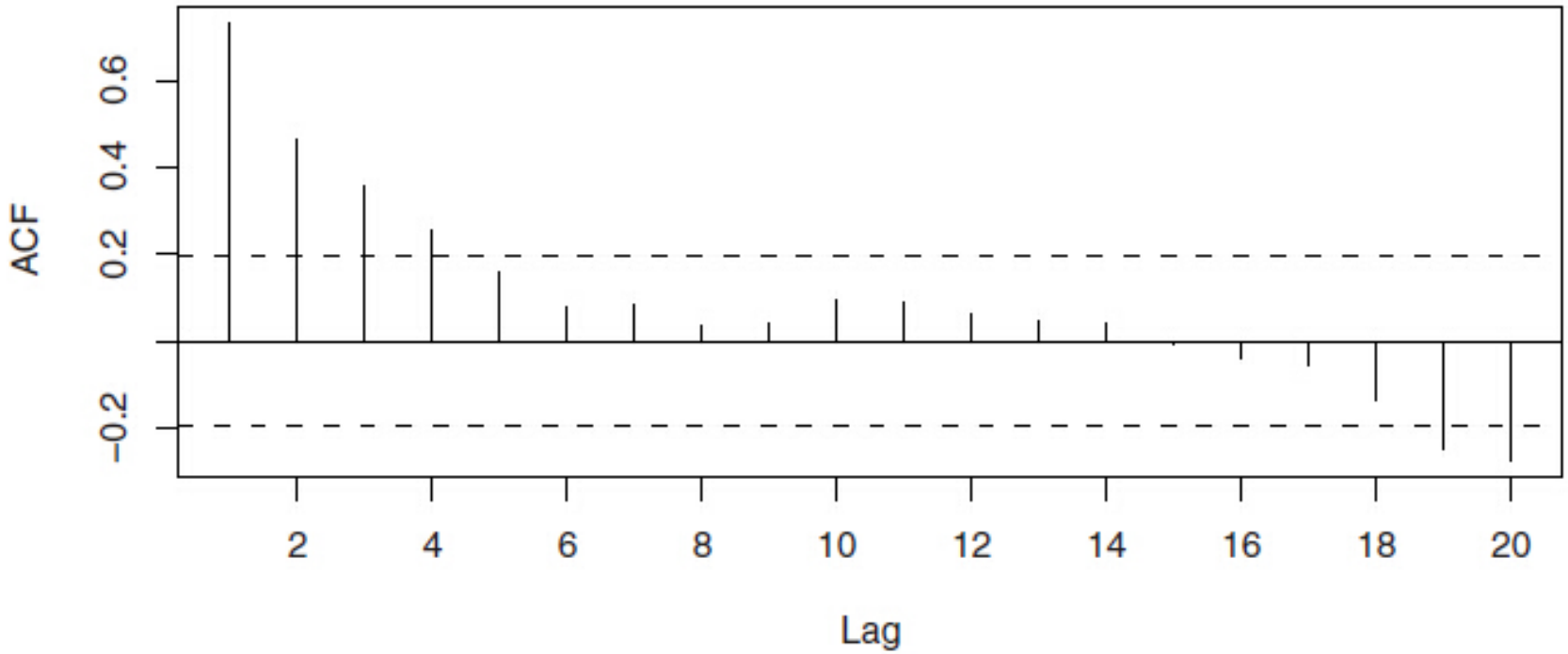
We simulate an series of mixed ARMA(1,1) model and analyze the data.

**Exhibit 6.14 Simulated ARMA(1,1) Series with $\phi = 0.6$ and $\theta = -0.3$.**



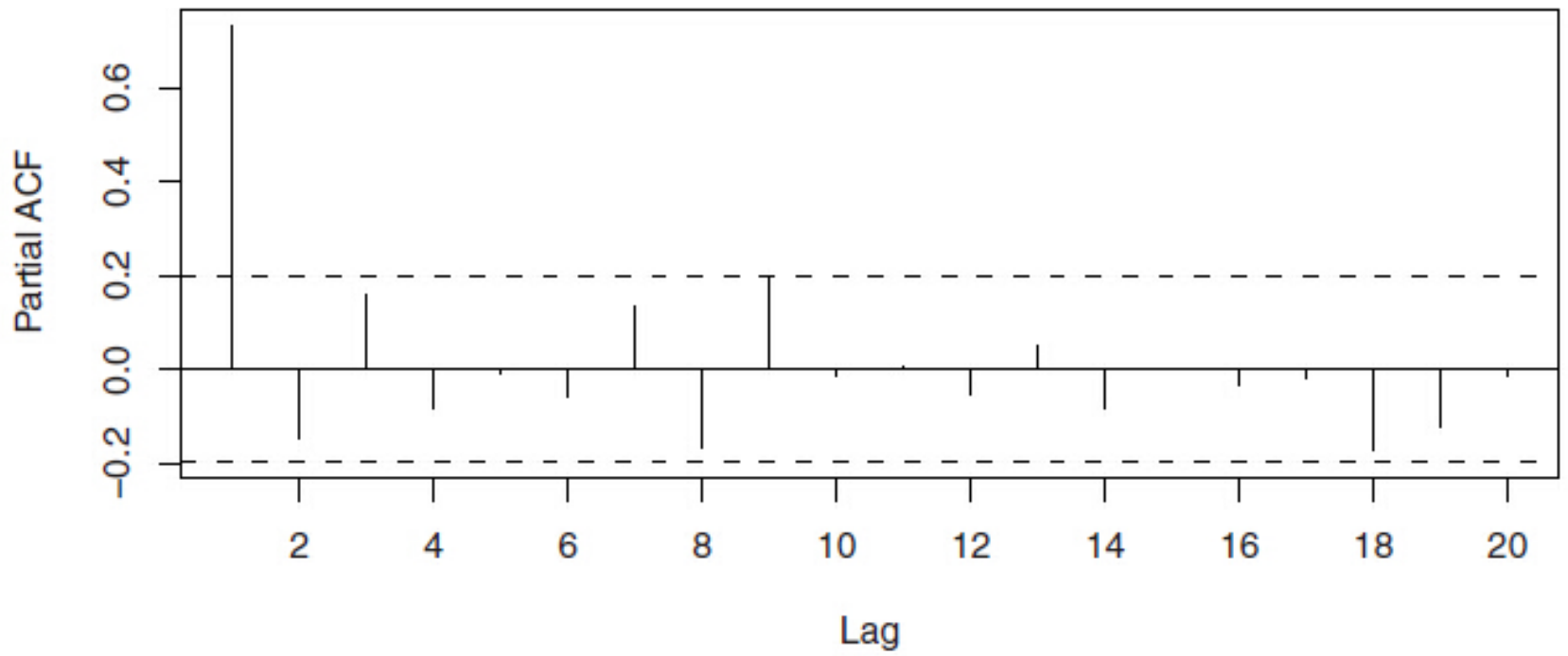

```
> data(arma11.s)
> plot(arma11.s, type='o',ylab=expression(Y[t]))
```

## Exhibit 6.15 Sample ACF for Simulated ARMA(1,1) Series




```
> acf(arma11.s,xaxp=c(0,20,10))
```

**Exhibit 6.16 Sample PACF for Simulated ARMA(1,1) Series**




```
> pacf(arma11.s,xaxp=c(0,20,10))
```

The sample ACF and the sample PACF seem to indicate that an AR(1) model should be specified. However, the triangular region of zeros shown in the sample EACF in Exhibit 6.17 indicates quite clearly that a mixed model with $q = 1$ and with $p = 1 or 2$ would be more appropriate.

