MODEL SPECIFICATION

YENNI ANGRAINI

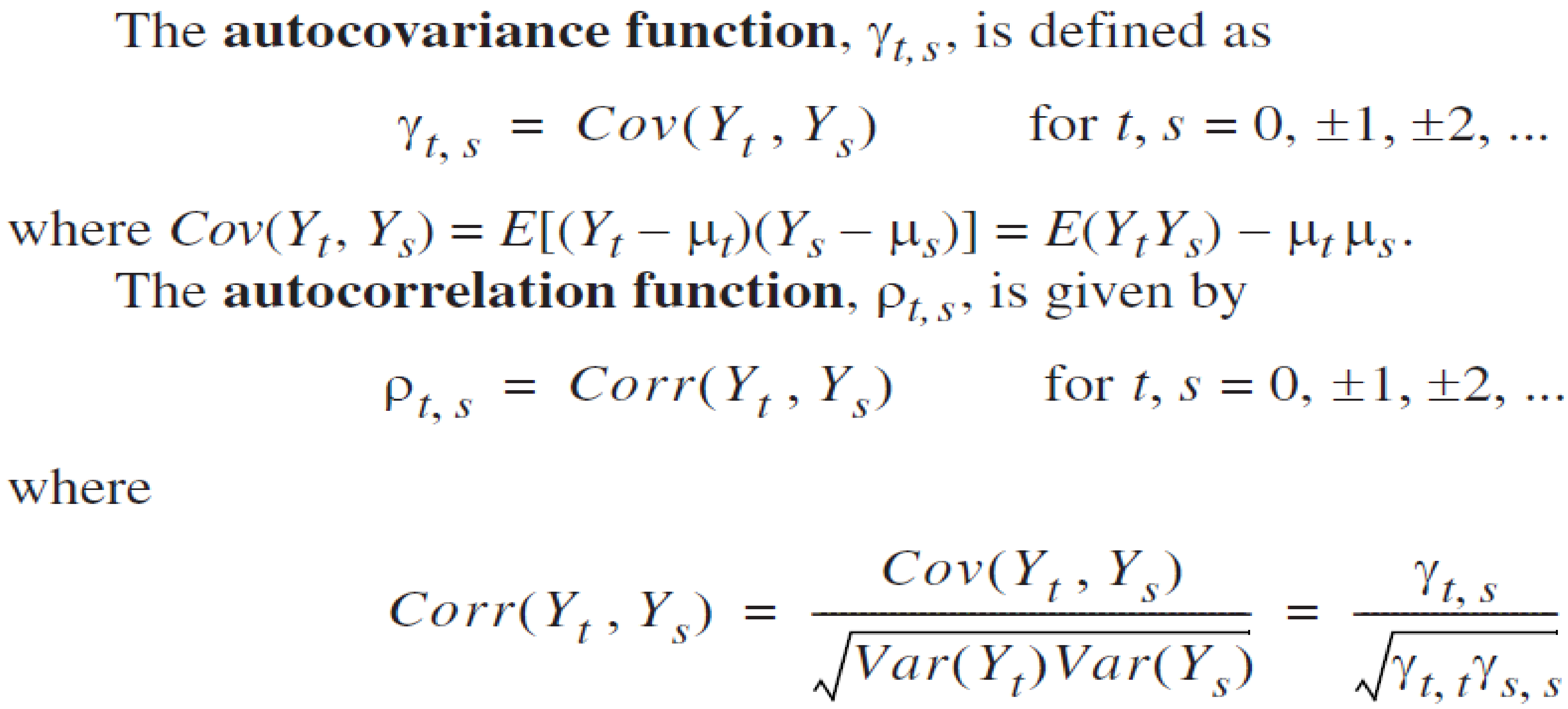
Outline

- Characteristic of ACF : AR(p), MA(q) and ARMA(p,q)
- Characteristic of PACF : AR(p), MA(q) and ARMA(p,q)
- Illustrations

The subjects of the next three meeting, respectively, are:

1. how to choose appropriate values for *p*, *d*, and *q* for a given series;
2. how to estimate the parameters of a specific ARIMA(*p*,*d*,*q*) model;
3. how to check on the appropriateness of the fitted model and improve it if needed.

The Autocorrelation Functions



The **autocovariance function**, $\gamma_{t,s}$, is defined as

$$\gamma_{t,s} = Cov(Y_t, Y_s) \qquad \text{for } t, s = 0, \pm 1, \pm 2, \ldots$$

where $Cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E(Y_t Y_s) - \mu_t \mu_s$.

The **autocorrelation function**, $\rho_{t,s}$, is given by

$$\rho_{t,s} = Corr(Y_t, Y_s) \qquad \text{for } t, s = 0, \pm 1, \pm 2, \ldots$$

where

$$Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}$$

we can simplify our notation and write

$$\gamma_k = Cov(Y_t, Y_{t-k}) \qquad \text{and} \qquad \rho_k = Corr(Y_t, Y_{t-k})$$

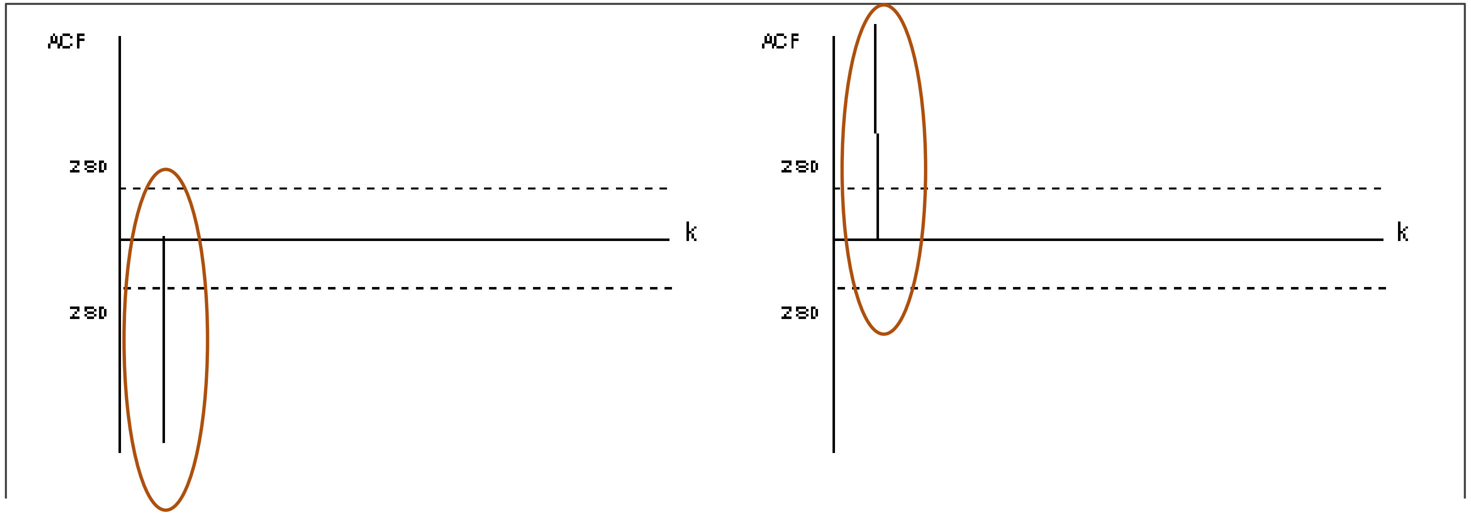
Note also that

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

$$\left.\begin{array}{ll} \gamma_0 = Var(Y_t) & \rho_0 = 1 \\ \gamma_k = \gamma_{-k} & \rho_k = \rho_{-k} \\ |\gamma_k| \leq \gamma_0 & |\rho_k| \leq 1 \end{array}\right\}$$

If a process is strictly stationary and has finite variance, then the covariance function must depend only on the time lag.

ACF for MA (1)



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

$$E(Y_t) = 0, Var(Y_t) = \gamma_0 = \sigma_e^2(1 + \theta^2)$$

$$Cov(Y_t, Y_{t-1}) = \gamma_1 = -\theta\sigma_e^2$$

$$Cov(Y_t, Y_{t-2}) = 0,\ \ Cov(Y_t, Y_{t-3}) = 0, \dots Cov(Y_t, Y_{t-k}) = 0\ for\ k = 2,3, \dots$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta}{(1+\theta^2)}, \rho_2 = 0, \dots, \rho_k = 0\ for\ k = 2,3, \dots$$

ACF for MA (2)

$$Y_t = e_t - \theta e_{t-1} - \theta_2 e_{t-2}$$

$$E(Y_t) = 0,$$

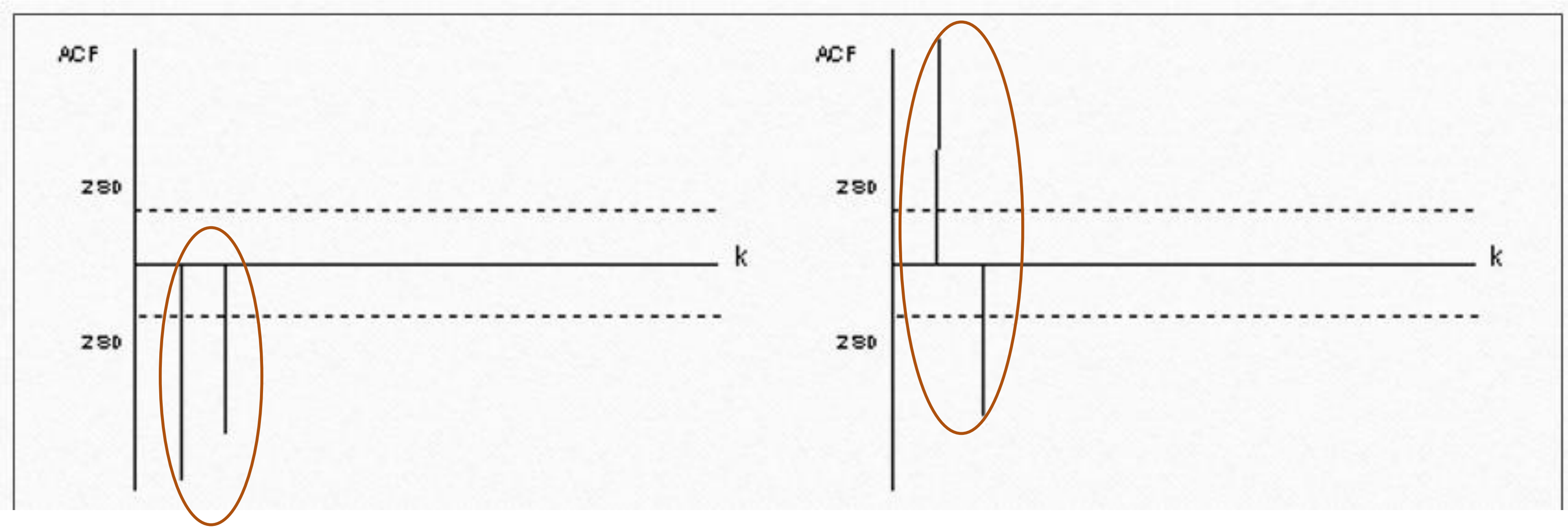
$$Var(Y_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma_e^2$$

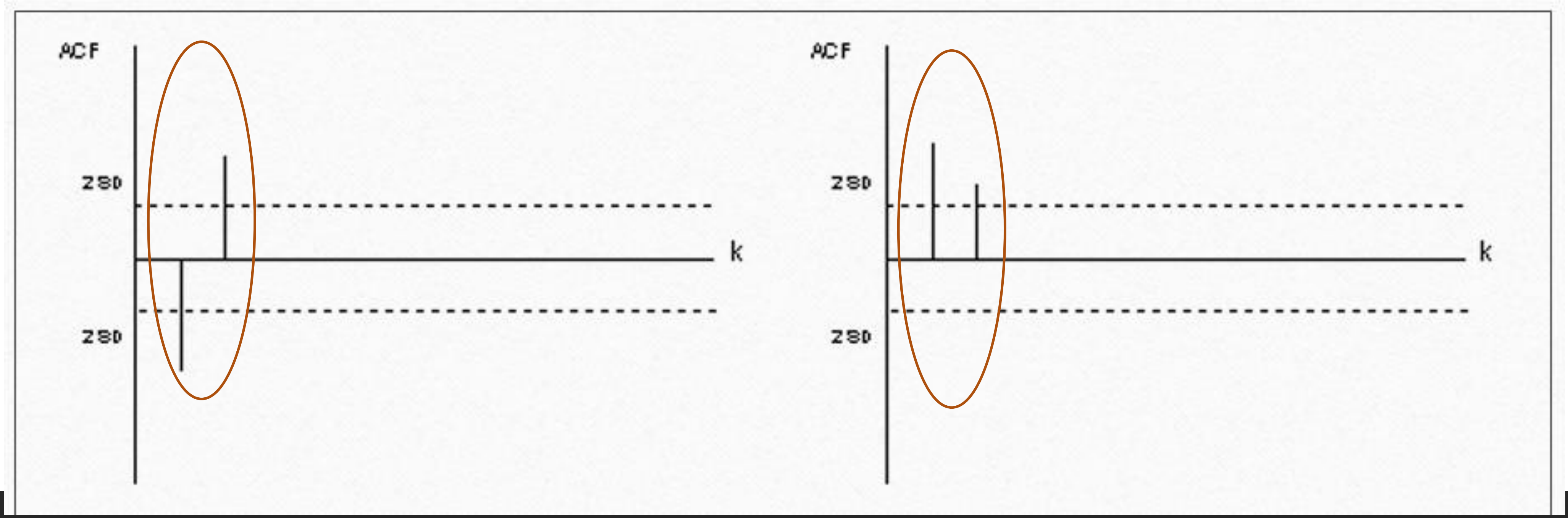
$$Cov(Y_t, Y_{t-1}) = \gamma_1 = (-\theta_1 + \theta_1\theta_2)\sigma_e^2$$

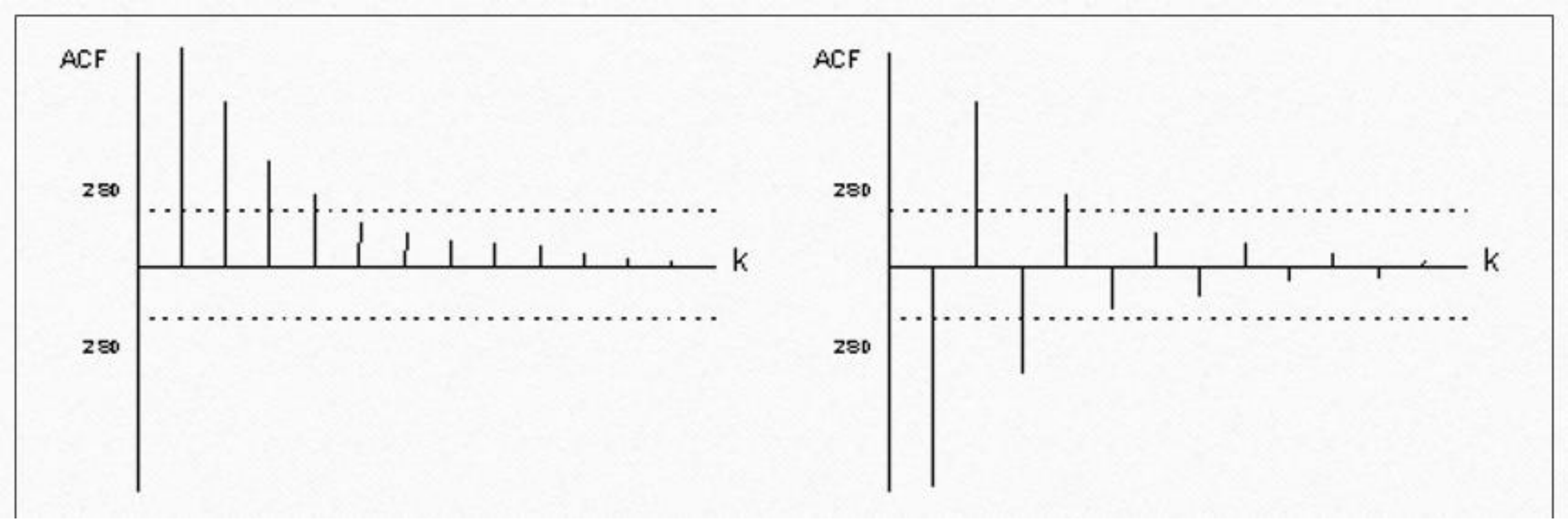
$$Cov(Y_t, Y_{t-2}) = \gamma_2 = -\theta_2\sigma_e^2,\ Cov(Y_t, Y_{t-3}) = \gamma_3 = 0$$

$$\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2}, \rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}, \rho_k = 0\ for\ k = 3,4,\ldots$$

ACF for MA (2)







ACF for AR (1)

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

$$E(Y_t) = 0,$$

$$Var(Y_t) = \gamma_0 = \frac{\sigma_e^2}{1-\phi^2}$$

$$\gamma_k = \phi\gamma_{k-1} = \phi^k \frac{\sigma_e^2}{1-\phi^2} \quad for\ k = 1,2,3,\ldots$$

$$\boxed{\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k} \quad for\ k = 1,2,3,\ldots$$

ACF for AR (2)

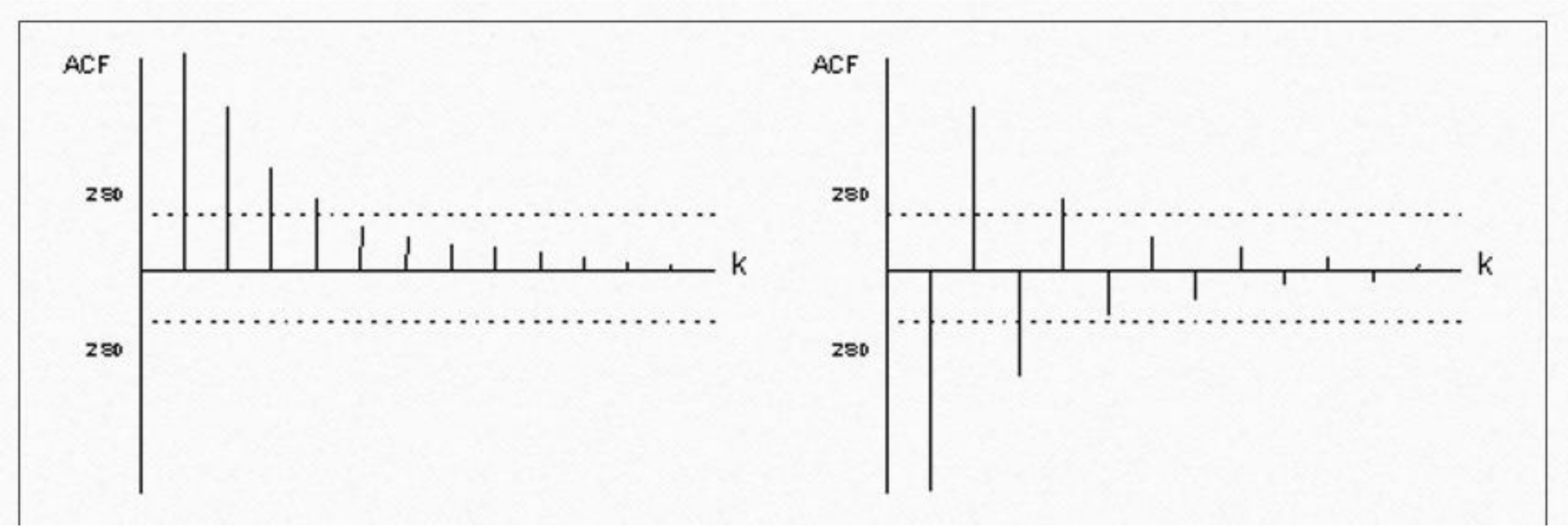
$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$$