**Exhibit 6.17 Sample EACF for Simulated ARMA(1,1) Series**

| AR / MA | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | x | x | x | x | o | o | o | o | o | o | o | o | o | o |
| 1 | x | o | o | o | o | o | o | o | o | o | o | o | o | o |
| 2 | x | o | o | o | o | o | o | o | o | o | o | o | o | o |
| 3 | x | x | o | o | o | o | o | o | o | o | o | o | o | o |
| 4 | x | o | x | o | o | o | o | o | o | o | o | o | o | o |
| 5 | x | o | o | o | o | o | o | o | o | o | o | o | o | o |
| 6 | x | o | o | o | x | o | o | o | o | o | o | o | o | o |
| 7 | x | o | o | o | x | o | o | o | o | o | o | o | o | o |

```
> eacf(arma11.s)
```

You may simulate and analyze a ARMA(1,1) series by yourself. (See TS-ch6.R)

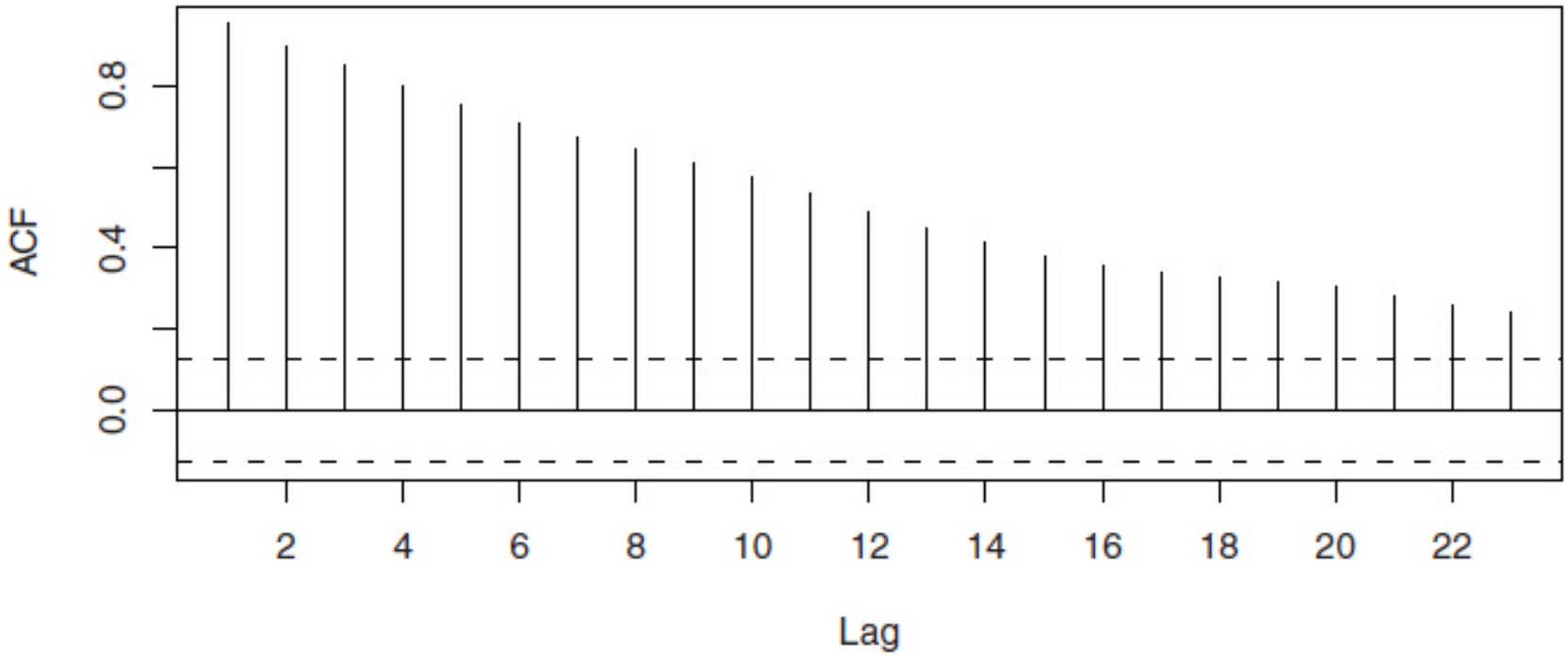
## 6.3 Specification of Some Simulated Time Series

The exhibits have been illustrated in 6.1 and 6.2.

## 6.4 Nonstationarity

The definition of sample ACF implicitly assumes stationarity. However, for nonstationary series, the sample ACF often fails to die out rapidly as the lags increase, due to the tendency for nonstationary series to drift slowly.

## Exhibit 6.18 Sample ACF for the Oil Price Time Series

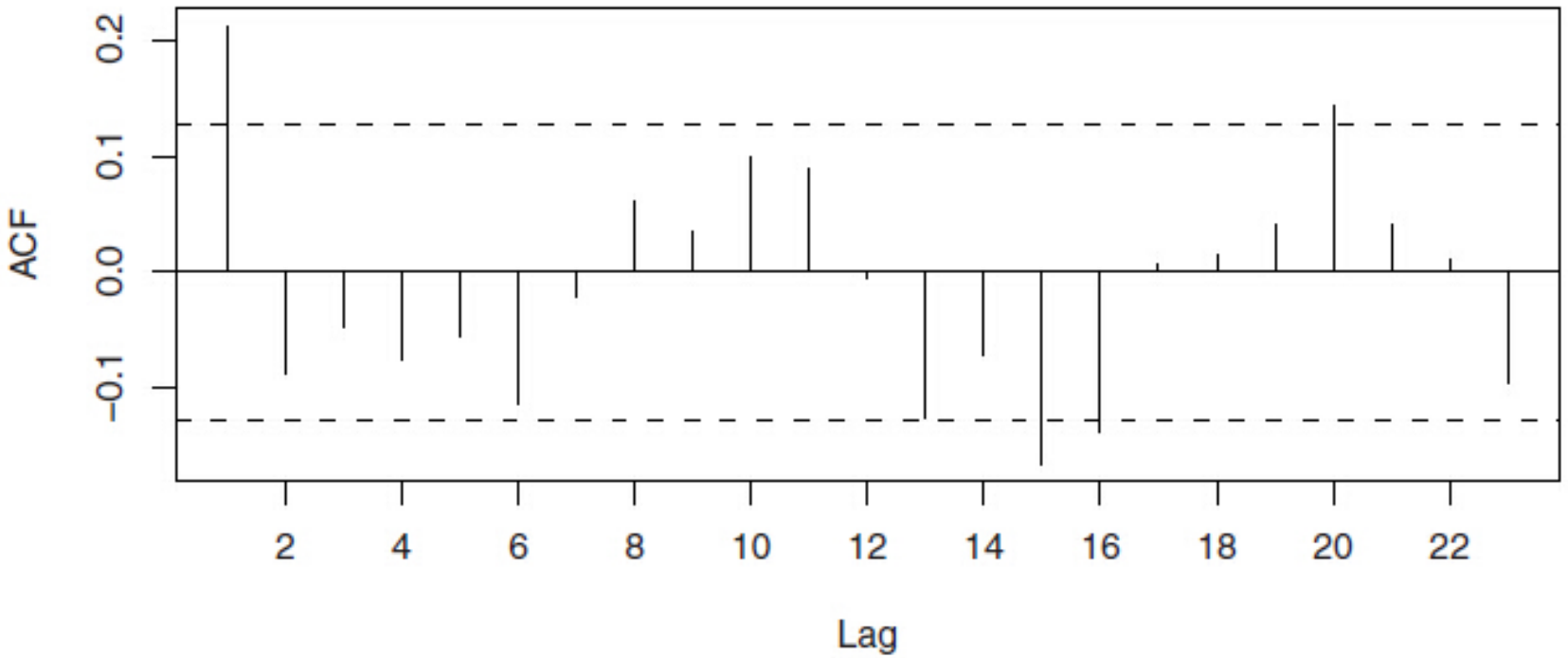



```
> data(oil.price)
> acf(as.vector(oil.price),xaxp=c(0,24,12))
```

The sample ACF values are "significantly far from zero" and they decrease slowly. The sample ACF is unlike that of a stationary one. Similar pattern happens for the series log(oil.price).

The PACF of oil.price seems strongly inclined to an AR(1) model. However, the plot of oil.price has shown that it is nonstationary.

## Exhibit 6.19 Sample ACF for the Difference of the Log Oil Price Series




```
> acf(diff(as.vector(log(oil.price))),xaxp=c(0,24,12))
```

The plot of diff(log(oil.price)) supports stationarity with possible outliners. The ACF looks like a damping sine wave with significant lag 1 value. Therefore, an MA(1) model for diff(log(oil.price)) seems appropriate. The EACF also supports an MA(1) model. Thus we may assign $diff(log(oil.price)) \sim MA(1), \quad log(oil.price) \sim IMA(1,1)$.

In the difference approach, we difference the observed series until the plot of resulting series and its sample ACF, PACF, or EACF supports a stationary process. Usually we will achieve stationarity after one or two differences, together with logarithm or some other transforms.

### 6.4.1 Overdifferencing

The difference of a stationary series is also stationary. However, overdifferencing introduces unnecessary correlations into a series.

**Ex.** Suppose an observed series

$$\{Y_t\} \sim ARIMA(p, d, q).$$

Then

$$\nabla^d Y_t \sim ARMA(p, q).$$

If we overdifference the series and get

$$\nabla^{d+1} Y_t \sim ARMA(p, q+1)$$

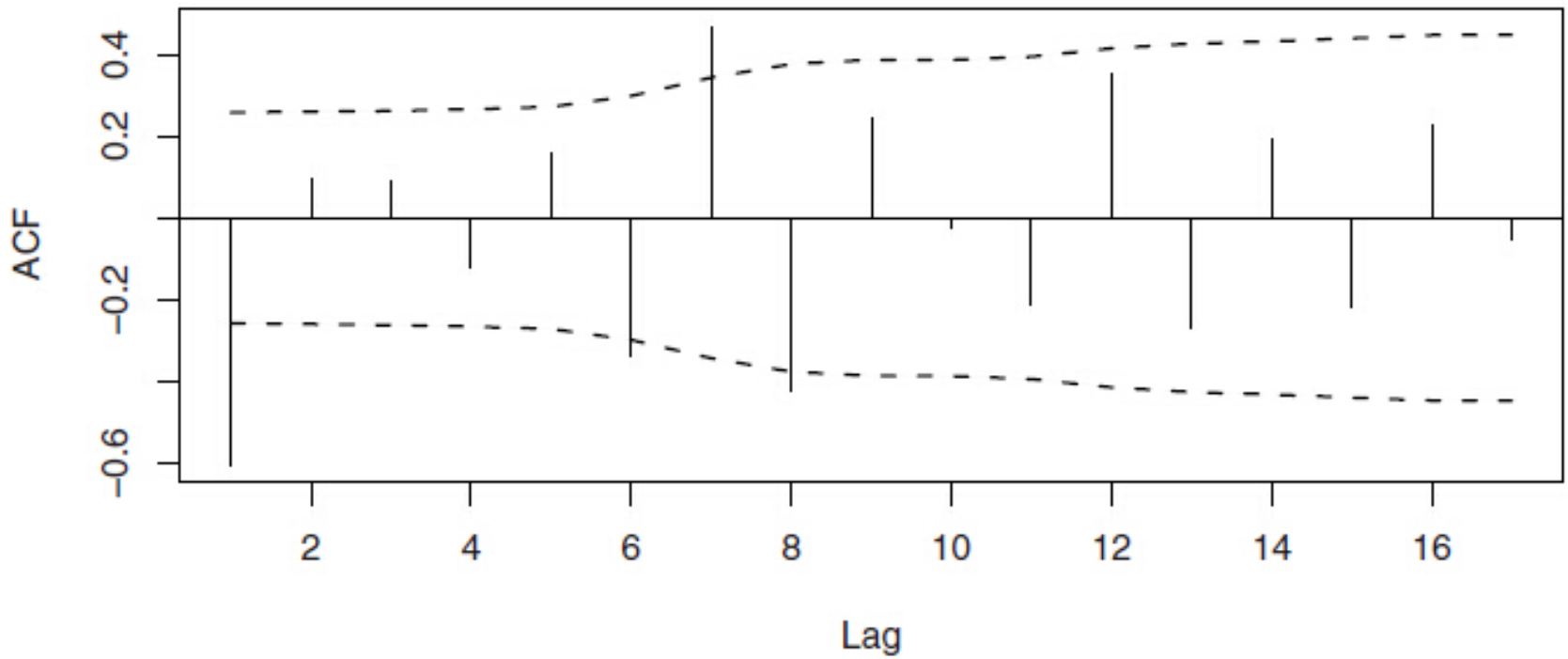
then we will mistakenly model the original series as

$$\{Y_t\} \sim ARIMA(p, d+1, q+1).$$

Drawbacks:

1. Overdifferecing makes the model more complicated.
2. Overdifferecing creates a noninvertible model, which will create serious problems when we attempt to estimate their parameters.

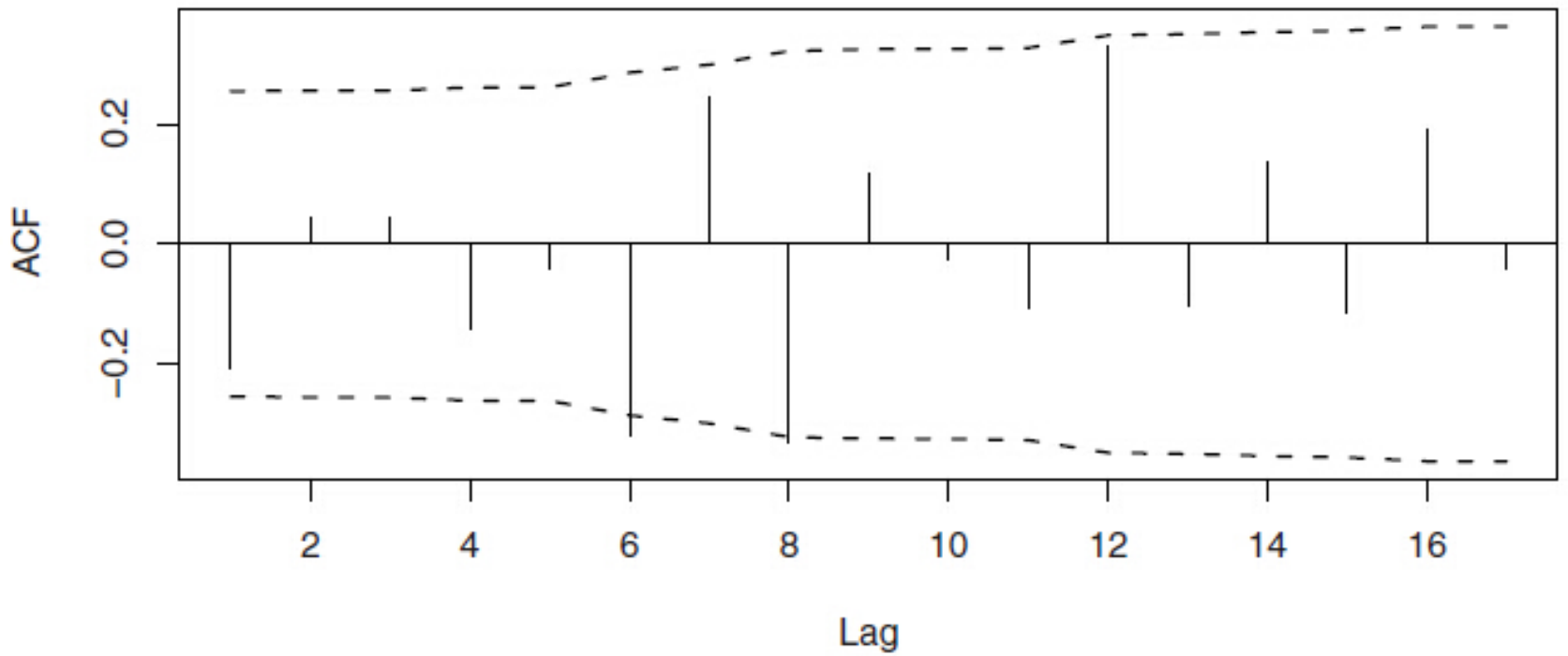
## Exhibit 6.20 Sample ACF of Overdifferenced Random Walk




```
> data(rwalk)
> acf(diff(rwalk,difference=2),ci.type='ma', xaxp=c(0,18,9))
```

If we overdifference the dataset `rwalk`, then we may model $\nabla^2(rwalk) \sim MA(1)$ and mistakenly model $rwalk \sim IMA(2,1)$.