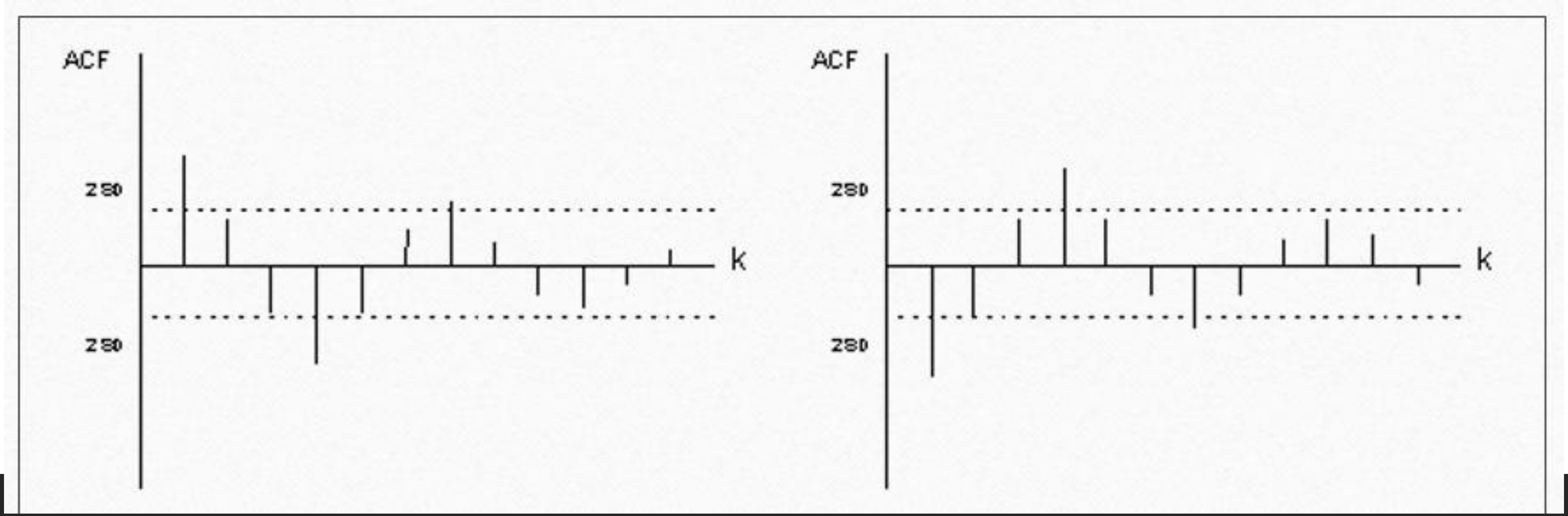
$$E(Y_t) = 0,$$
$$Var(Y_t) = \gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma_e^2$$

$$Cov(Y_t, Y_{t-k}) = \gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2}\ for\ k = 1,2,3, \dots$$

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} for\ k = 1,2,3, \dots$$

ACF
for AR (2)



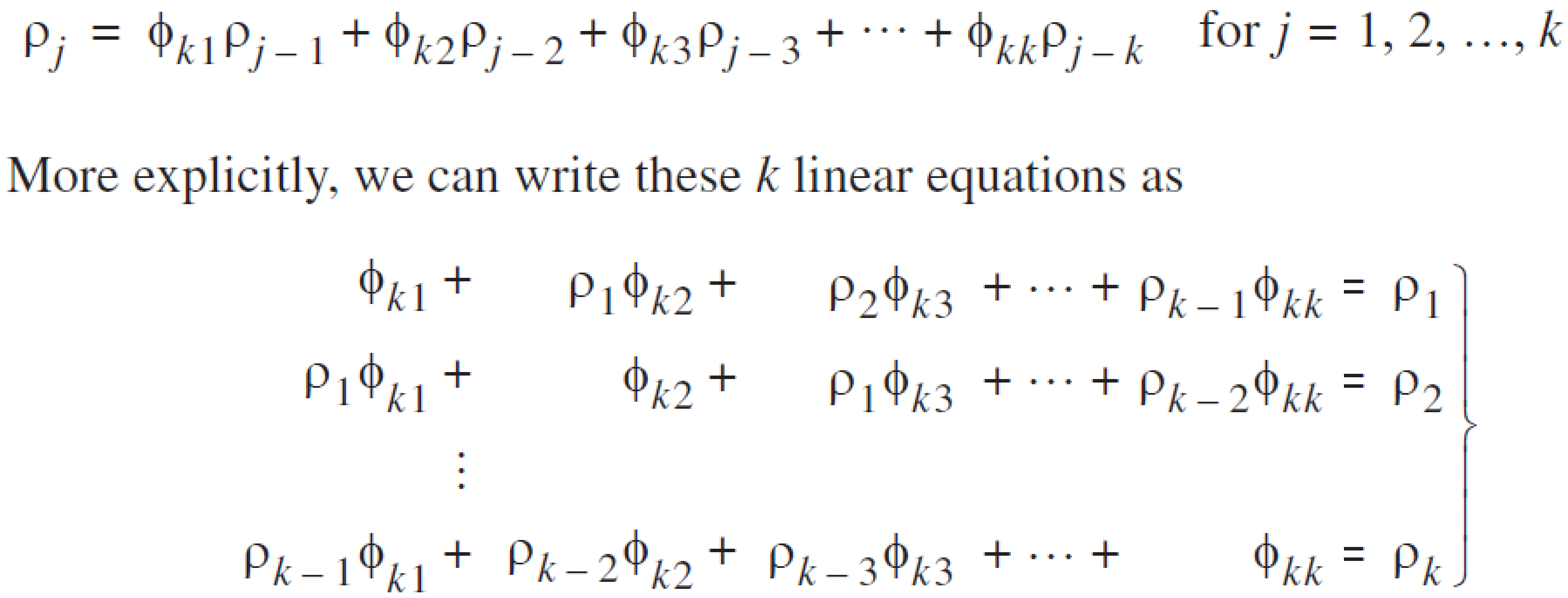


The Partial Autocorrelation Functions

- Since for MA(q) models the autocorrelation function is zero for lags beyond q, the sample autocorrelation is a good indicator of the order of the process.
- However, the autocorrelations of an AR(p) model do not become zero after a certain number of lags—they die off rather than cut off. **So a different function is needed to help determine the order of autoregressive models**.
- A function may be defined as the correlation between Y_t and Y_{t-k} *after removing the effect of the intervening variables* $Y_{t-1}, Y_{t-2}, Y_{t-3}, \ldots, Y_{t-k+1}$
- This coefficient is called **the partial autocorrelation at lag k** and will be denoted by ϕ_{kk}.

The Partial Autocorrelation Functions

A general method for finding the partial autocorrelation function for any stationary process with autocorrelation function ρ_k is as follows (see Anderson 1971, pp. 187–188, for example). For a given lag k, it can be shown that the ϕ_{kk} satisfy the Yule-Walker equations

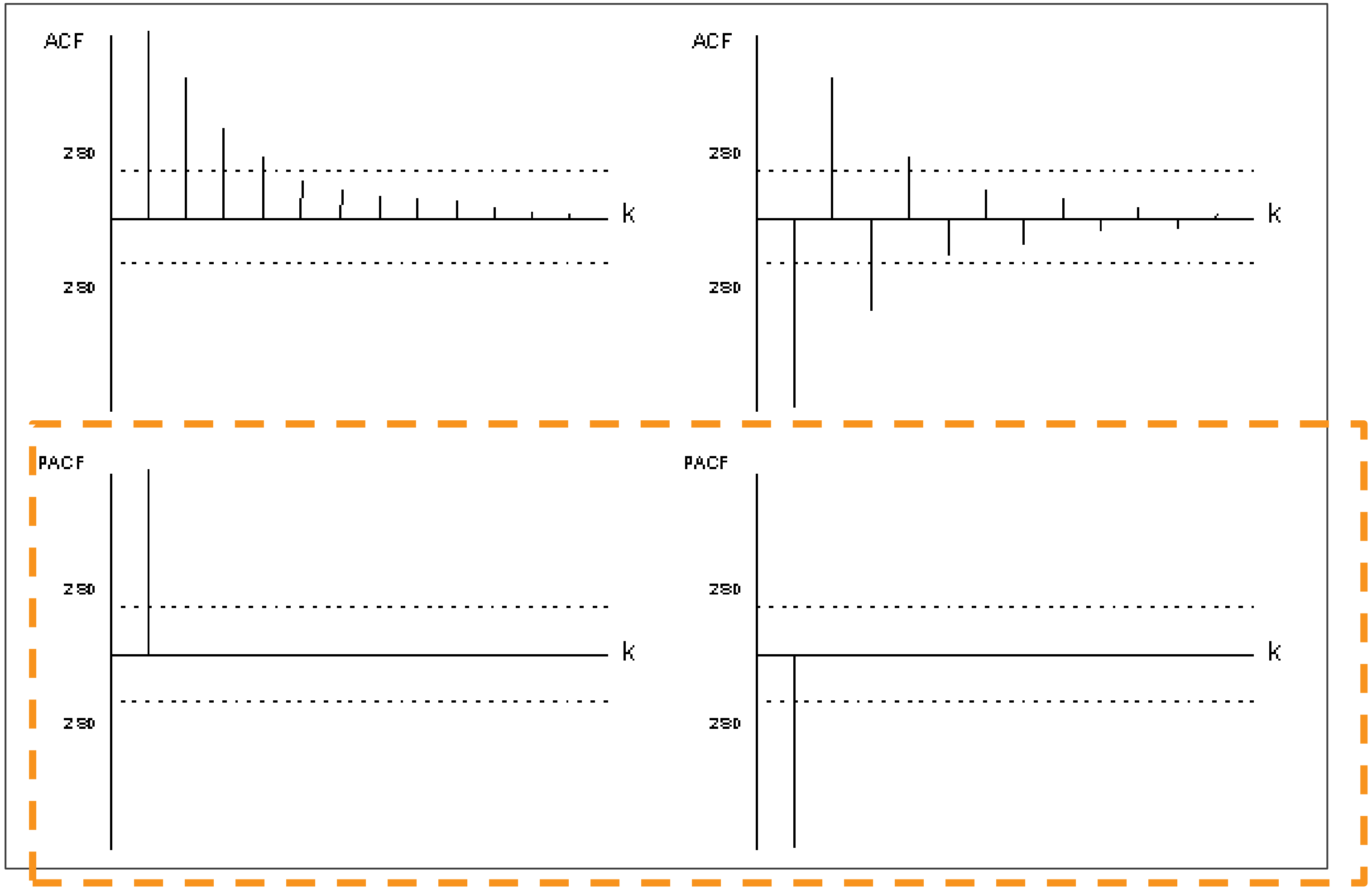


$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \phi_{k3}\rho_{j-3} + \cdots + \phi_{kk}\rho_{j-k} \quad \text{for } j = 1, 2, \ldots, k$$

More explicitly, we can write these k linear equations as

$$\left.\begin{array}{rrrrr}
\phi_{k1} + & \rho_1\phi_{k2} + & \rho_2\phi_{k3} & + \cdots + \rho_{k-1}\phi_{kk} = & \rho_1 \\
\rho_1\phi_{k1} + & \phi_{k2} + & \rho_1\phi_{k3} & + \cdots + \rho_{k-2}\phi_{kk} = & \rho_2 \\
\vdots & & & & \\
\rho_{k-1}\phi_{k1} + & \rho_{k-2}\phi_{k2} + & \rho_{k-3}\phi_{k3} & + \cdots + \qquad \phi_{kk} = & \rho_k
\end{array}\right\}$$

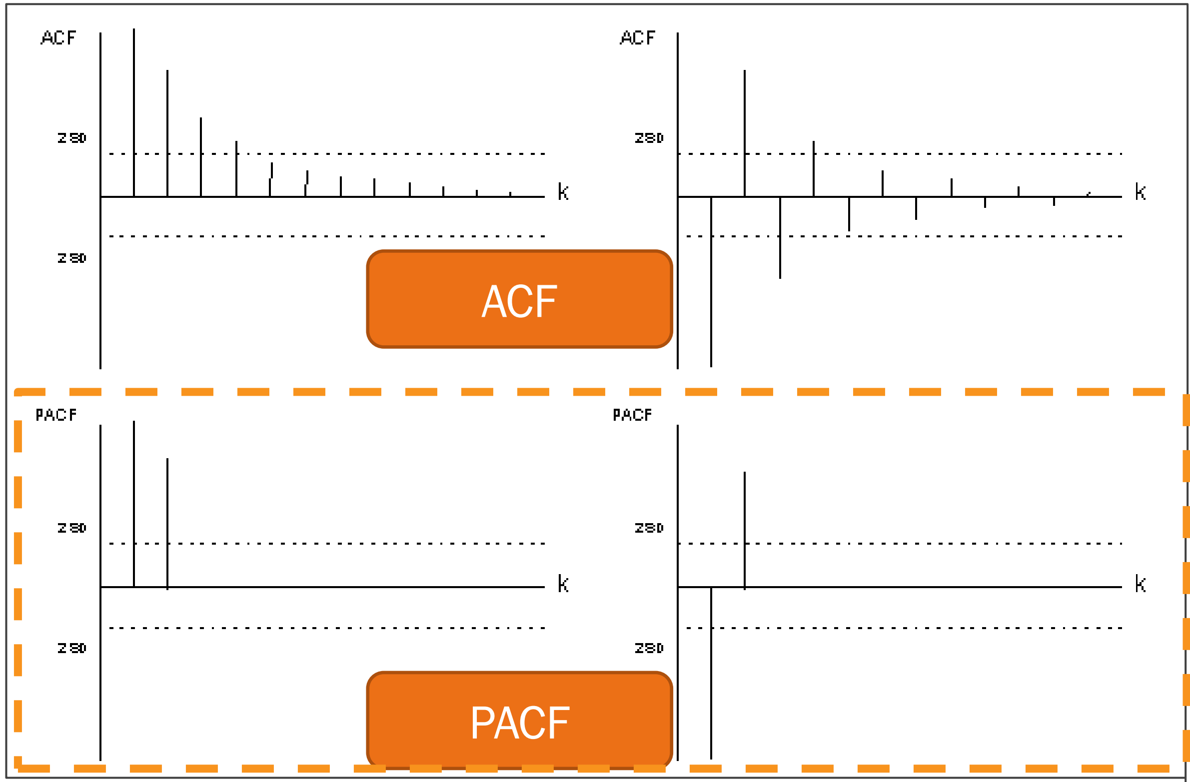
PACF for AR (1)

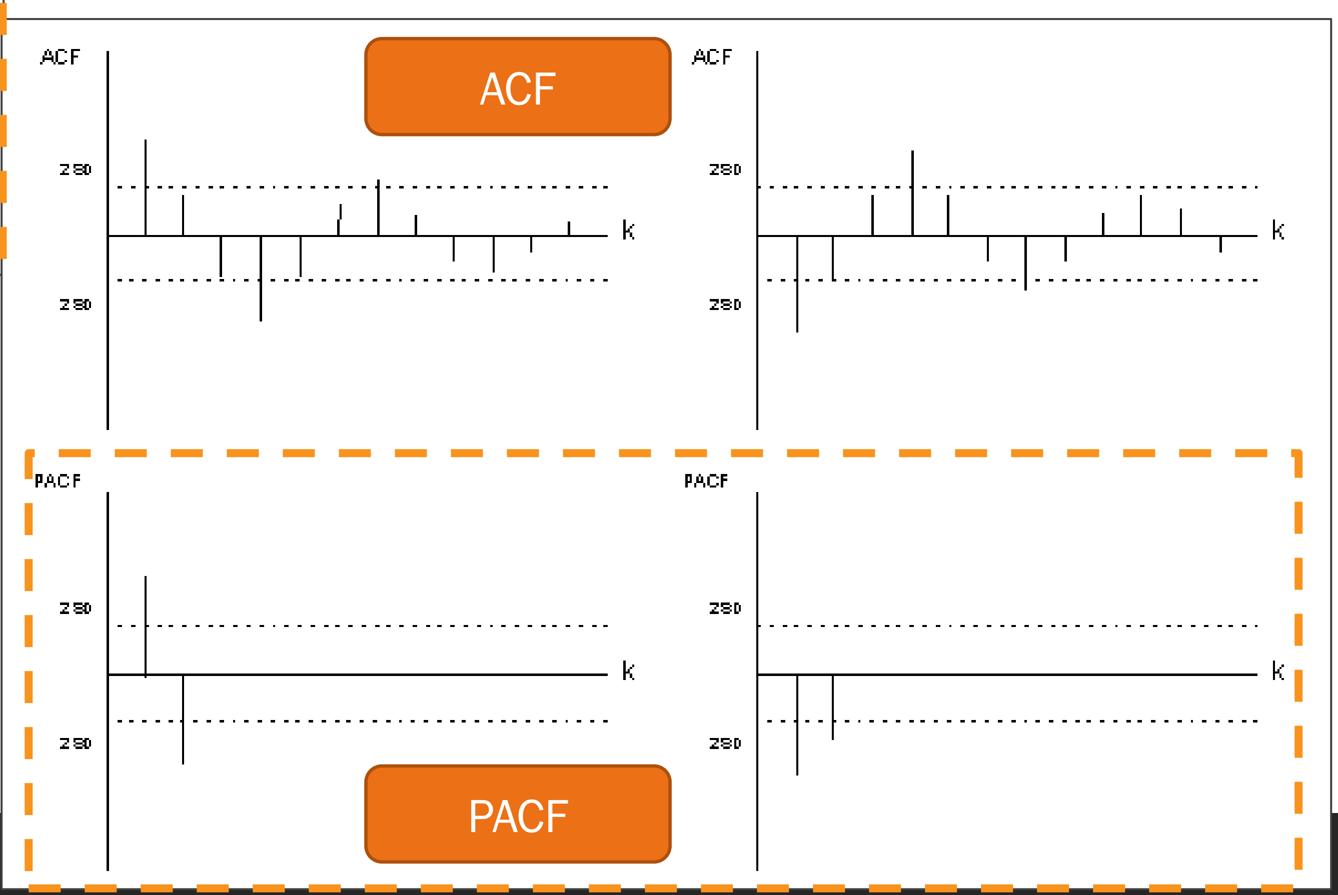


ACF

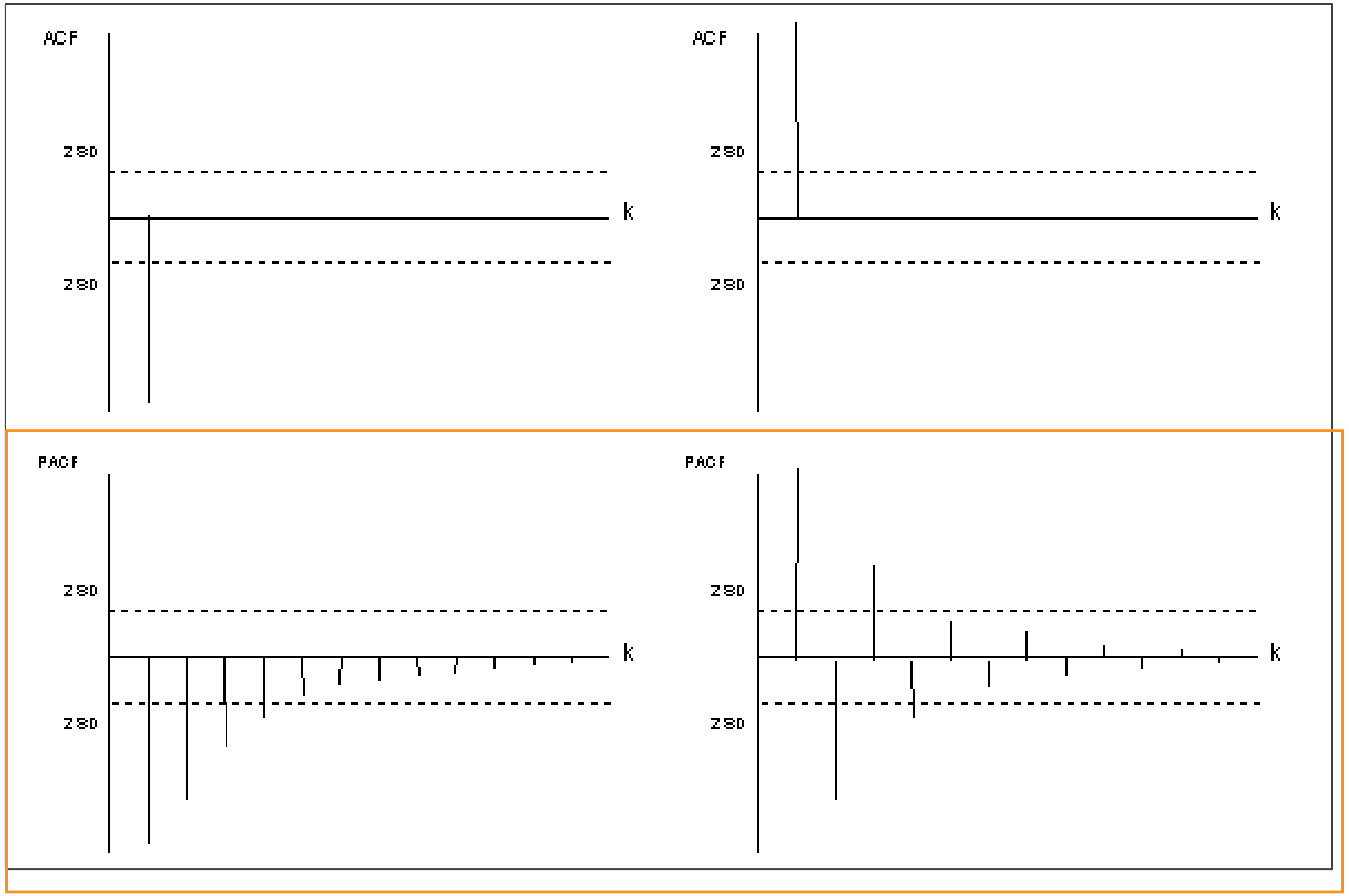
PACF

PACF
for AR (2)


ACF
ACF
2SD
2SD
k
k
ACF
PACF
PACF
2SD
2SD
k
k
PACF


ACF
ACF
ACF
2SD
2SD
k
k
PACF
PACF
2SD
2SD
k
k
PACF

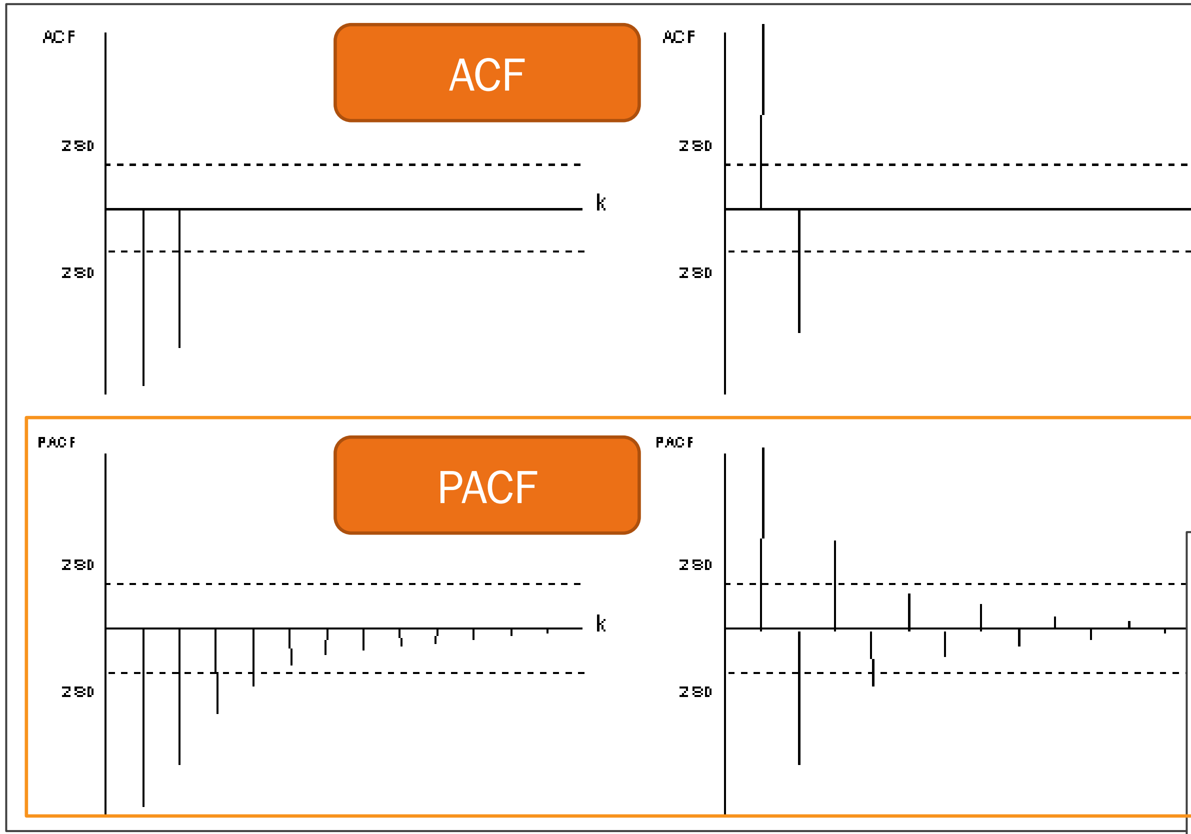
PACF for MA (1)

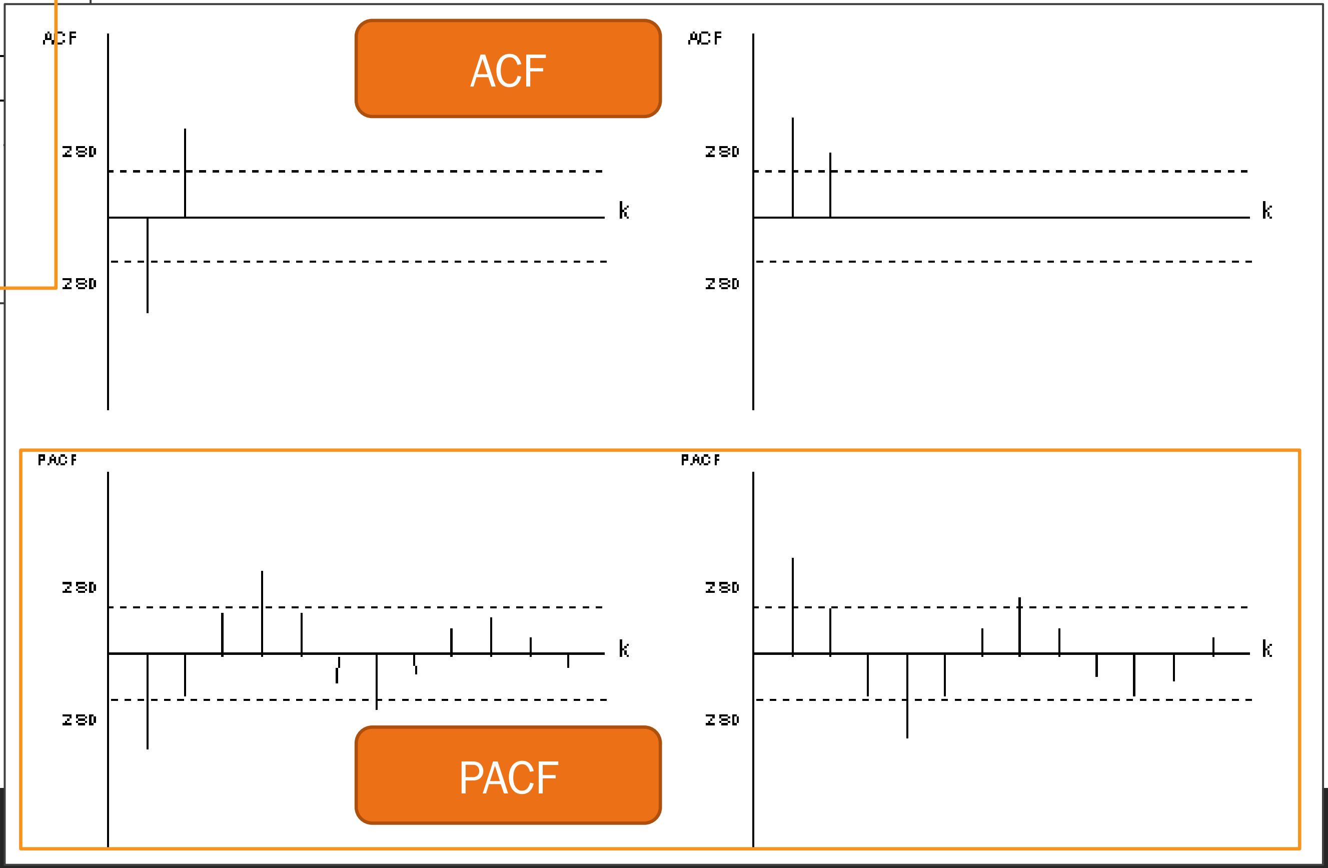


ACF

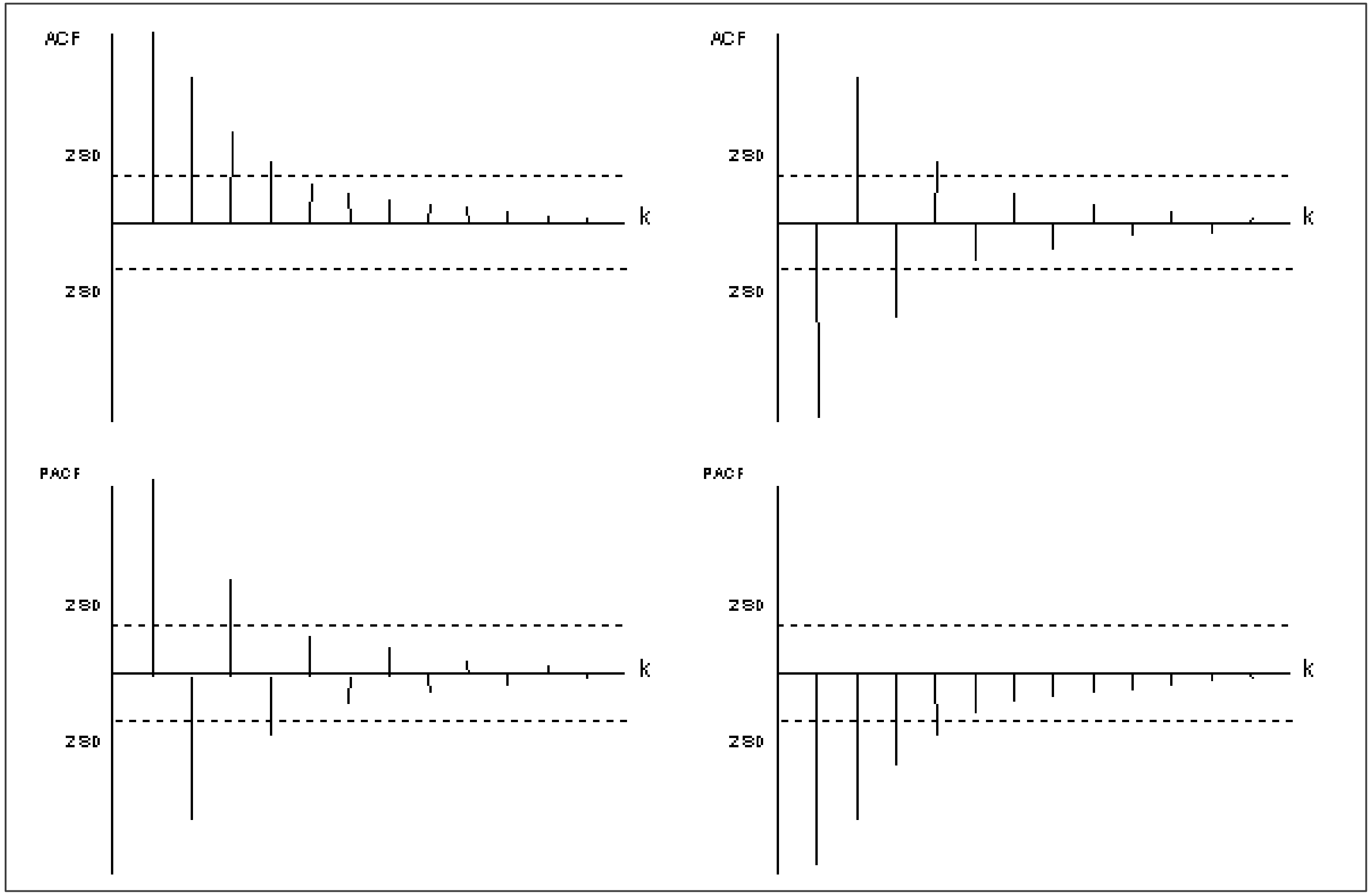
PACF

PACF for MA (2)


ACF
ACF
ACF
2SD
2SD
k
PACF
PACF
PACF
2SD
2SD
k


ACF
ACF
ACF
2SD
2SD
k
PACF
PACF
PACF
2SD
2SD
k

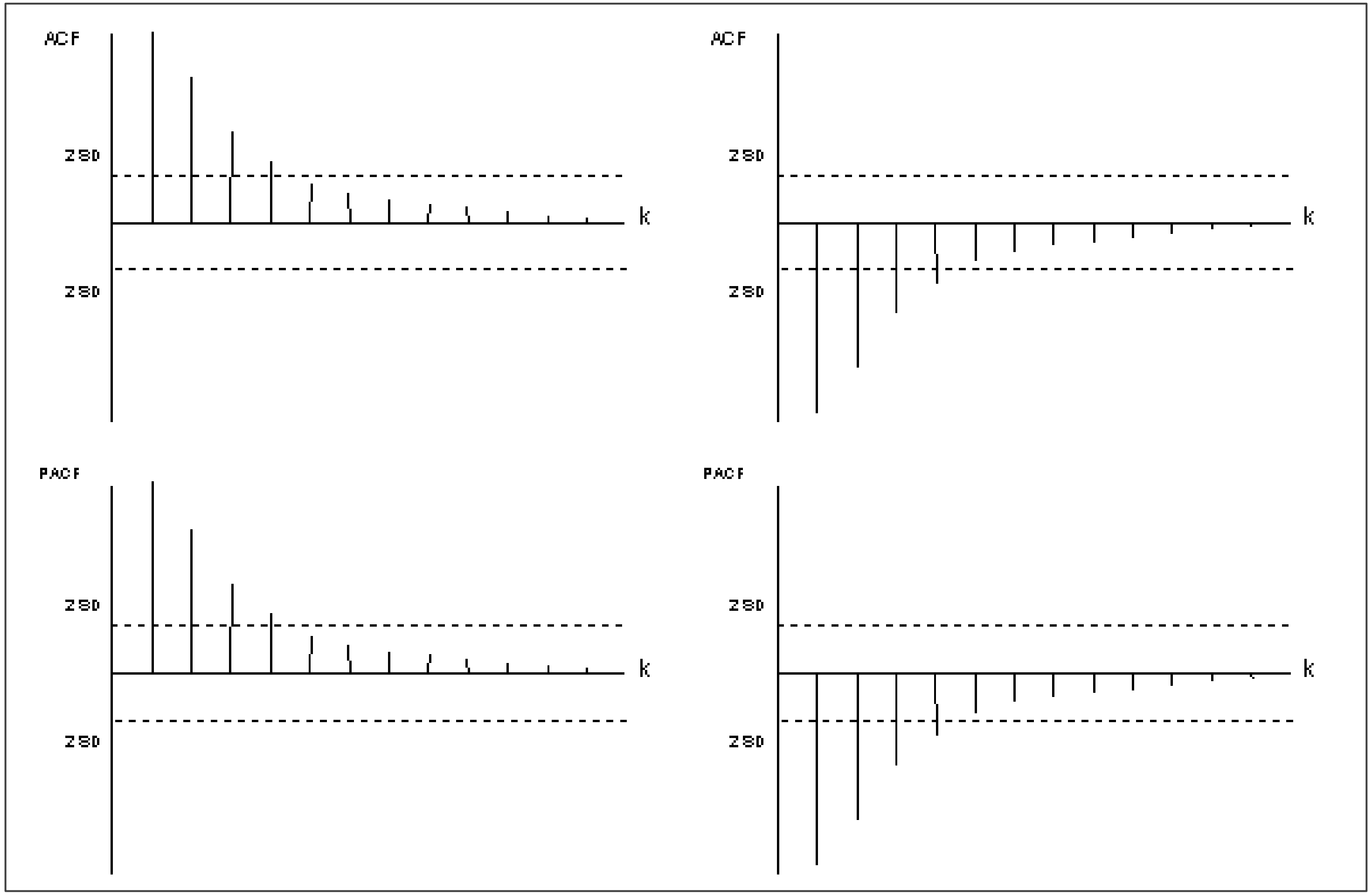
ACF and PACF for ARMA(1,1)