**Exhibit 6.21 Sample ACF of Correctly Differenced Random Walk**




```
> acf(diff(rwalk),ci.type='ma',xaxp=c(0,18,9))
```

The right difference suggests that $\nabla(rwalk) \sim e_t$ and thus $rwalk \sim ARIMA(0, 1, 0)$ (a random walk process).

Remember the principle of parsimony:

**models should be simple, but not too simple.**

## 6.5 Other Specification Methods

Some popular approaches to determine the ARMA orders $p$ and $q$:

1. **Akaike's Information Criterion (AIC):** to select the model that minimizes

$$\text{AIC} = -2\log(\text{maximal likelihood}) + 2k$$

where $k = p + q + 1$ if the model contains an intercept or constant term and $k = p + q$ otherwise. The AIC is an estimator of the average Kullback-Leibler divergence of the estimated model from the true model.

2. **The Schwarz Bayesian Information Criterion (BIC):** to select the model that minimizes

$$\text{BIC} = -2\log(\text{maximal likelihood}) + k\log(n).$$

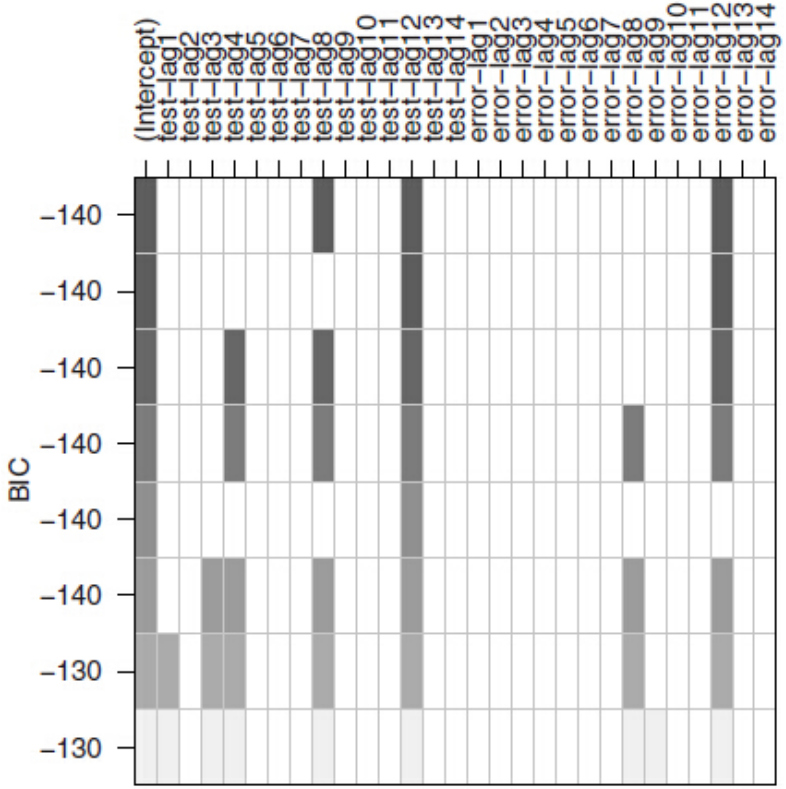
A **subset ARMA(p,q) model** is an ARMA(p,q) model with a subset of its coefficients known to be zero. For a subset ARMA(12,12) model like

$$Y_t = 0.8Y_{t-12} + e_t + 0.7e_{t-12},$$

finding a subset ARMA model that adequately approximates the underlying process is more important than simply determining the ARMA orders. It is prudent to examine a few best subset ARMA models (in terms of, for example, BIC) in order to arrive at some helpful tentative models for further study.

We show an example of the above model in the next exhibit.

# Exhibit 6.22 Best Subset ARMA Selection Based on BIC




```
> set.seed(92397)
> test=arima.sim(model=list(ar=c(rep(0,11),.8),
    ma=c(rep(0,11),0.7)),n=120)
> res=armasubsets(y=test,nar=14,nma=14,y.name='test',
    ar.method='ols')
> plot(res)
```

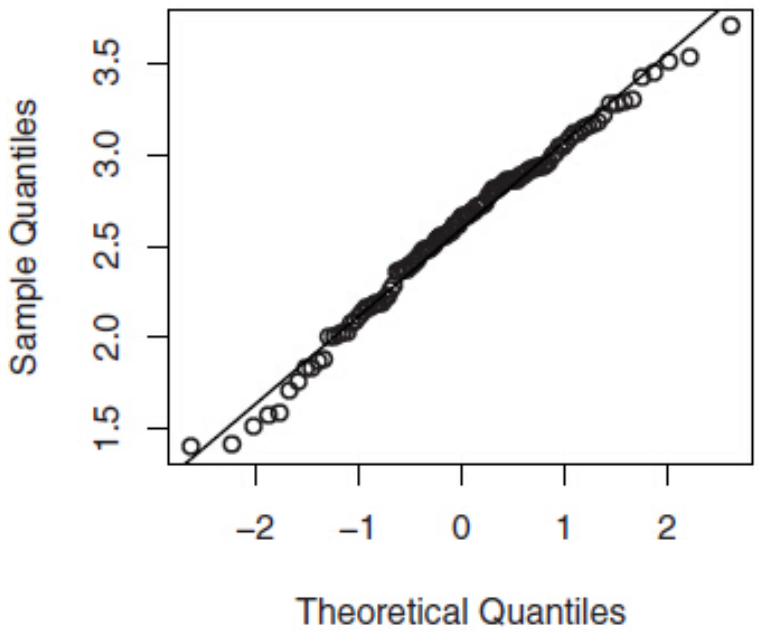
Each row in the exhibit corresponds to a subset ARMA model where the cells of the variables selected for the model are shaded. The models are sorted according to their BIC, with better models (lower BIC) placed in higher rows and with darker shades.

The best Subset ARMA selection only suggests some possible models. These models need further examinations. In the above exhibit, if we use `set.seed(3399)` or some other numbers, we may have different outcomes.