ACF

PACF

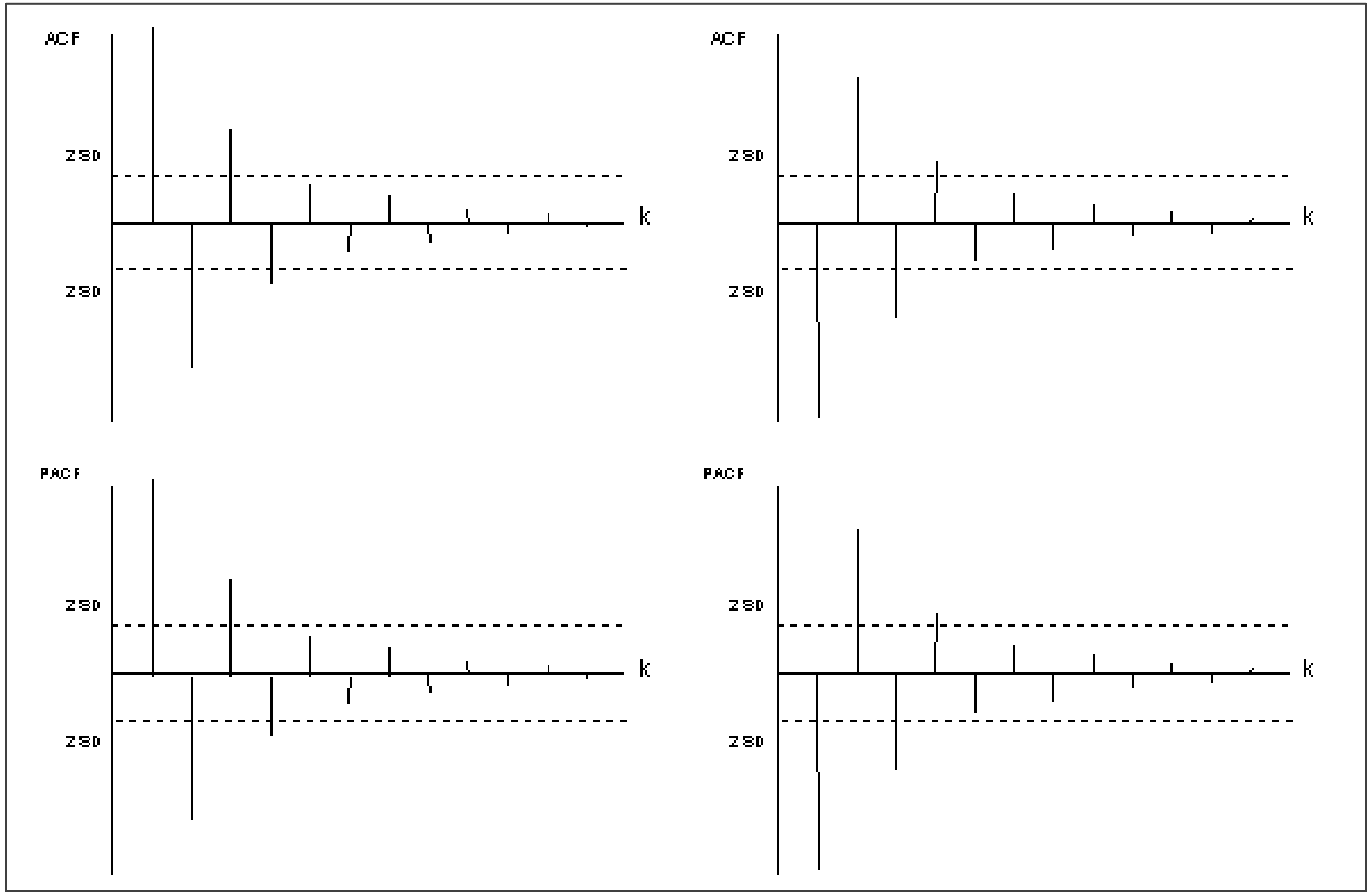
ACF and PACF for ARMA(1,1)



ACF

PACF

ACF and PACF for ARMA(1,1)



ACF

PACF

General Behavior of the ACF and PACF for ARMA Models

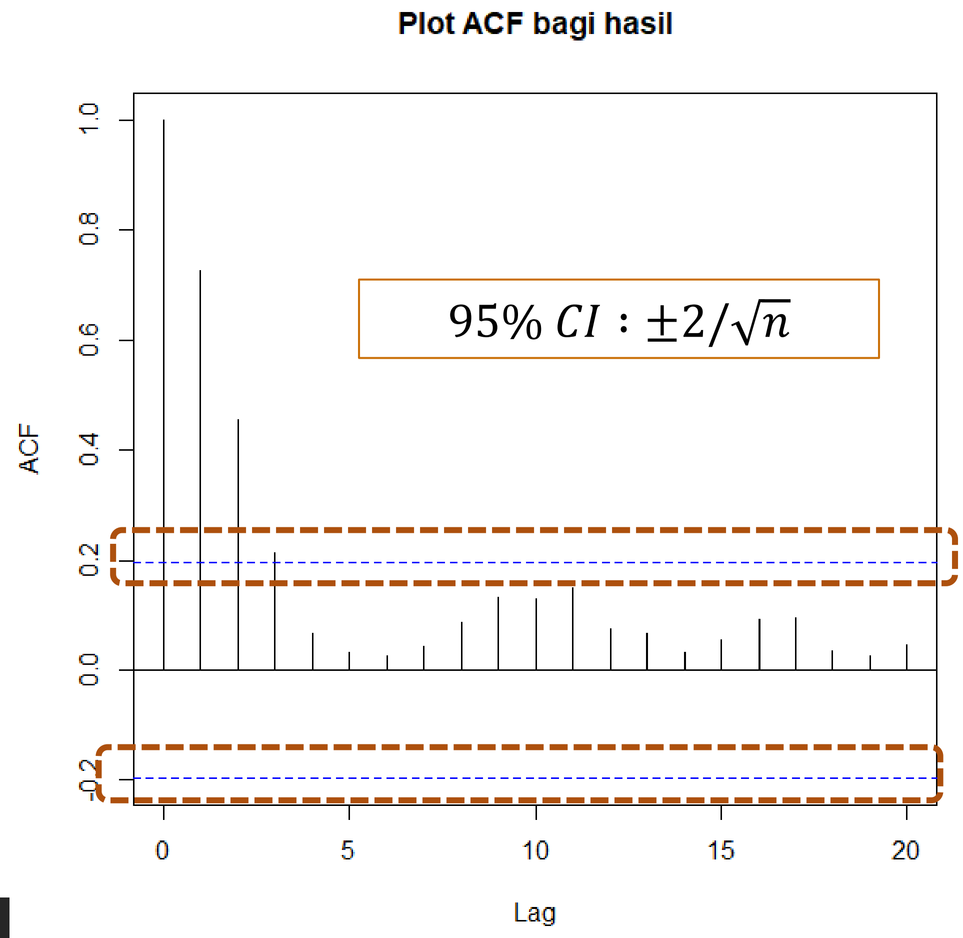
	AR(p)	MA(q)	ARMA(p, q), $p>0$, and $q>0$
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

The **sample autocorrelation function**, r_k, at lag k as

$$r_k = \frac{\sum_{t=k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2} \qquad \text{for } k = 1, 2, \ldots$$

$$Var(r_k) \approx \frac{1}{n}$$

$$S(r_k) = 1/\sqrt{n}$$



95% CI : $\pm 2/\sqrt{n}$

The **sample partial autocorrelation function**, $\hat{\phi}_{kk}$

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

where

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

$$Var(\hat{\phi}_{kk}) = 1/n$$

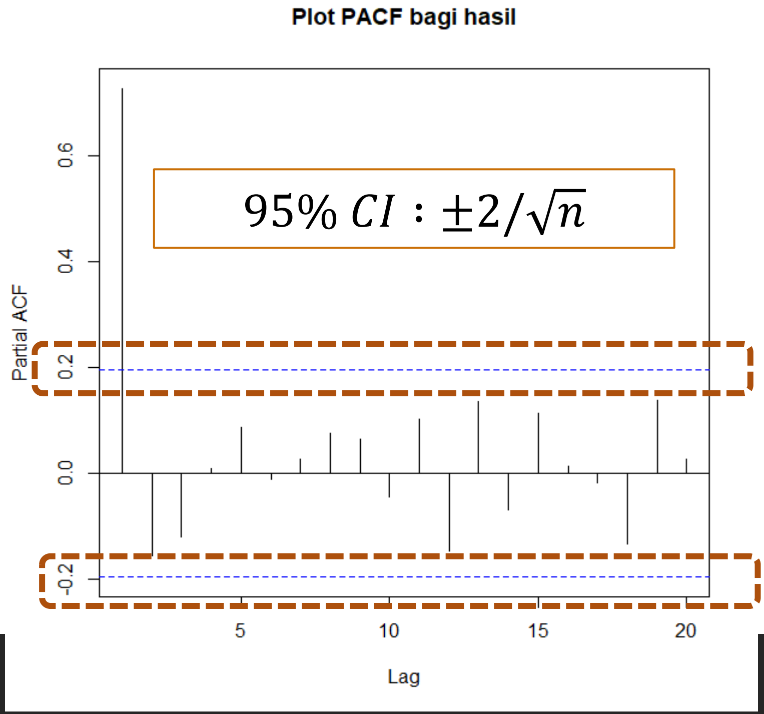
$$S(\hat{\phi}_{kk}) = 1/\sqrt{n}$$

For example, using $\phi_{11} = \rho_1$ to get started, we have

$$\phi_{22} = \frac{\rho_2 - \phi_{11}\rho_1}{1 - \phi_{11}\rho_1} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

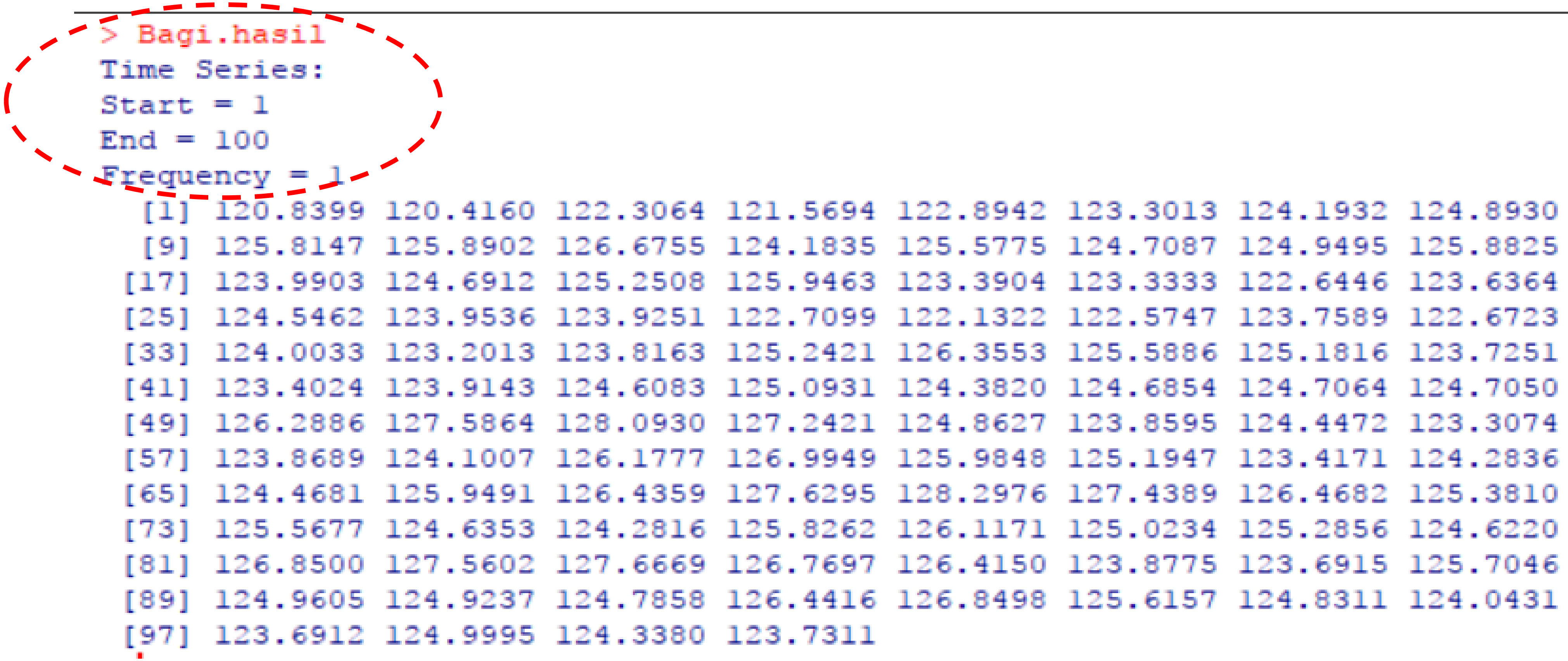
(as before) with $\phi_{21} = \phi_{11} - \phi_{22}\phi_{11}$, which is needed for the next step. Then

$$\phi_{33} = \frac{\rho_3 - \phi_{21}\rho_2 - \phi_{22}\rho_1}{1 - \phi_{21}\rho_1 - \phi_{22}\rho_2}$$

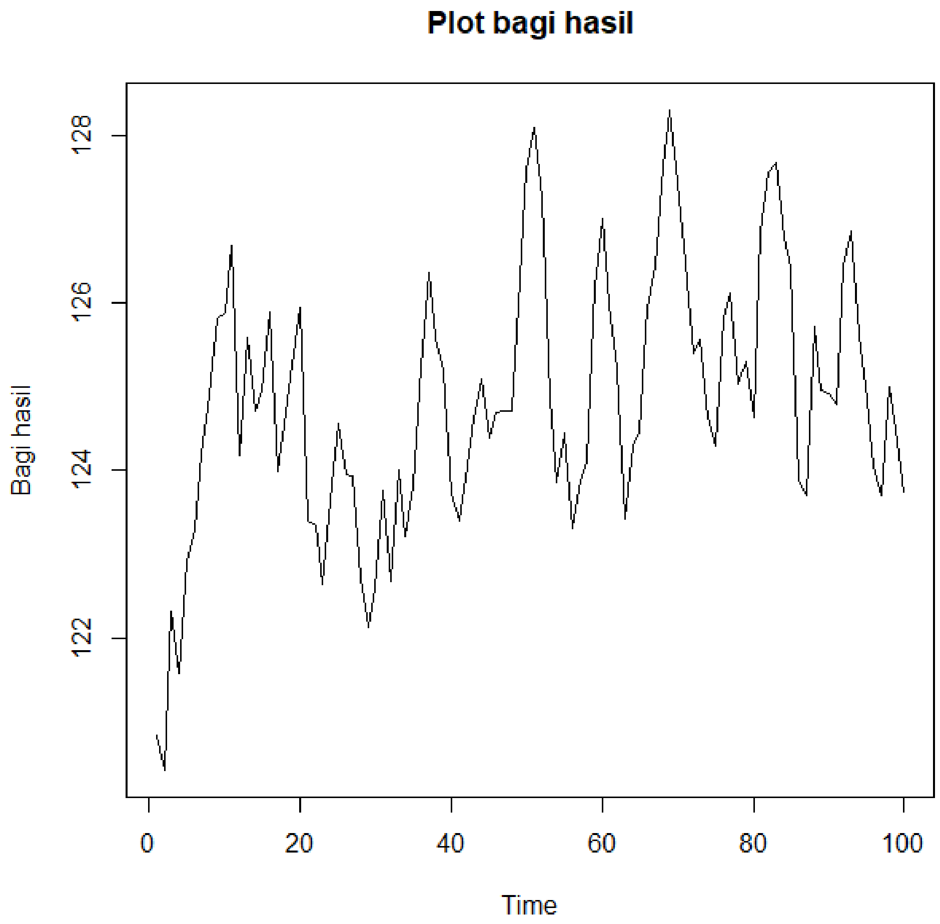


95% CI : $\pm 2/\sqrt{n}$

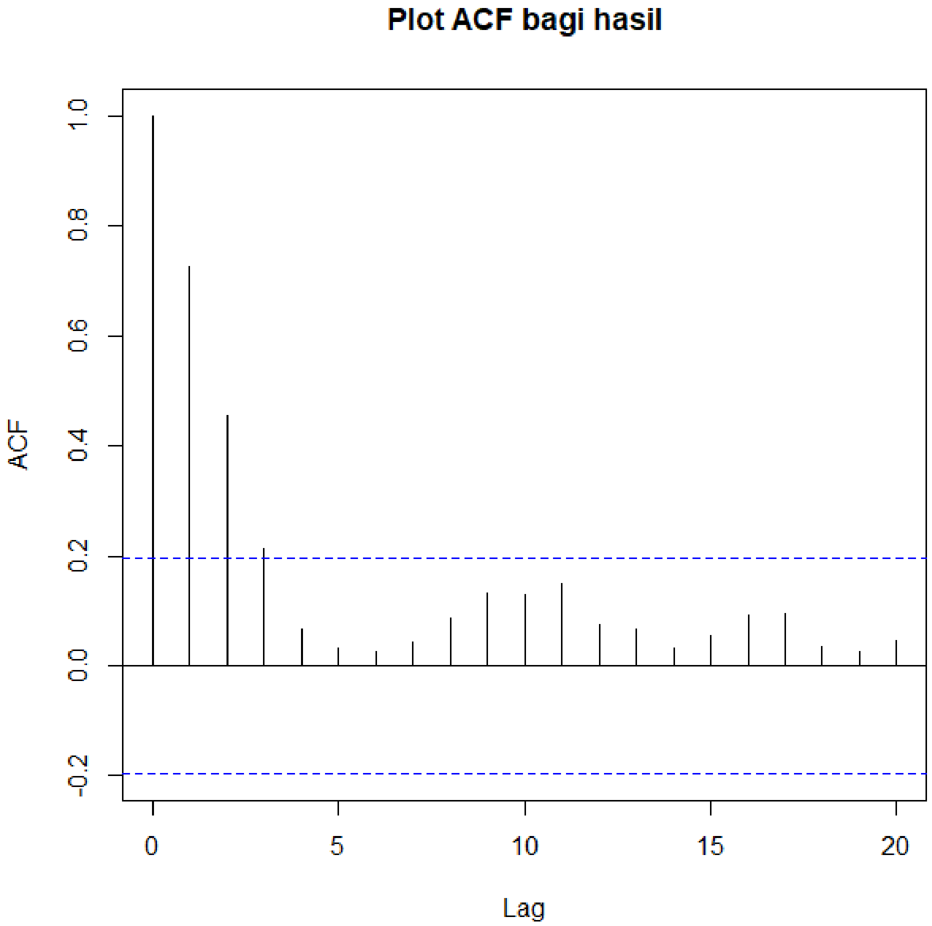
Illustration 1



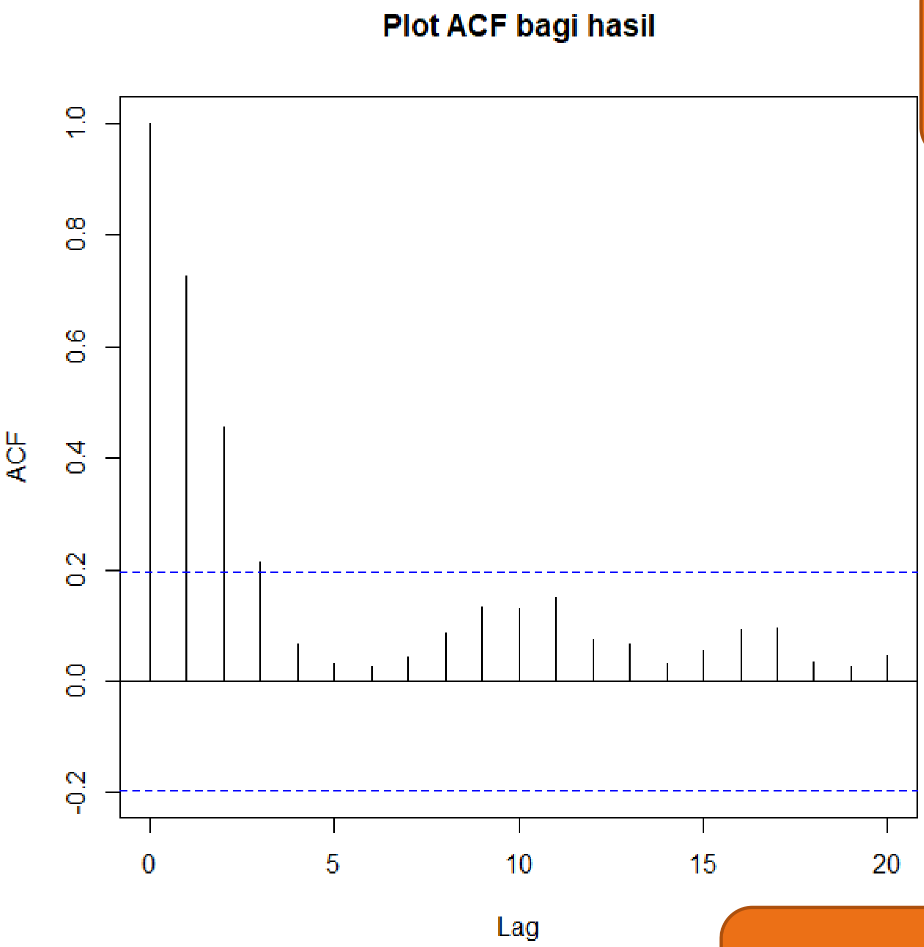
```
> Bagi.hasil
Time Series:
Start = 1
End = 100
Frequency = 1
  [1] 120.8399 120.4160 122.3064 121.5694 122.8942 123.3013 124.1932 124.8930
  [9] 125.8147 125.8902 126.6755 124.1835 125.5775 124.7087 124.9495 125.8825
 [17] 123.9903 124.6912 125.2508 125.9463 123.3904 123.3333 122.6446 123.6364
 [25] 124.5462 123.9536 123.9251 122.7099 122.1322 122.5747 123.7589 122.6723
 [33] 124.0033 123.2013 123.8163 125.2421 126.3553 125.5886 125.1816 123.7251
 [41] 123.4024 123.9143 124.6083 125.0931 124.3820 124.6854 124.7064 124.7050
 [49] 126.2886 127.5864 128.0930 127.2421 124.8627 123.8595 124.4472 123.3074
 [57] 123.8689 124.1007 126.1777 126.9949 125.9848 125.1947 123.4171 124.2836
 [65] 124.4681 125.9491 126.4359 127.6295 128.2976 127.4389 126.4682 125.3810
 [73] 125.5677 124.6353 124.2816 125.8262 126.1171 125.0234 125.2856 124.6220
 [81] 126.8500 127.5602 127.6669 126.7697 126.4150 123.8775 123.6915 125.7046
 [89] 124.9605 124.9237 124.7858 126.4416 126.8498 125.6157 124.8311 124.0431
 [97] 123.6912 124.9995 124.3380 123.7311
```



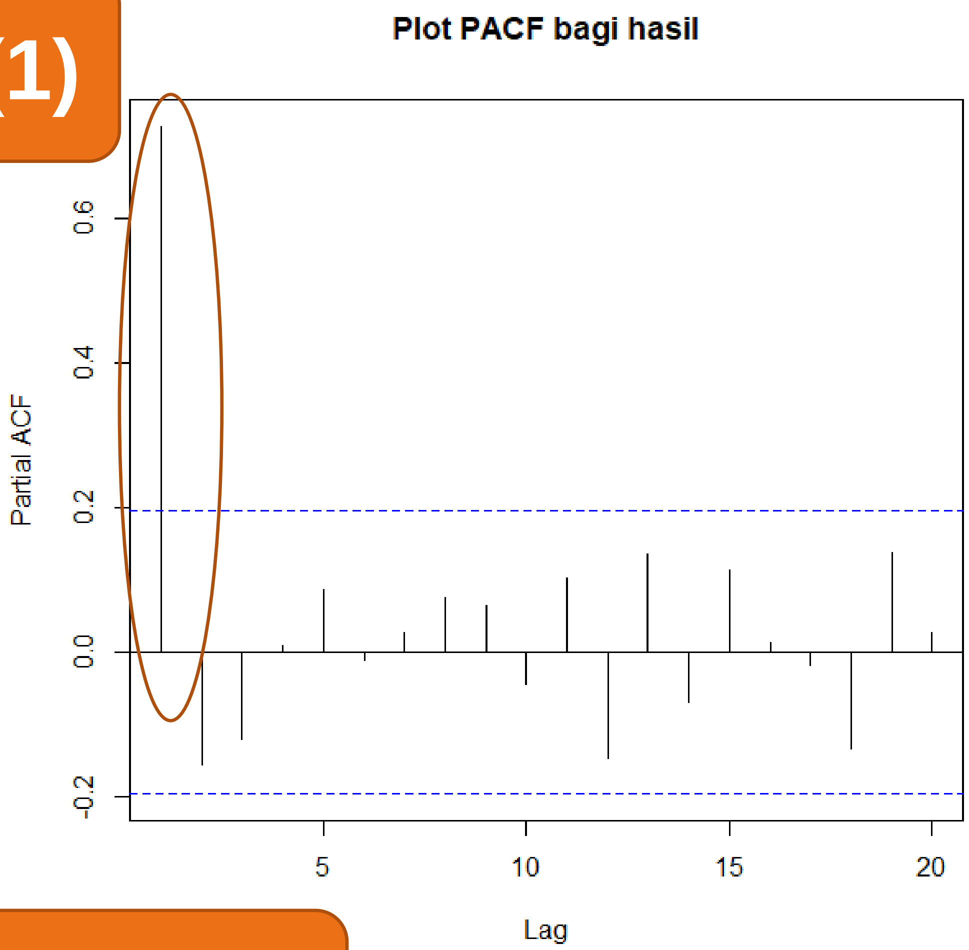
Plot bagi hasil



Plot ACF bagi hasil

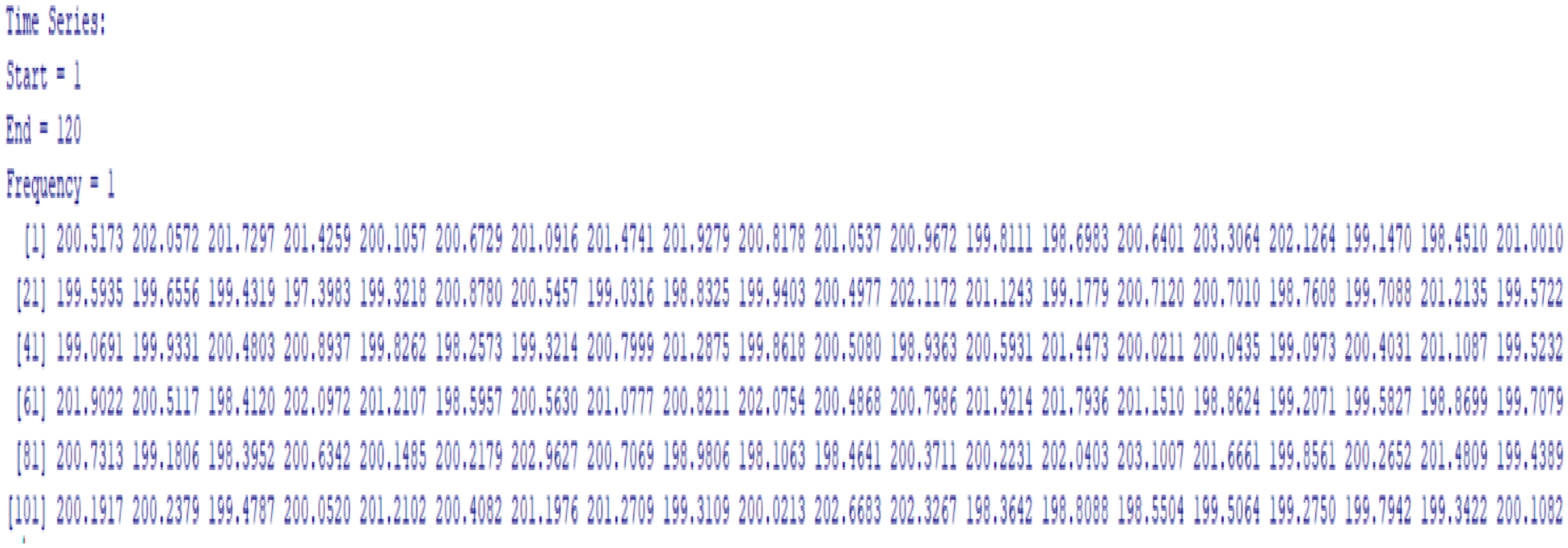

Plot ACF bagi hasil

AR(1)

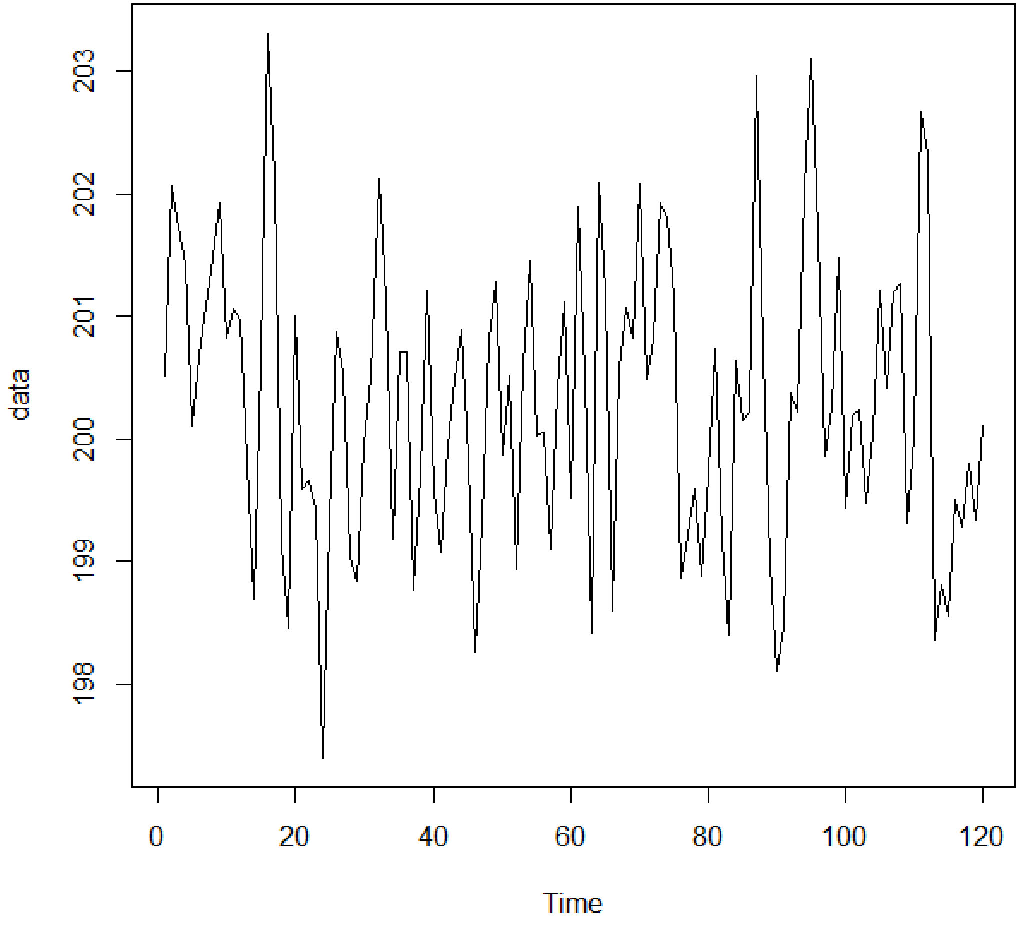

Plot PACF bagi hasil

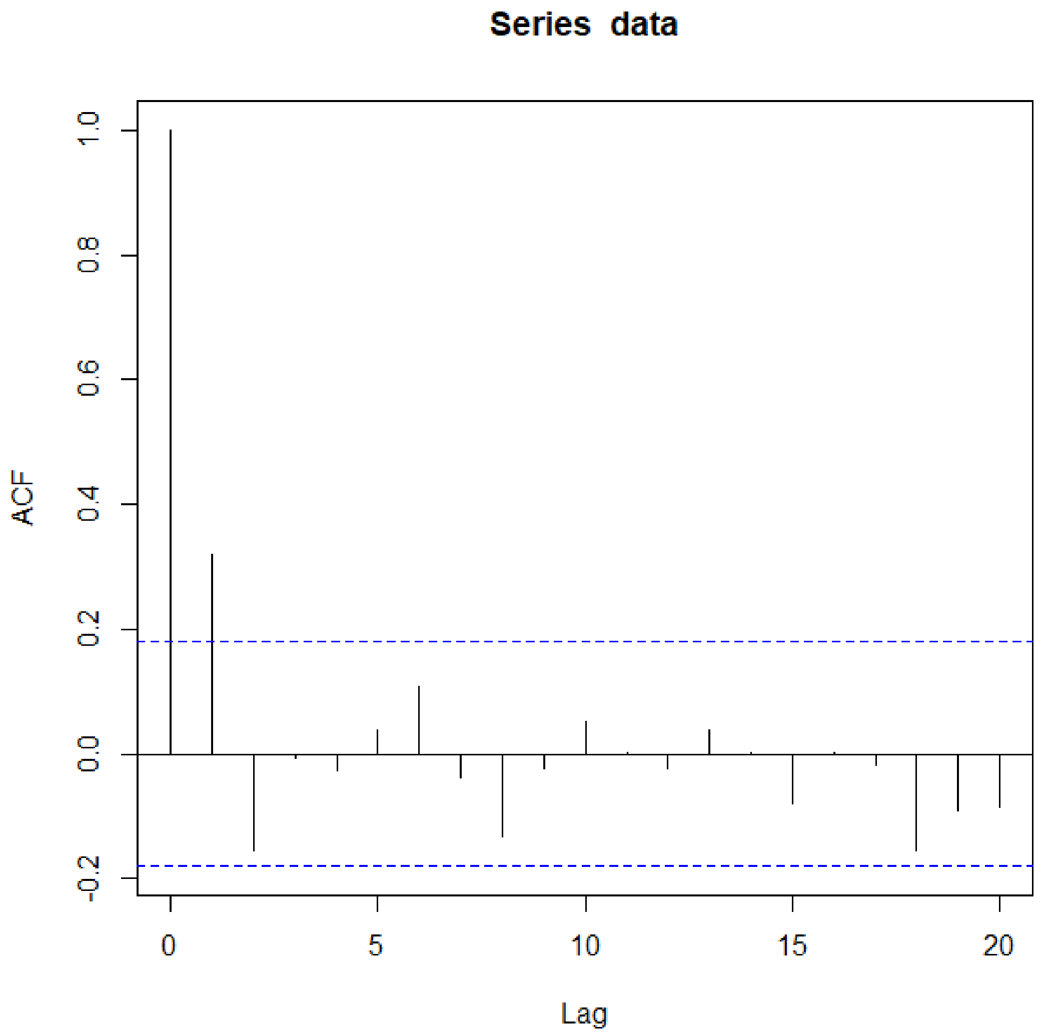
→ **ACF tails off**
→ **PACF cuts off after lag 1**