# 6.6 Specification of Some Actual Time Series

## The Los Angeles Annual Rainfall Series `larain`

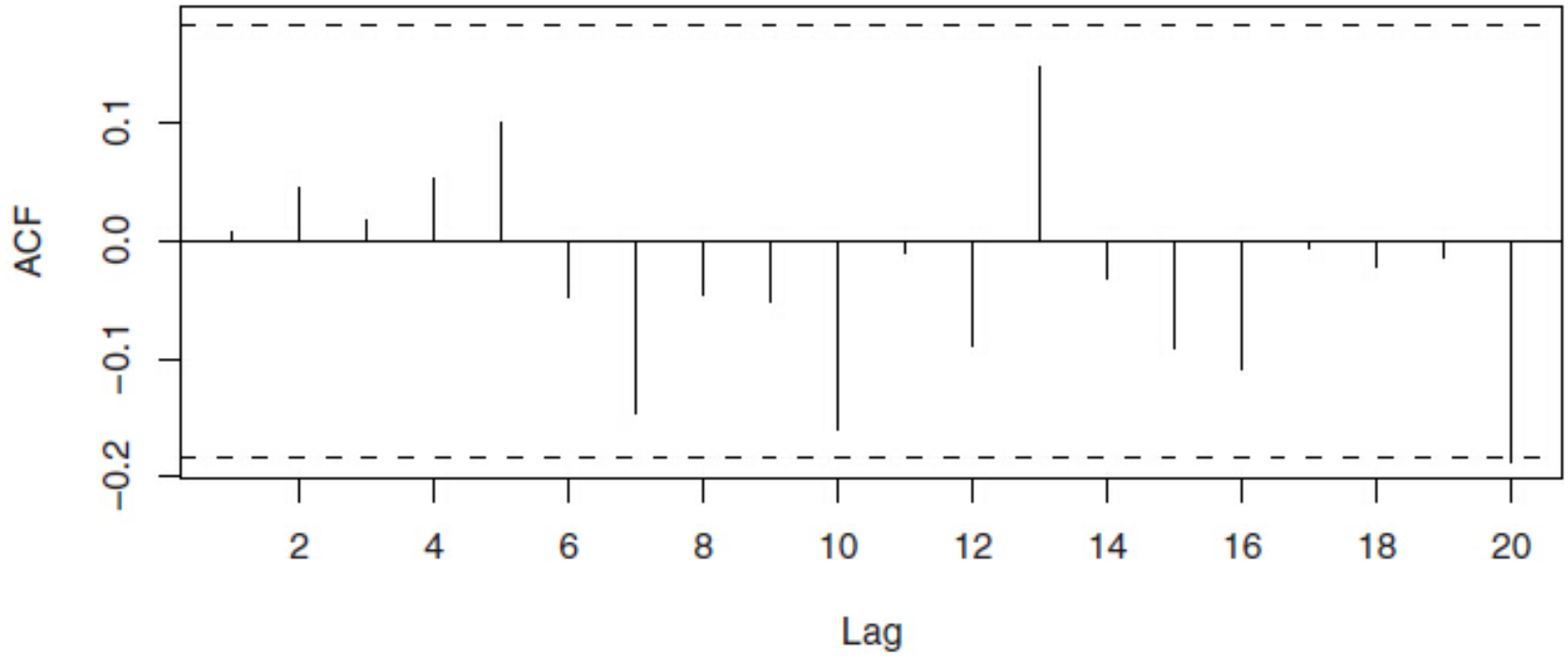
**Exhibit 6.23 QQ Normal Plot of the Logarithms of LA Annual Rainfall**




```
> data(larain); win.graph(width=2.5,height=2.5,pointsize=8)
> qqnorm(log(larain)); qqline(log(larain))
```

Taking logarithms significantly improves the normality. You may also try `shapiro.test` to check.

## Exhibit 6.24 Sample ACF of the Logarithms of LA Annual Rainfall

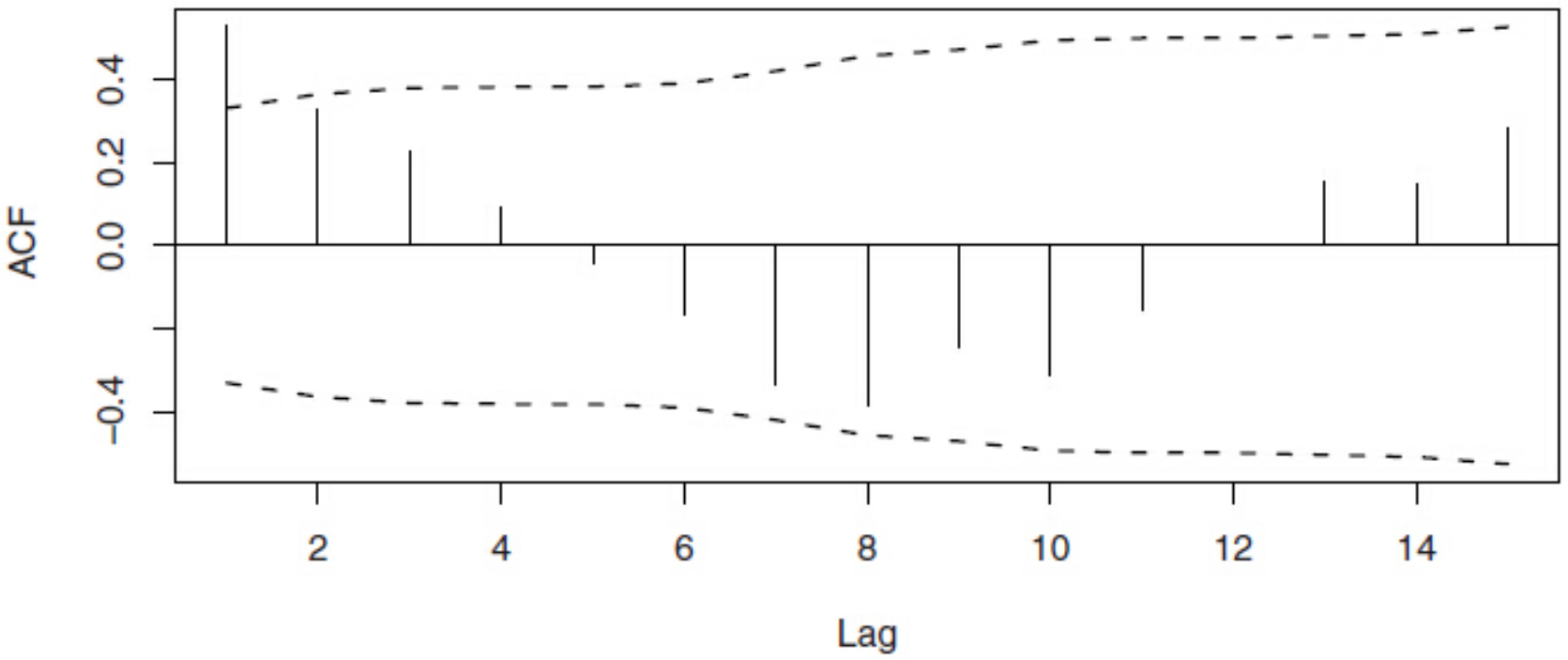



```
> win.graph(width=4.875,height=3,pointsize=8)
> acf(log(larain),xaxp=c(0,20,10))
```

Based on the sample ACF, we could model the logarithm of annual rainfall amount as independent, normal random variables with mean 2.58 and standard deviation 0.478.

# The Chemical Process Color Property Series `color`

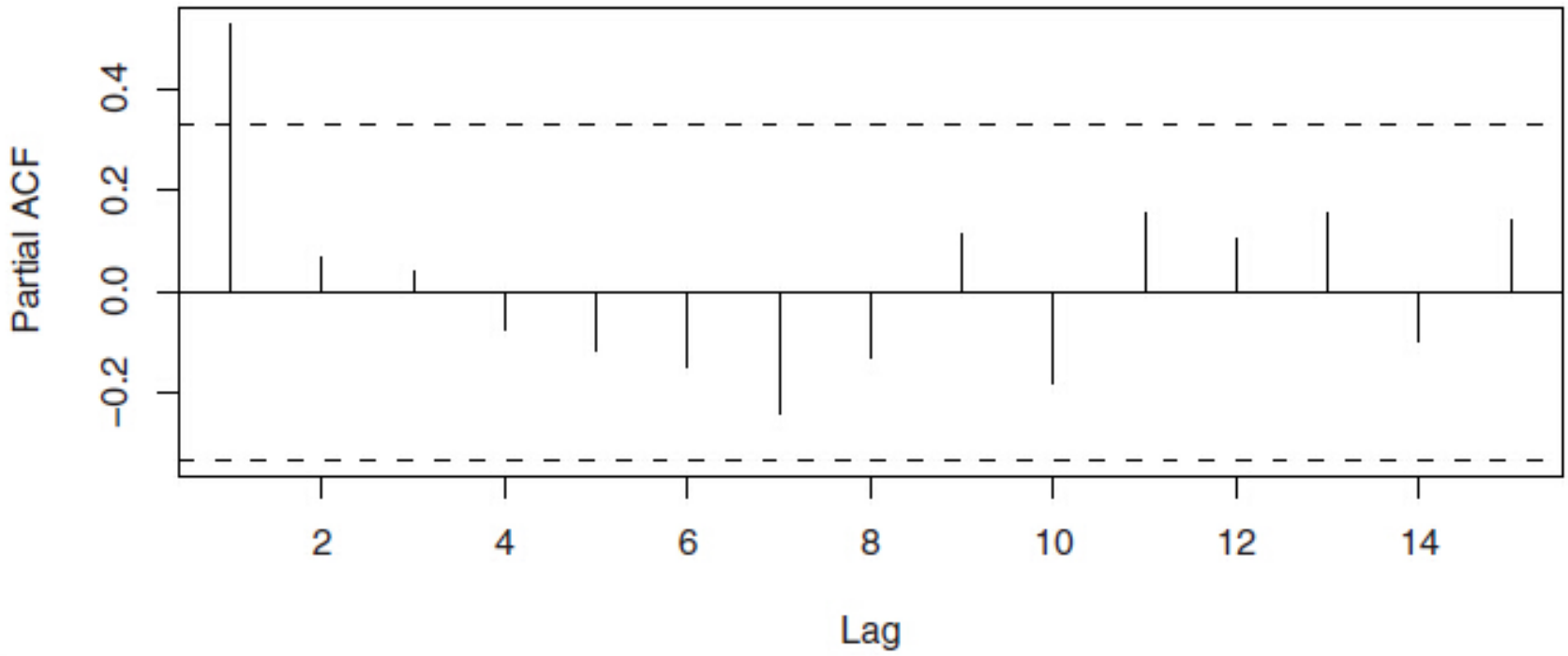
**Exhibit 6.25 Sample ACF for the Color Property Series**




```
> data(color); acf(color,ci.type='ma')
```