Illustration 2

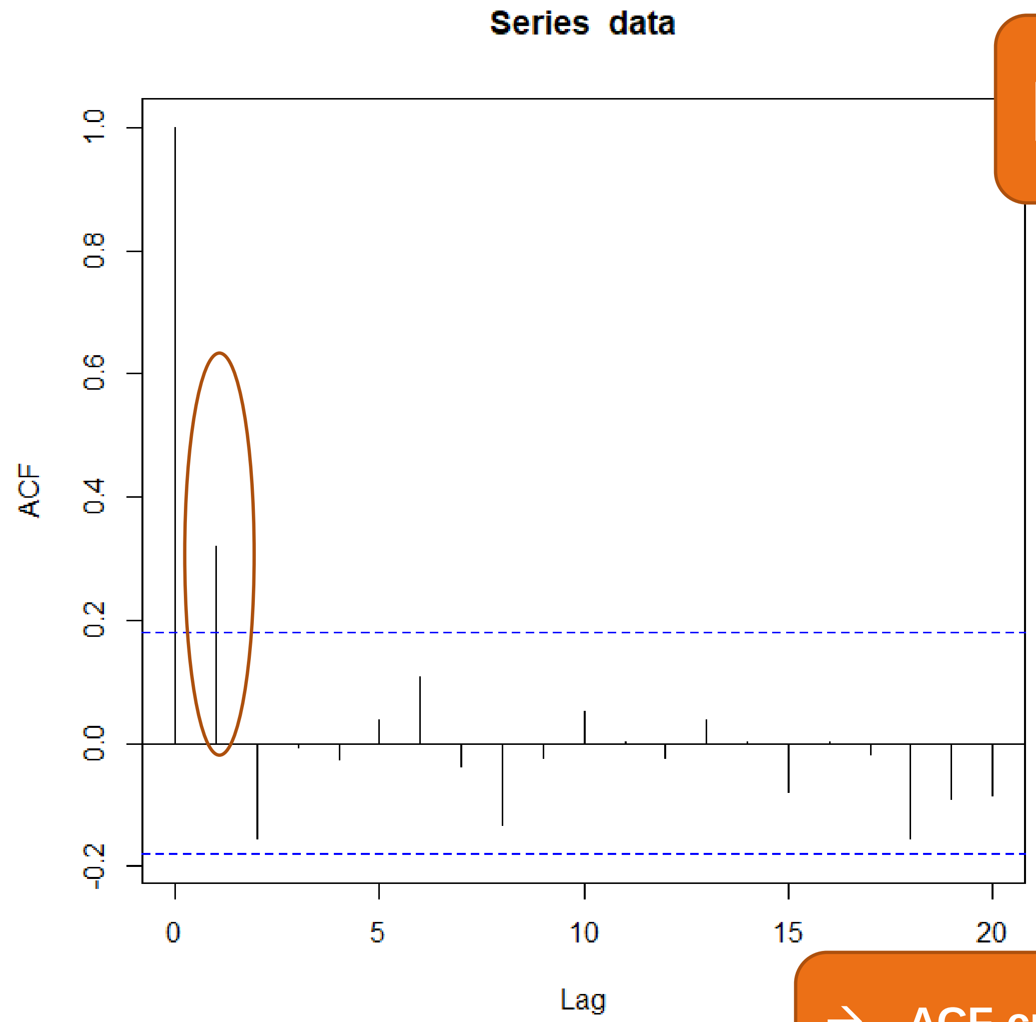


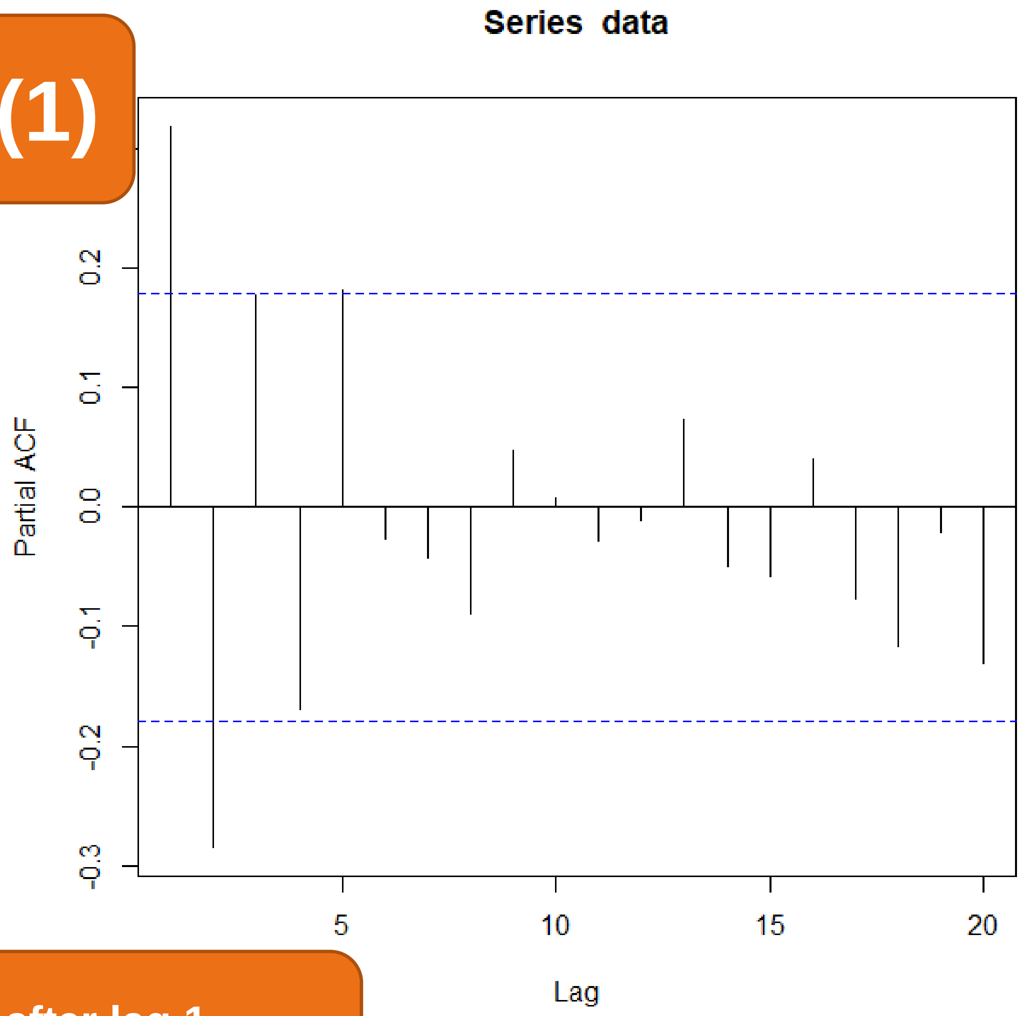
```
Time Series:
Start = 1
End = 120
Frequency = 1
  [1] 200.5173 202.0572 201.7297 201.4259 200.1057 200.6729 201.0916 201.4741 201.9279 200.8178 201.0537 200.9672 199.8111 198.6983 200.6401 203.3064 202.1264 199.1470 198.4510 201.0010
 [21] 199.5935 199.6556 199.4319 197.3983 199.3218 200.8780 200.5457 199.0316 198.8325 199.9403 200.4977 202.1172 201.1243 199.1779 200.7120 200.7010 198.7608 199.7088 201.2135 199.5722
 [41] 199.0691 199.9331 200.4803 200.8937 199.8262 198.2573 199.3214 200.7999 201.2875 199.8618 200.5080 198.9363 200.5931 201.4473 200.0211 200.0435 199.0973 200.4031 201.1087 199.5232
 [61] 201.9022 200.5117 198.4120 202.0972 201.2107 198.5957 200.5630 201.0777 200.8211 202.0754 200.4868 200.7986 201.9214 201.7936 201.1510 198.8624 199.2071 199.5827 198.8699 199.7079
 [81] 200.7313 199.1806 198.3952 200.6342 200.1485 200.2179 202.9627 200.7069 198.9806 198.1063 198.4641 200.3711 200.2231 202.0403 203.1007 201.6661 199.8561 200.2652 201.4809 199.4389
[101] 200.1917 200.2379 199.4787 200.0520 201.2102 200.4082 201.1976 201.2709 199.3109 200.0213 202.6683 202.3267 198.3642 198.8088 198.5504 199.5064 199.2750 199.7942 199.3422 200.1082
```


data
198
199
200
201
202
203
0
20
40
60
80
100
120
Time


Series data
ACF
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
Lag

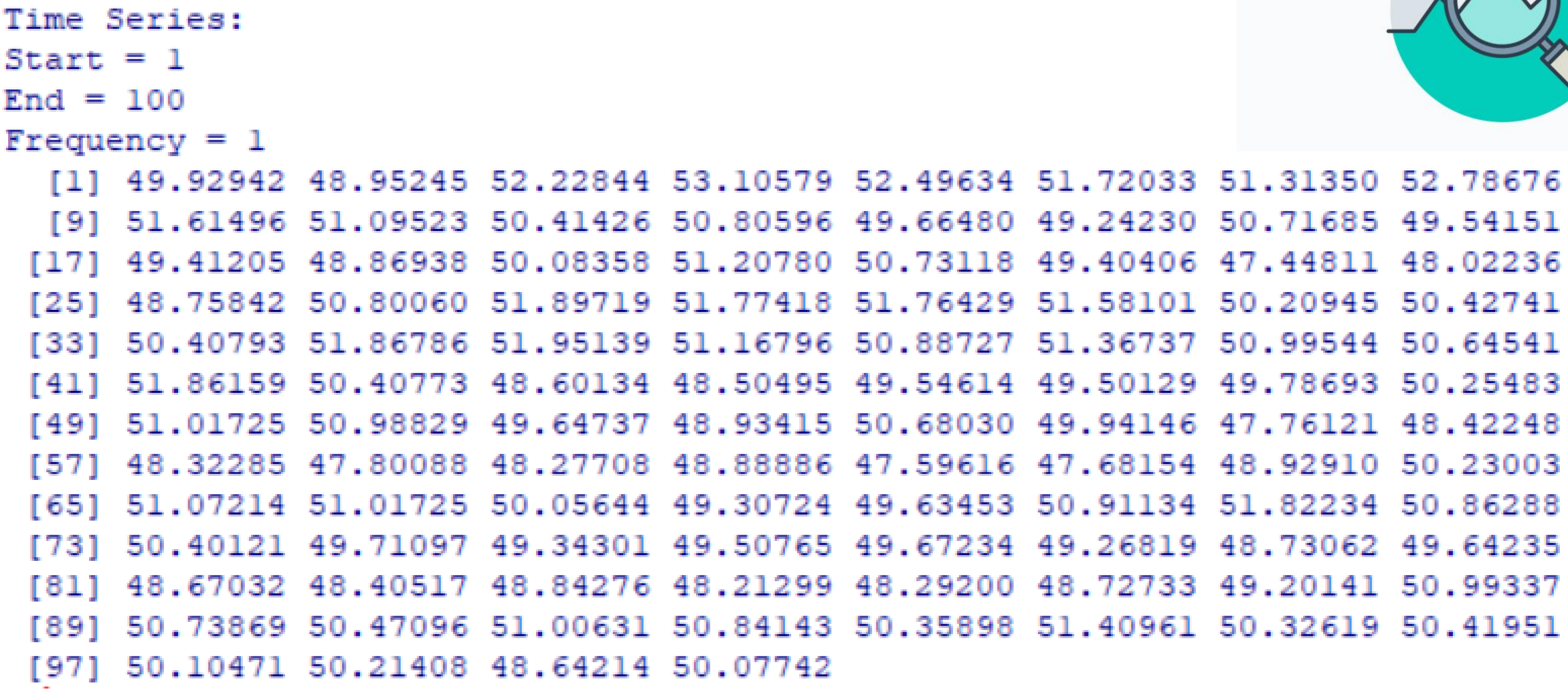
MA(1)



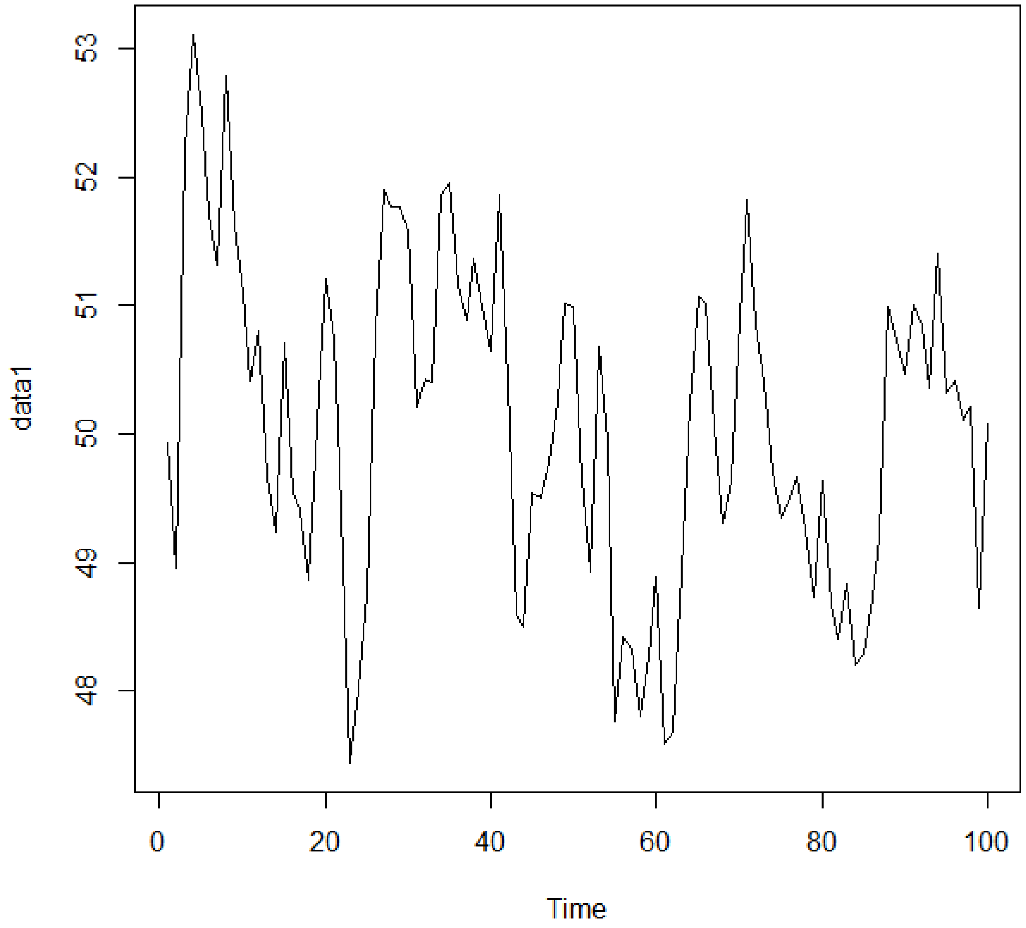


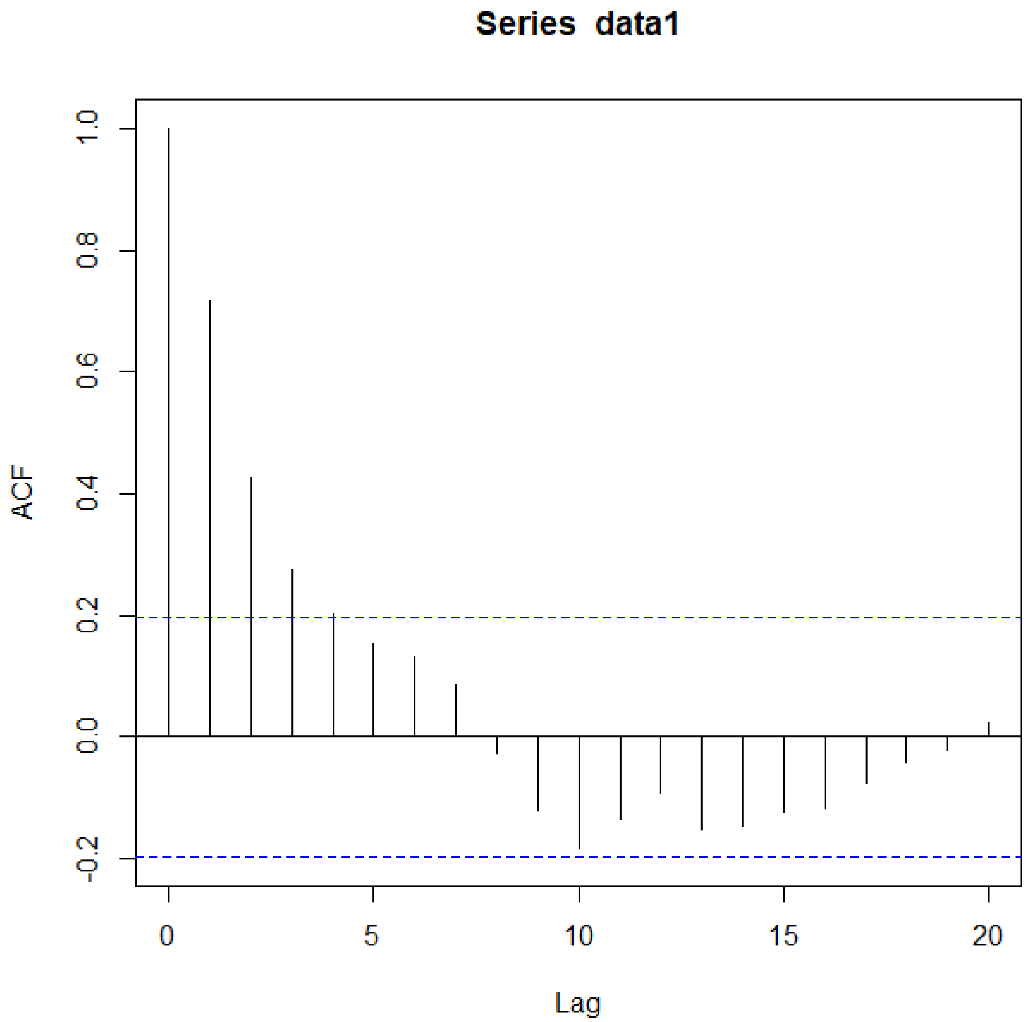
→ **ACF cuts off after lag 1**
→ **PACF tails off**