The ACF plot looks like an MA(1) model. However, there is a damp sine wave pattern. Maybe an AR or ARMA model is more appropriate.

## Exhibit 6.26 Sample Partial ACF for the Color Property Series

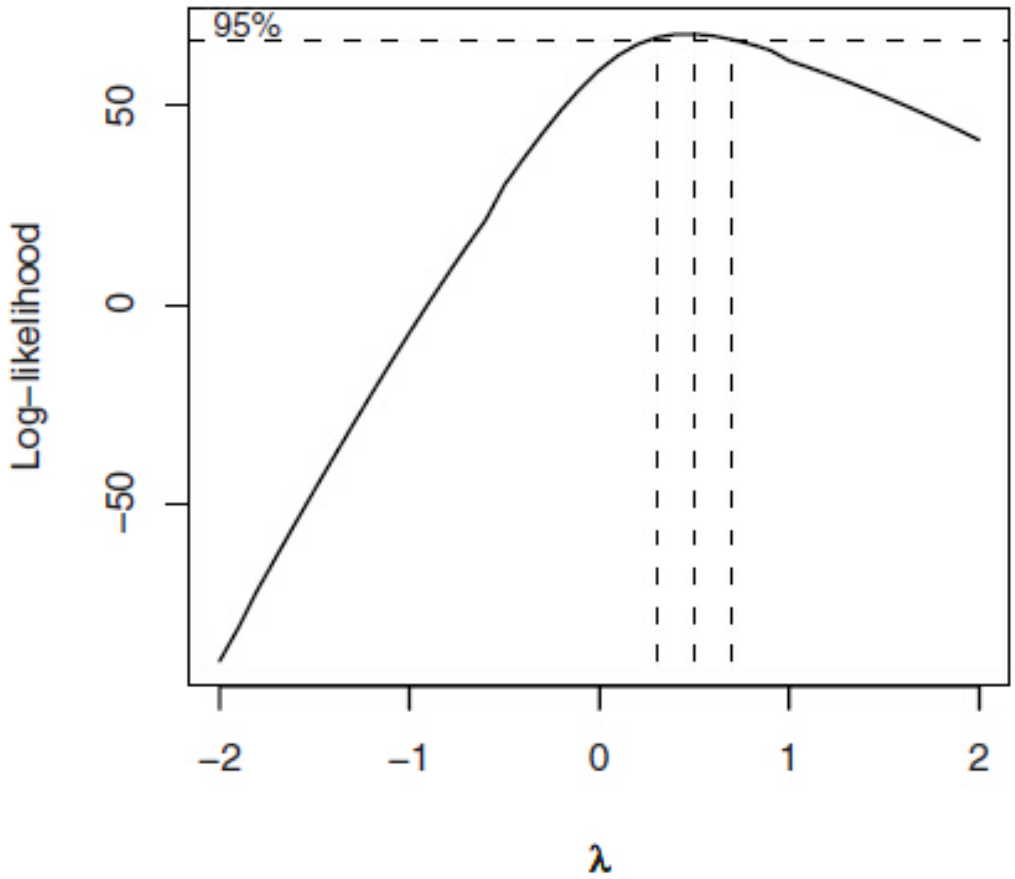



```
> pacf(color)
```

The PACF plot suggests an AR(1) model.

# The Annual Abundance of Canadian Hare Series `hare`

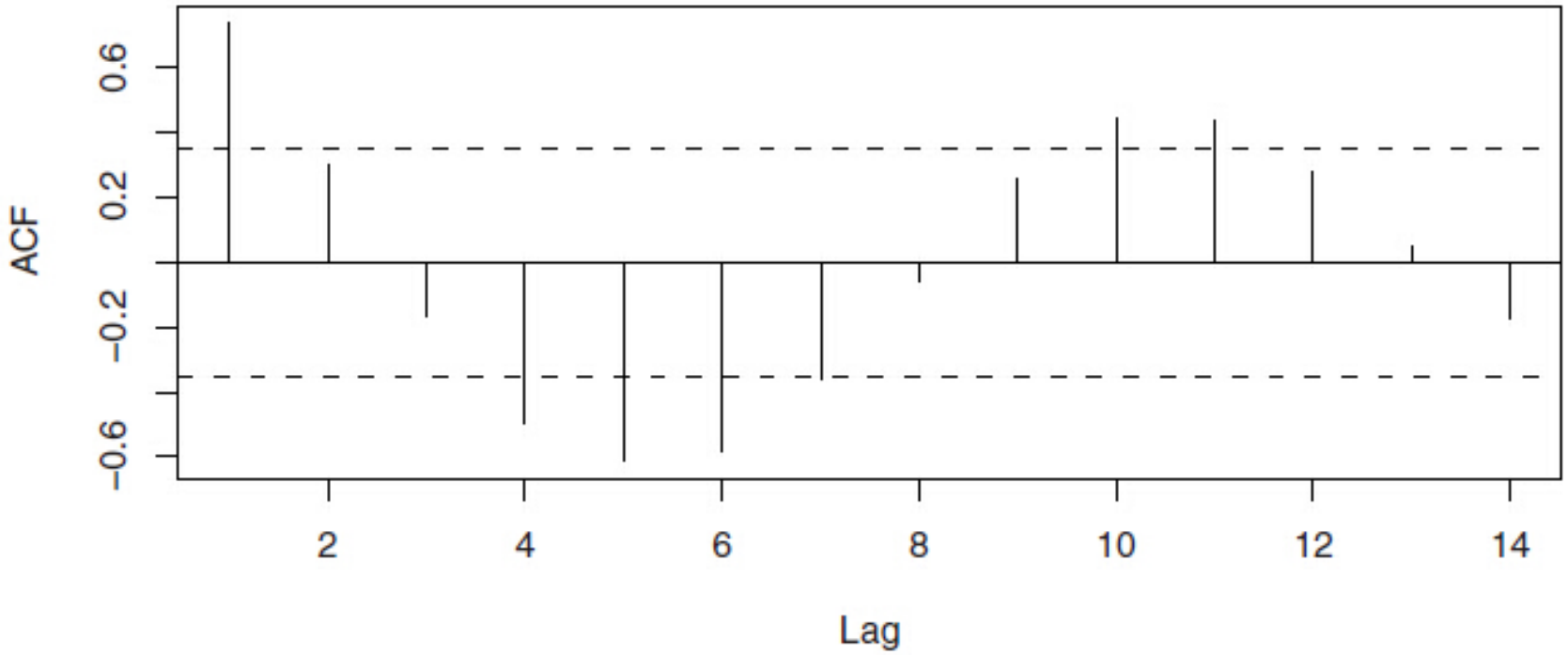
**Exhibit 6.27 Box-Cox Power Transformation Results for Hare Abundance**




```
> win.graph(width=3,height=3,pointsize=8)
> data(hare); BoxCox.ar(hare)
```

The confidence interval contains $\lambda = 0.5$. We will take square roots on `hare` for further analysis.