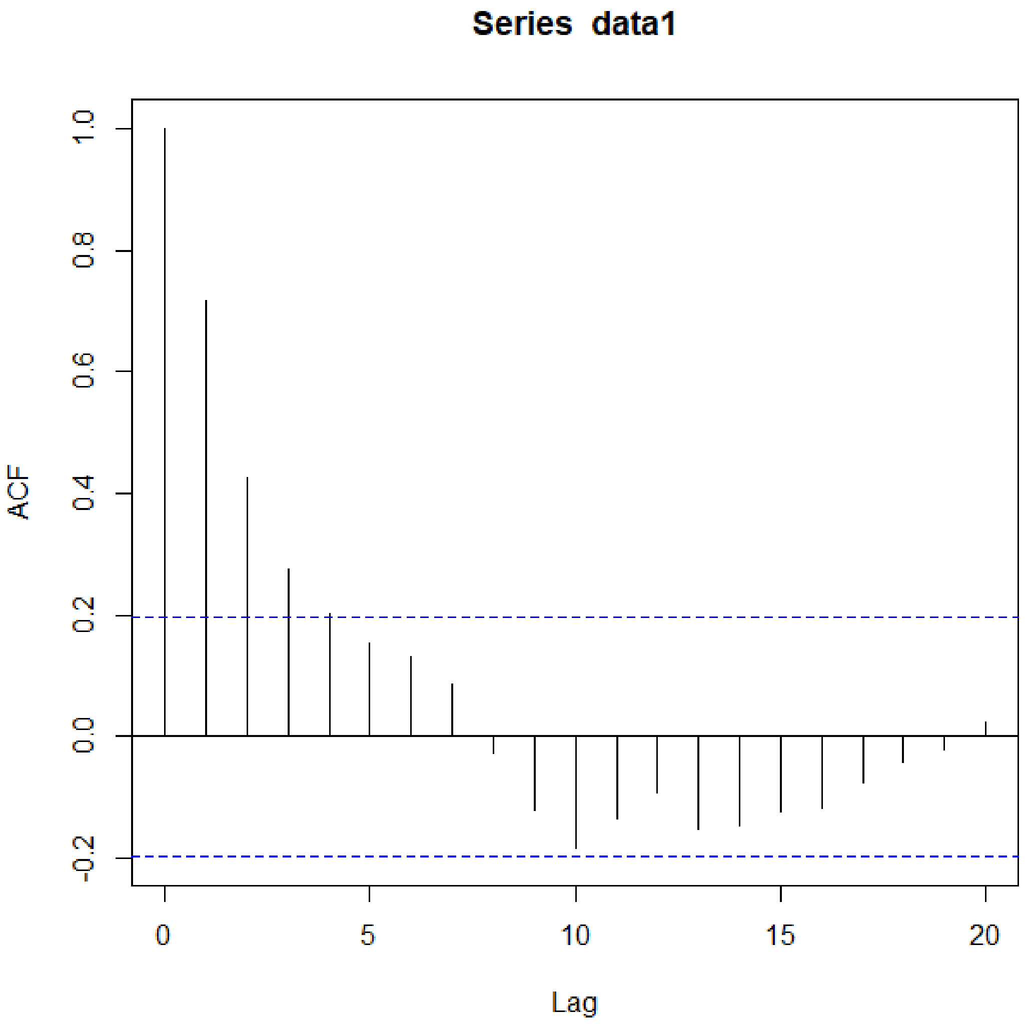
Illustration 3

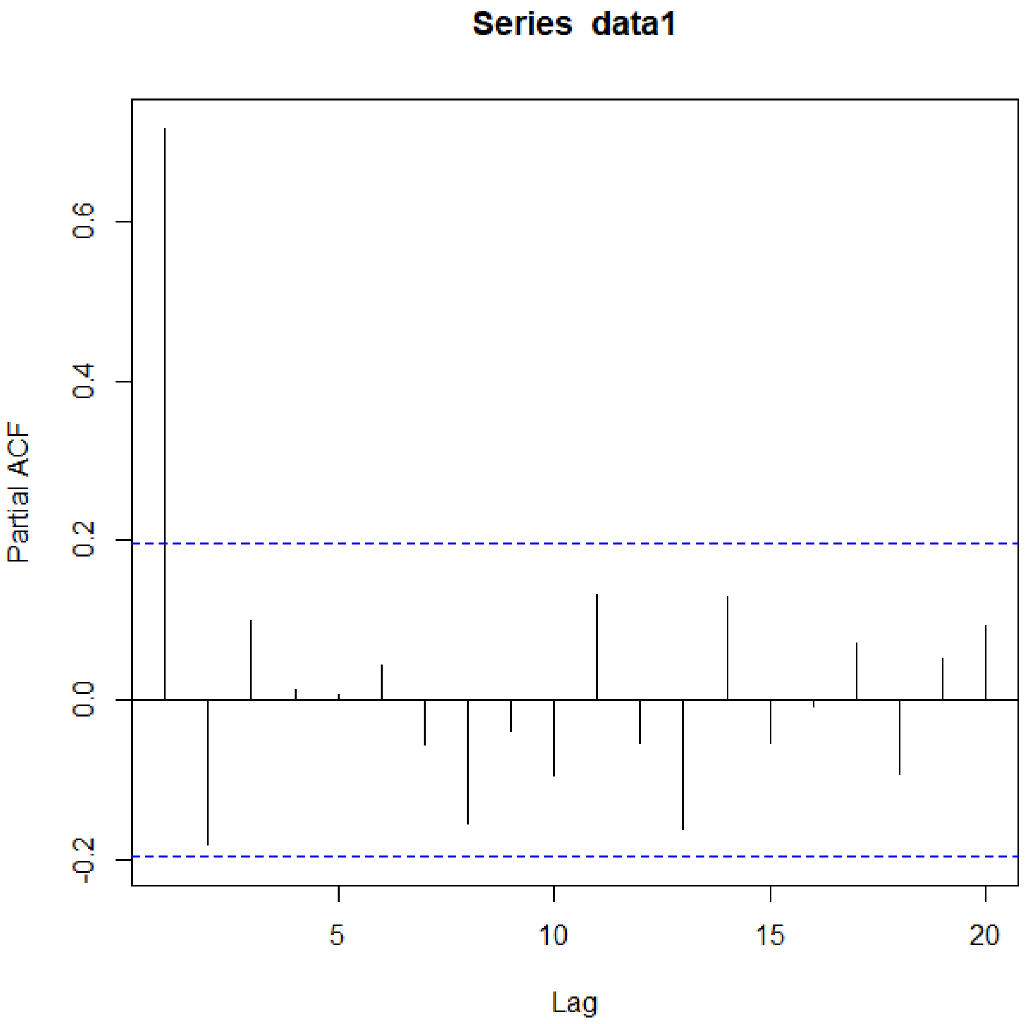


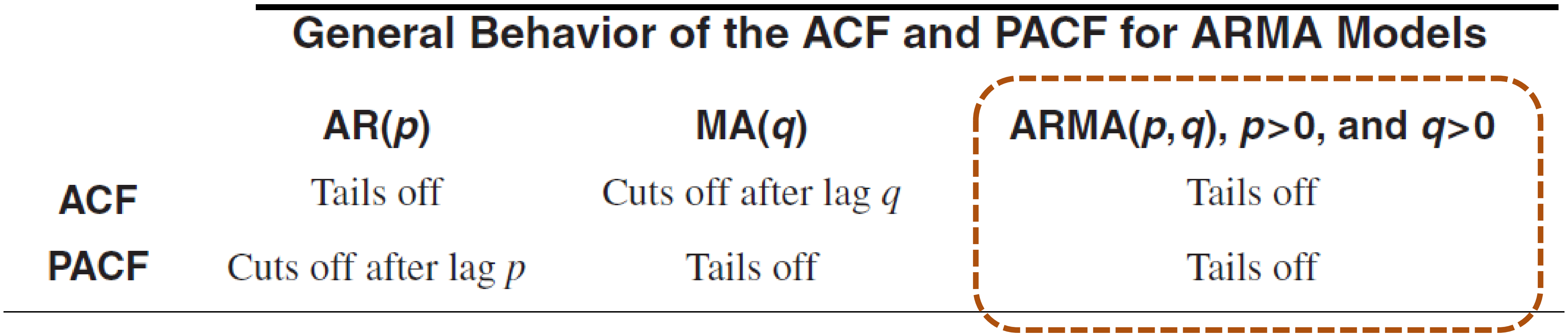
```
Time Series:
Start = 1
End = 100
Frequency = 1
  [1] 49.92942 48.95245 52.22844 53.10579 52.49634 51.72033 51.31350 52.78676
  [9] 51.61496 51.09523 50.41426 50.80596 49.66480 49.24230 50.71685 49.54151
 [17] 49.41205 48.86938 50.08358 51.20780 50.73118 49.40406 47.44811 48.02236
 [25] 48.75842 50.80060 51.89719 51.77418 51.76429 51.58101 50.20945 50.42741
 [33] 50.40793 51.86786 51.95139 51.16796 50.88727 51.36737 50.99544 50.64541
 [41] 51.86159 50.40773 48.60134 48.50495 49.54614 49.50129 49.78693 50.25483
 [49] 51.01725 50.98829 49.64737 48.93415 50.68030 49.94146 47.76121 48.42248
 [57] 48.32285 47.80088 48.27708 48.88886 47.59616 47.68154 48.92910 50.23003
 [65] 51.07214 51.01725 50.05644 49.30724 49.63453 50.91134 51.82234 50.86288
 [73] 50.40121 49.71097 49.34301 49.50765 49.67234 49.26819 48.73062 49.64235
 [81] 48.67032 48.40517 48.84276 48.21299 48.29200 48.72733 49.20141 50.99337
 [89] 50.73869 50.47096 51.00631 50.84143 50.35898 51.40961 50.32619 50.41951
 [97] 50.10471 50.21408 48.64214 50.07742
```


data1
53
52
51
50
49
48
0
20
40
60
80
100
Time


Series data1
ACF
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
5
10
15
20
Lag


Series data1
ACF
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
5
10
15
20
Lag

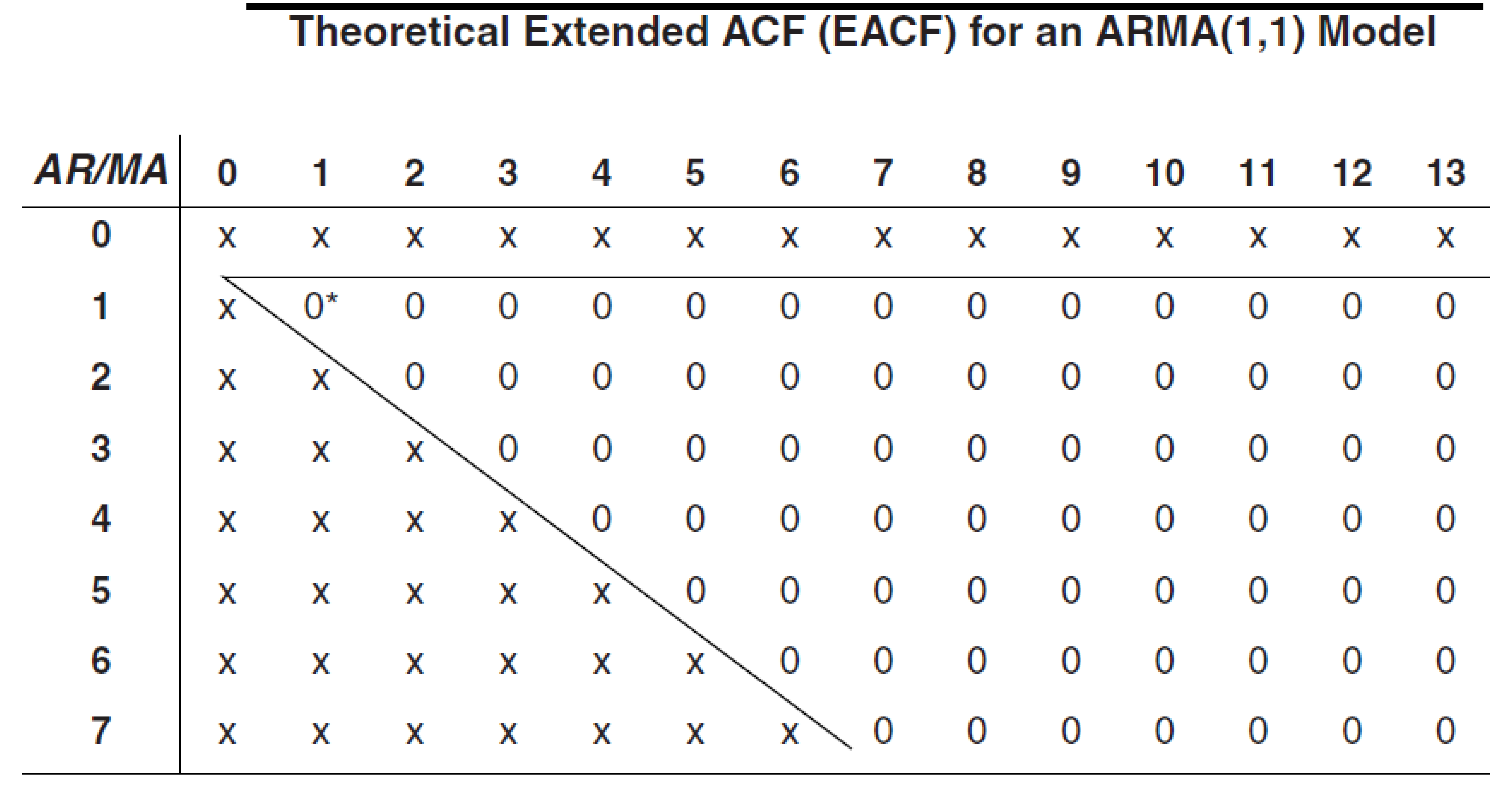

Series data1
Partial ACF
0.6
0.4
0.2
0.0
-0.2
5
10
15
20
Lag



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 Extended Autocorrelation Function

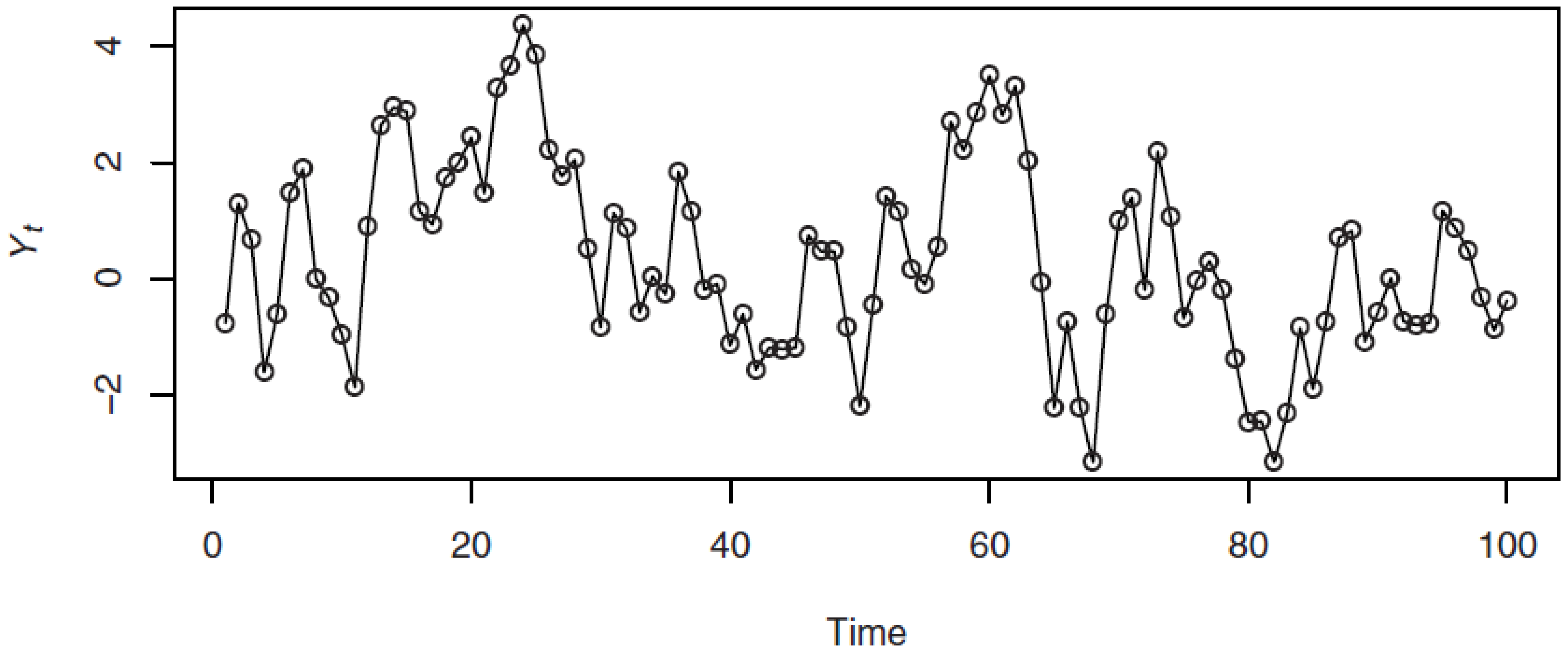


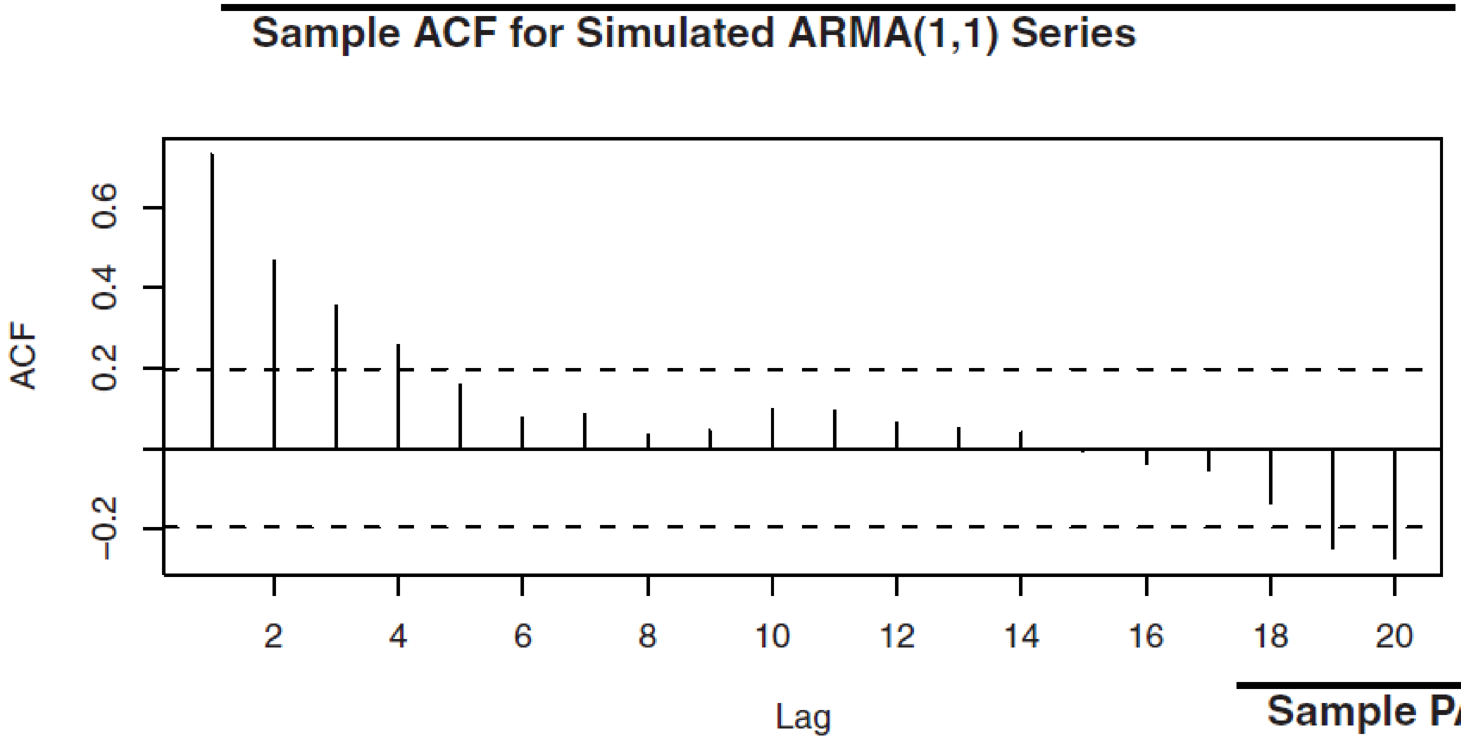
Theoretical Extended ACF (EACF) for an ARMA(1,1) Model