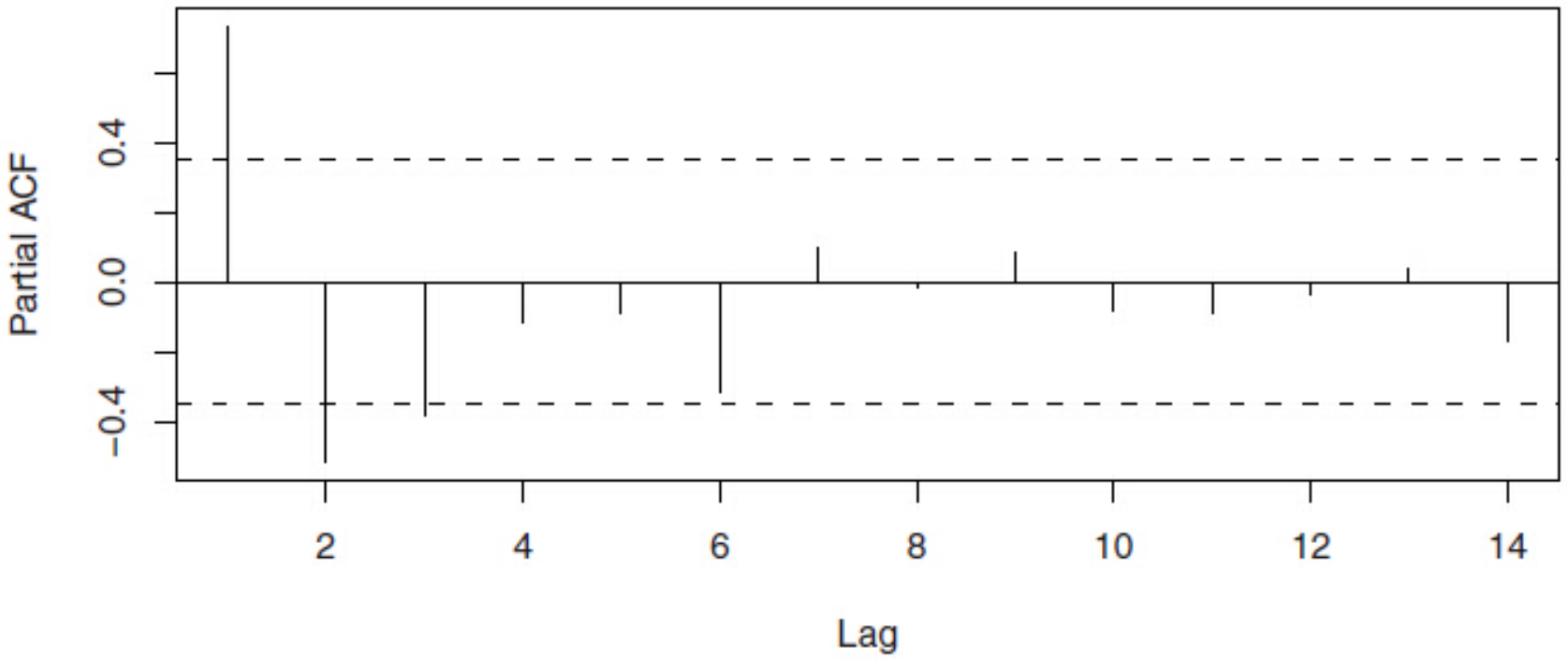
**Exhibit 6.28 Sample ACF for Square Root of Hare Abundance**




```
> acf(hare^.5)
```

The damped sine wave pattern motivate us to look for PACF.

**Exhibit 6.29 Sample Partial ACF for Square Root of Hare Abundance**




```
> pacf(hare^.5)
```

The PACF plot suggests an AR(2) or AR(3) model.

## The Oil Price Series `oil.price`

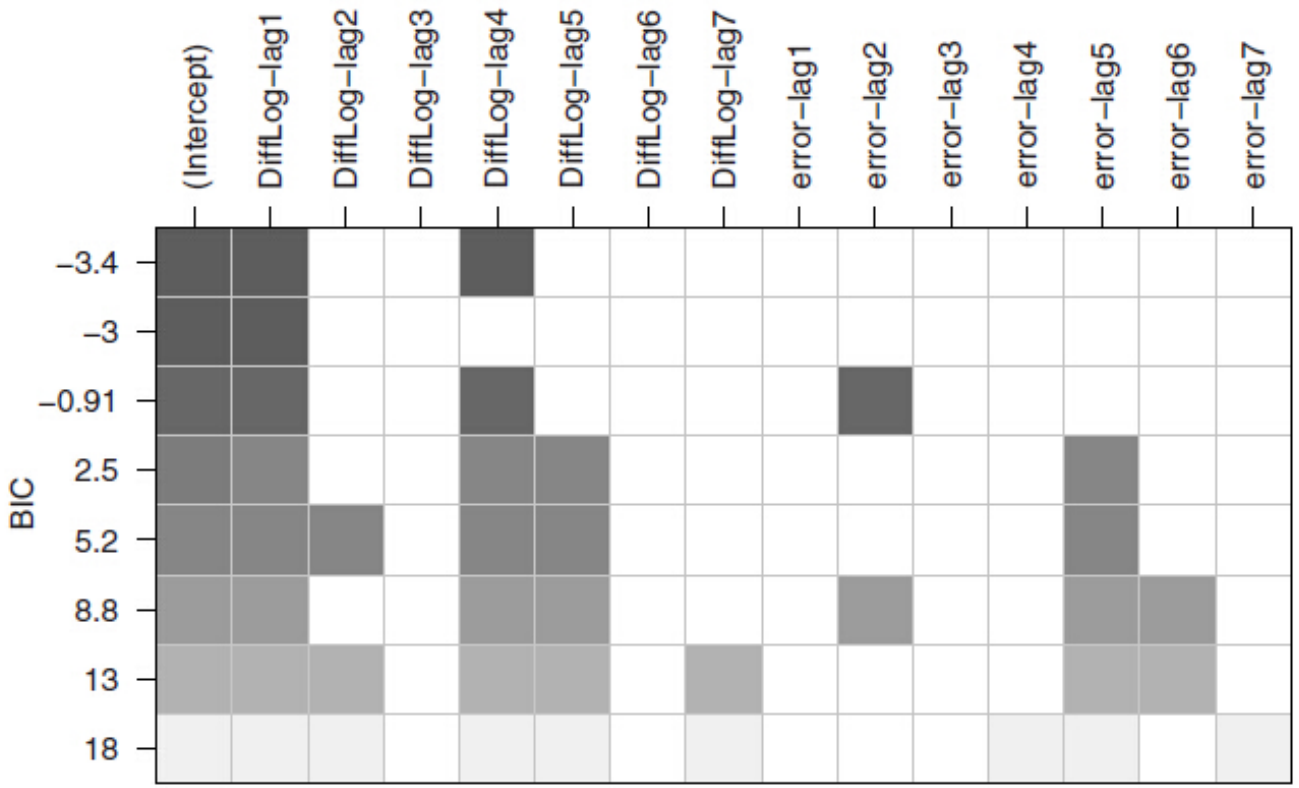
**Exhibit 6.30 Extended ACF for Difference of Logarithms of Oil Price Series**

| AR / MA | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | x | o | o | o | o | o | o | o | o | o | o | o | o | o |
| 1 | x | x | o | o | o | o | o | o | o | o | x | o | o | o |
| 2 | o | x | o | o | o | o | o | o | o | o | o | o | o | o |
| 3 | o | x | o | o | o | o | o | o | o | o | o | o | o | o |
| 4 | o | x | x | o | o | o | o | o | o | o | o | o | o | o |
| 5 | o | x | o | x | o | o | o | o | o | o | o | o | o | o |
| 6 | o | x | o | x | o | o | o | o | o | o | o | o | o | o |
| 7 | x | x | o | x | o | o | o | o | o | o | o | o | o | o |

```
> eacf(diff(log(oil.price)))
```

This table suggests an ARMA model with $p = 0$ and $q = 1$, that is, MA(1).

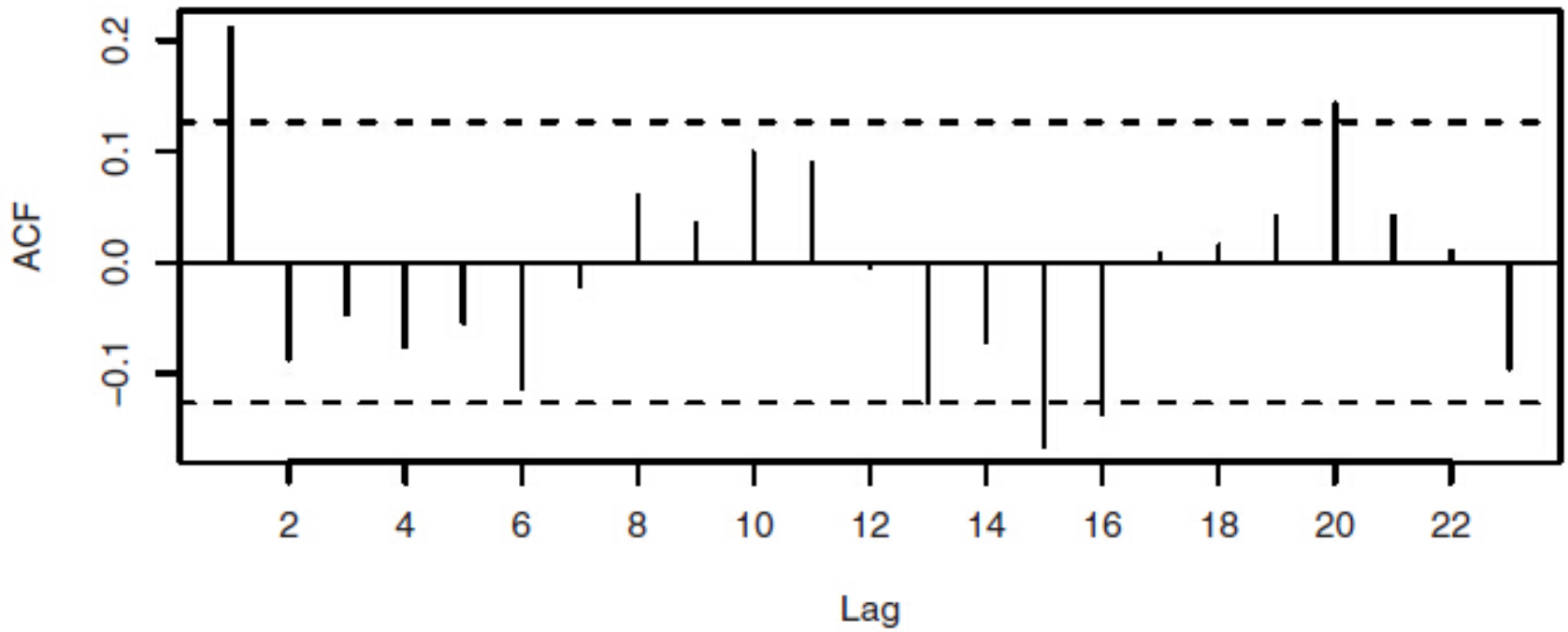
**Exhibit 6.31 Best Subset ARMA Model for Difference of Log(Oil)**




```
> res=armasubsets(y=diff(log(oil.price)),nar=7,nma=7,
    y.name='test', ar.method='ols')
> plot(res)
```

For $Y_t := \nabla \log(\text{Oil}_t)$, the best model is in terms of $Y_{t-1}$ and $Y_{t-4}$ and no lags are needed in error terms. The second best model is AR(1) (where $\log(\text{Oil}_t) \sim ARI(1,1)$).

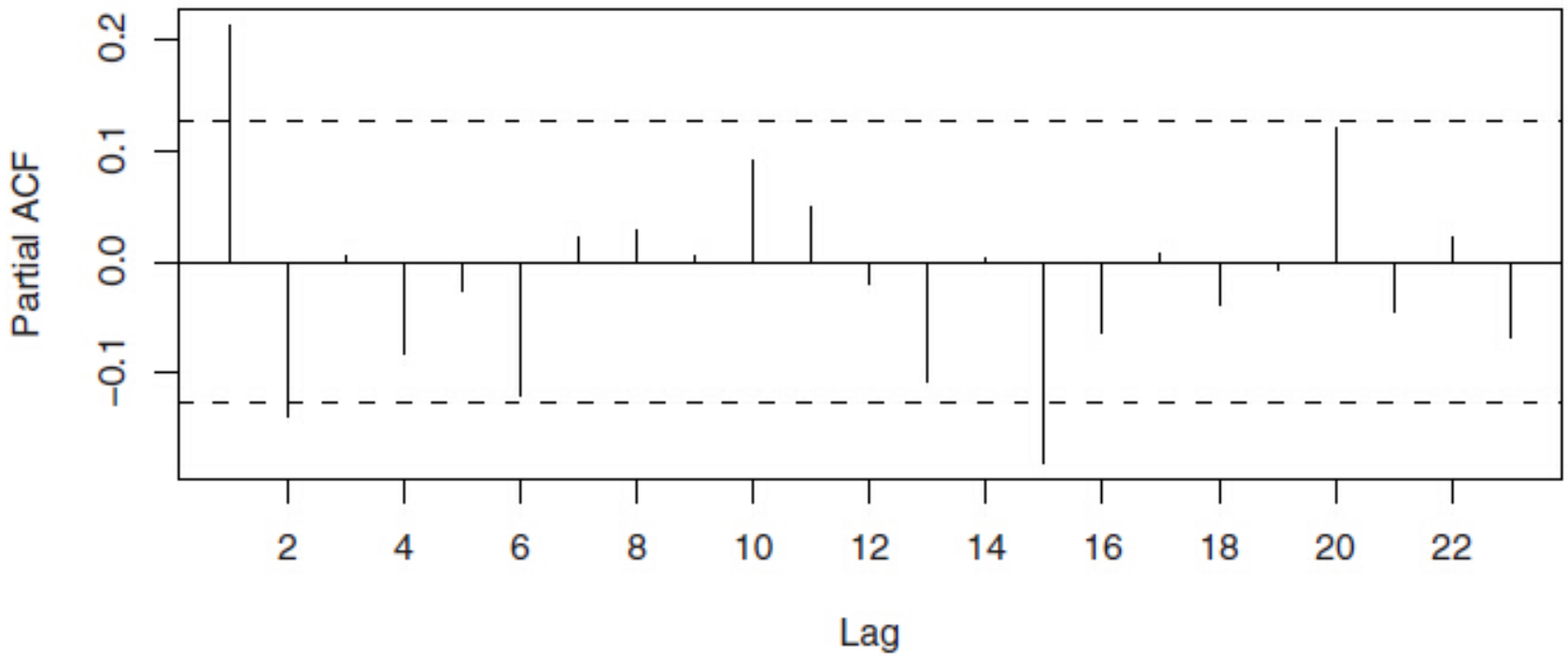
## Exhibit 6.32 Sample ACF of Difference of Logged Oil Prices




```
> acf(as.vector(diff(log(oil.price))),xaxp=c(0,22,11))
```

The plot suggests an MA(1) model for $Y_t := \nabla \log(\text{Oil}_t)$. The sample ACF seems to have some damped sine wave pattern.

## Exhibit 6.33 Sample PACF of Difference of Logged Oil Prices




```
> pacf(as.vector(diff(log(oil.price))),xaxp=c(0,22,11))
```

The plot suggests an AR(2) model for $Y_t := \nabla \log(\text{Oil}_t)$.

We will investigate all the above models further.