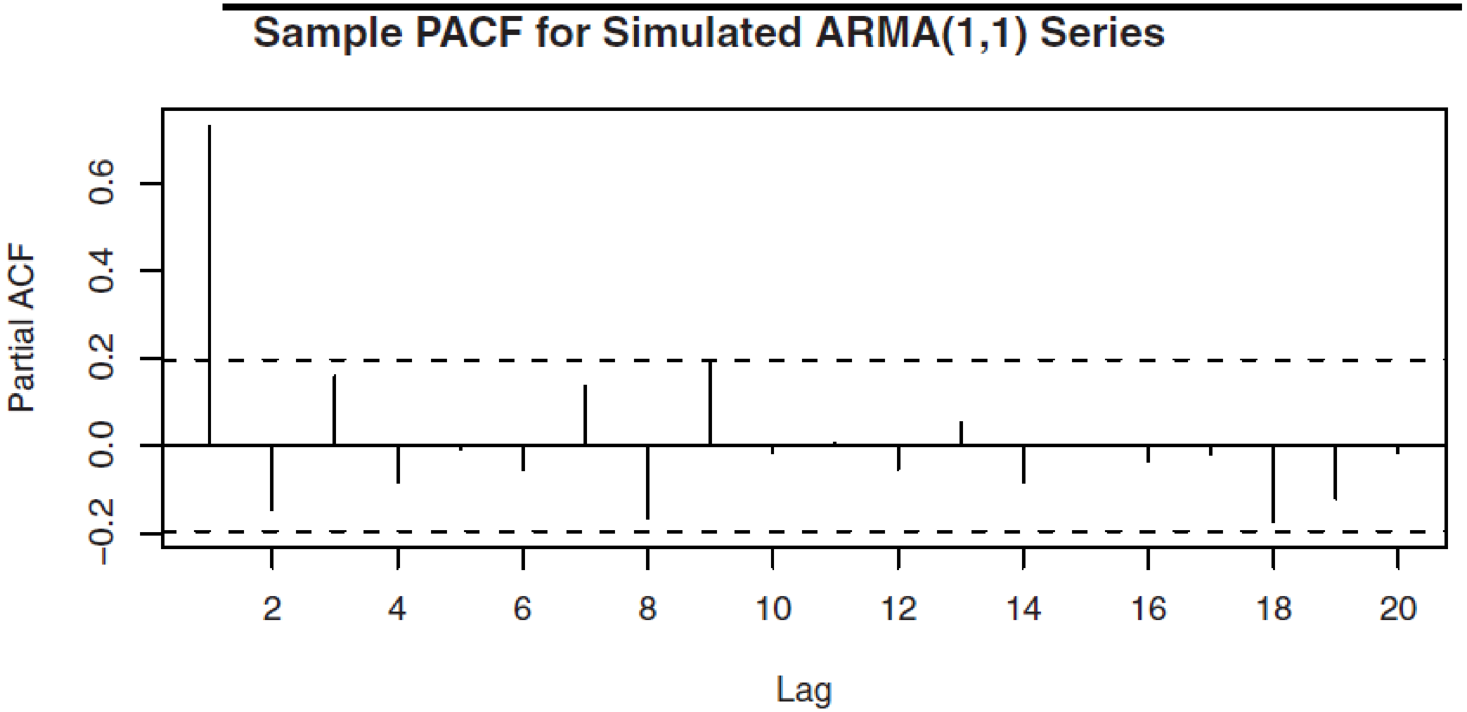
AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1	x	0*	0	0	0	0	0	0	0	0	0	0	0	0
2	x	x	0	0	0	0	0	0	0	0	0	0	0	0
3	x	x	x	0	0	0	0	0	0	0	0	0	0	0
4	x	x	x	x	0	0	0	0	0	0	0	0	0	0
5	x	x	x	x	x	0	0	0	0	0	0	0	0	0
6	x	x	x	x	x	x	0	0	0	0	0	0	0	0
7	x	x	x	x	x	x	x	0	0	0	0	0	0	0

Specification of ARMA model

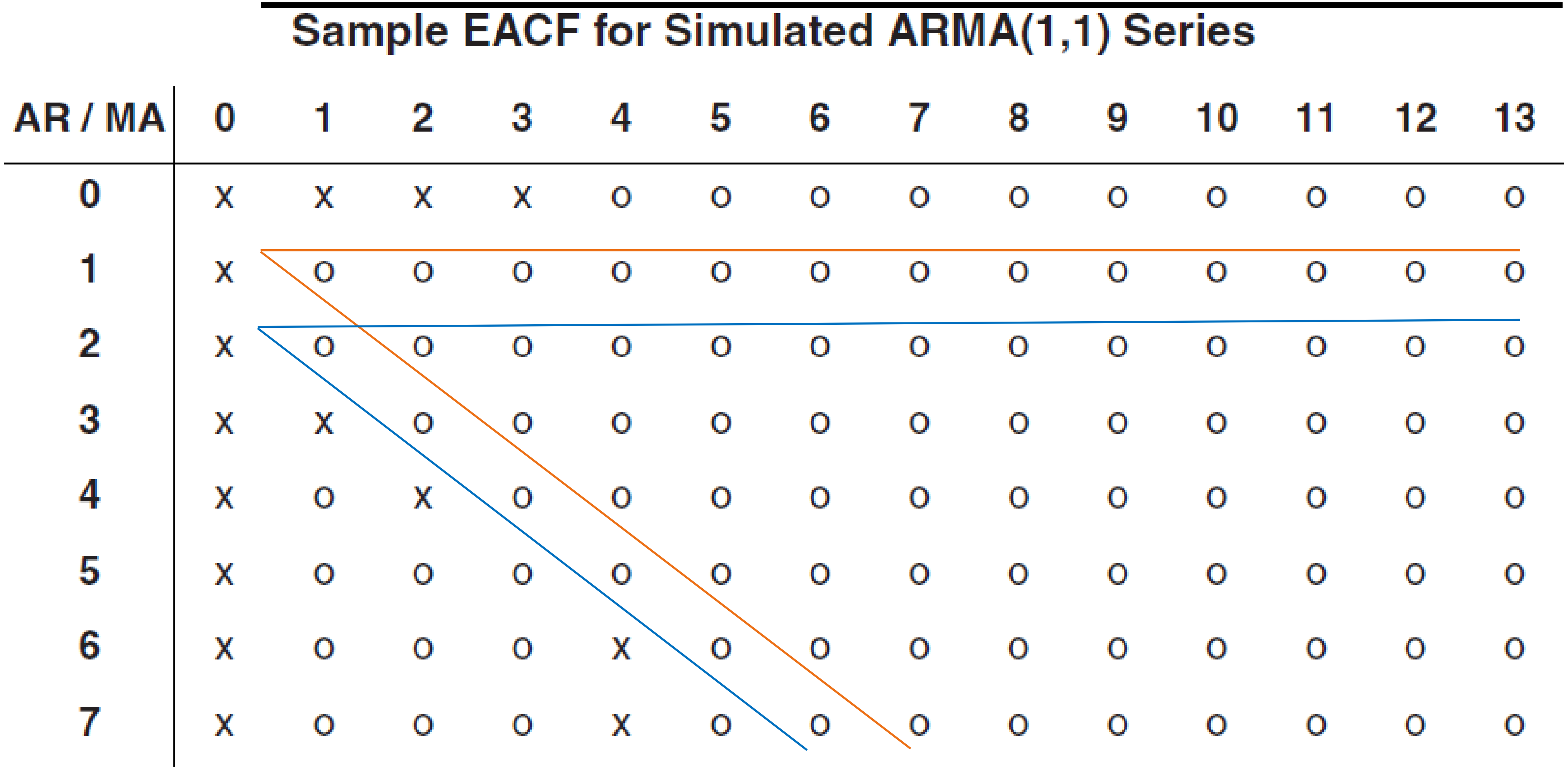
Simulated ARMA(1,1) Series with $\phi = 0.6$ and $\theta = -0.3$.




Sample ACF for Simulated ARMA(1,1) Series


Sample PACF for Simulated ARMA(1,1) Series

→ ACF tails off
→ PACF cuts off after lag 1



Sample EACF for Simulated ARMA(1,1) Series

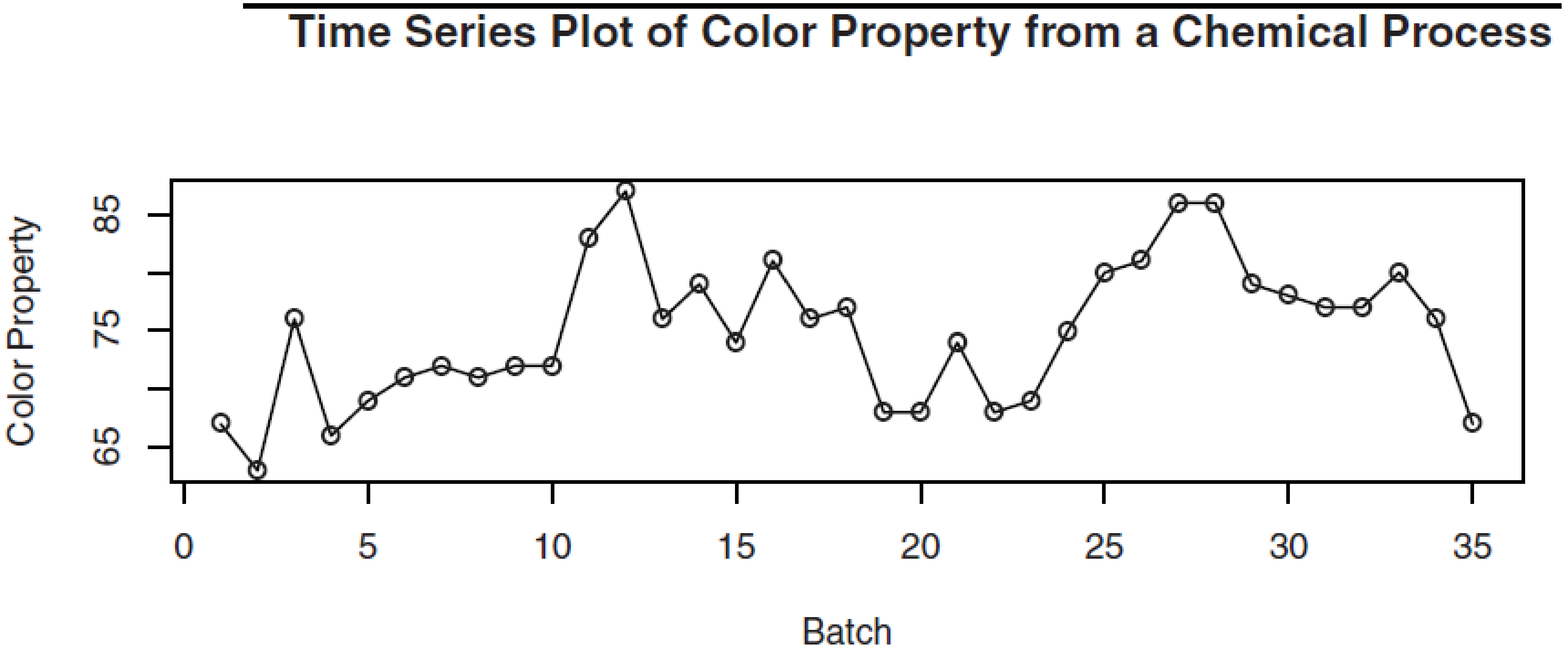
AR / MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	o	o	o	o	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	o	o	o	o	o	o	o	o	o	o	o	o	o
3	x	x	o	o	o	o	o	o	o	o	o	o	o	o
4	x	o	x	o	o	o	o	o	o	o	o	o	o	o
5	x	o	o	o	o	o	o	o	o	o	o	o	o	o
6	x	o	o	o	x	o	o	o	o	o	o	o	o	o
7	x	o	o	o	x	o	o	o	o	o	o	o	o	o

$p = 1, q = 1$

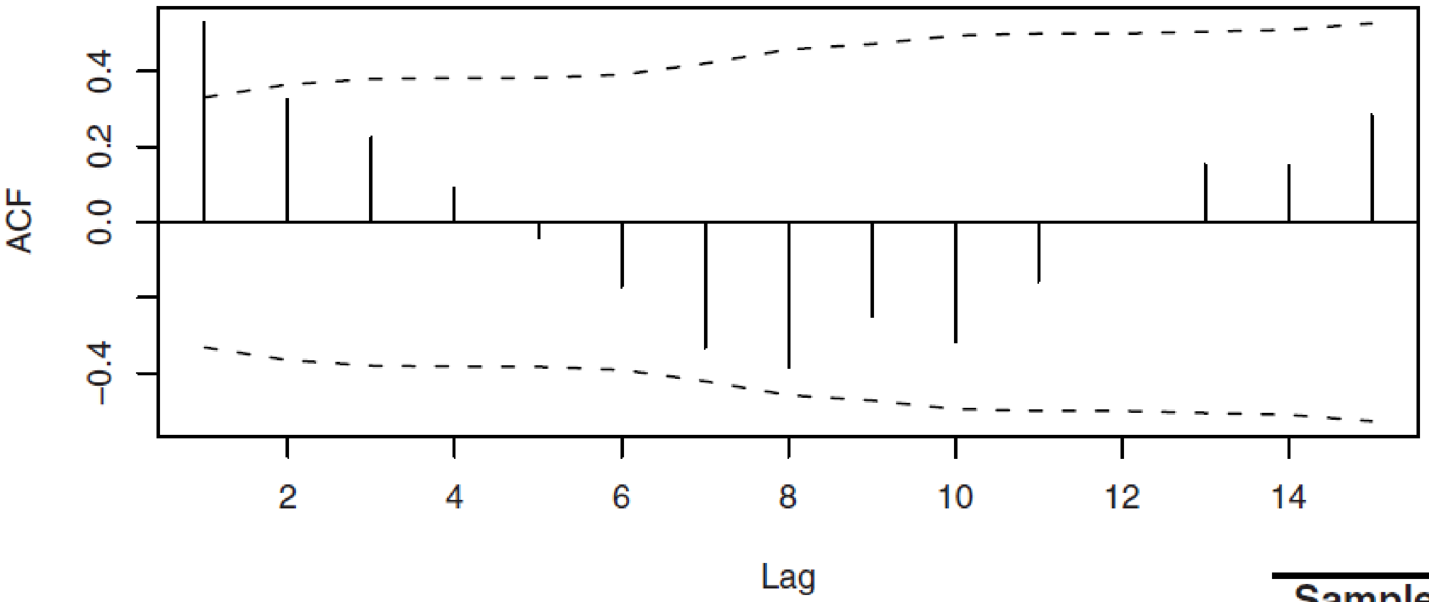
$p = 2, q = 1$

Specification for Some actual time series data

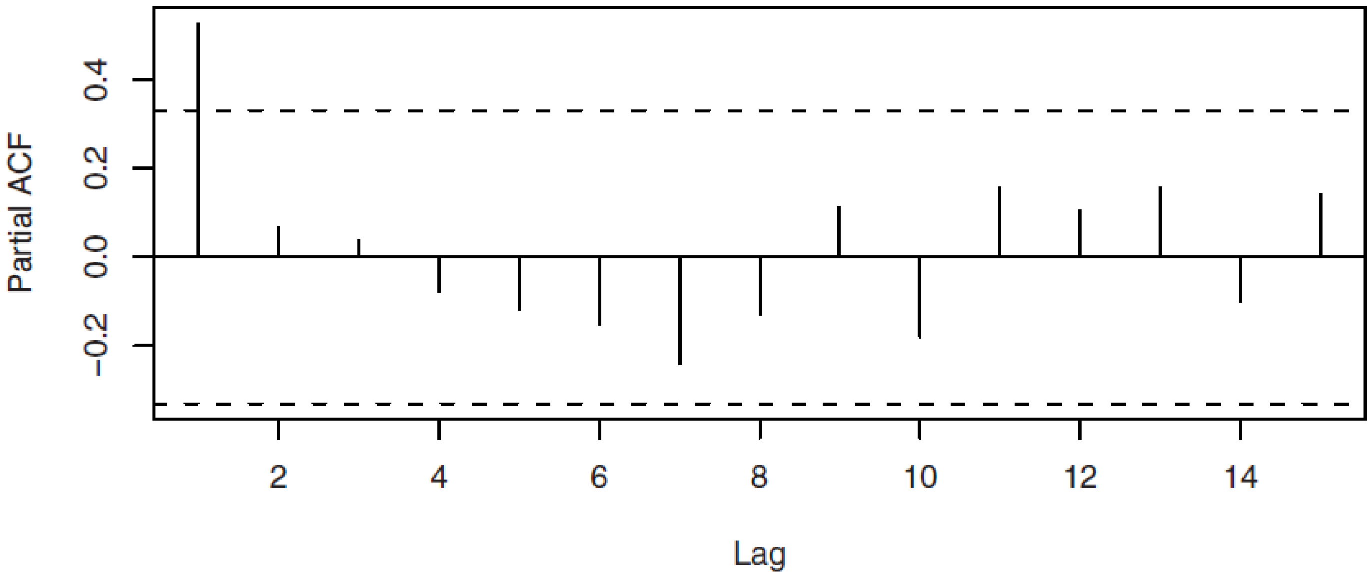
The Chemical Process Color Property Series



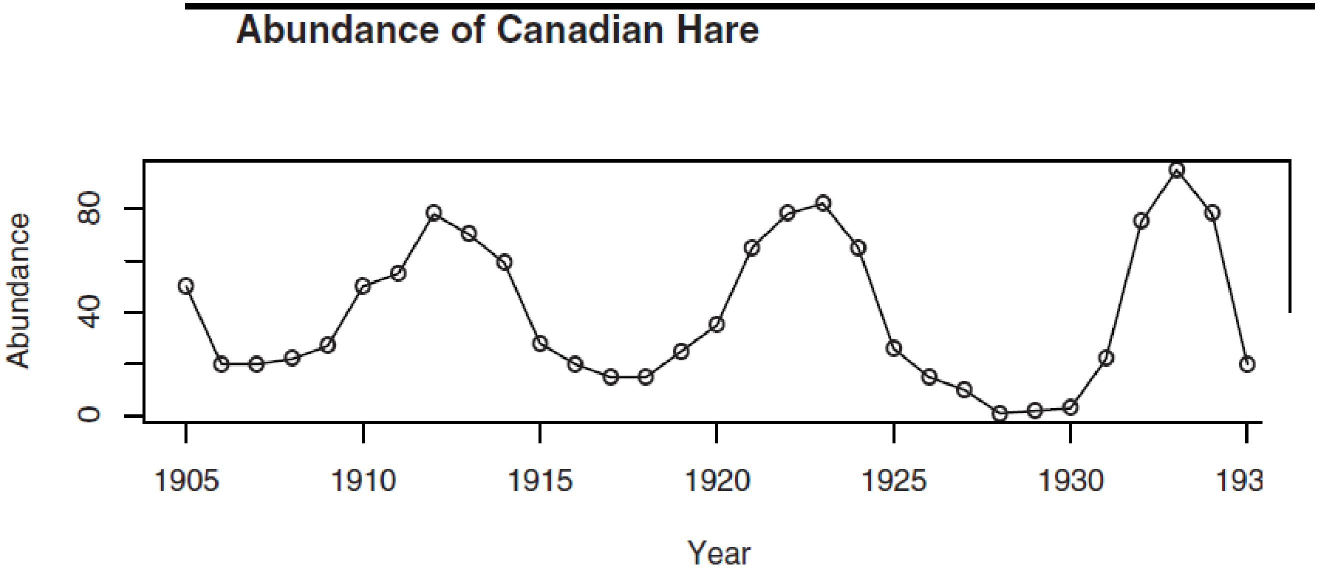
Sample ACF for the Color Property Series



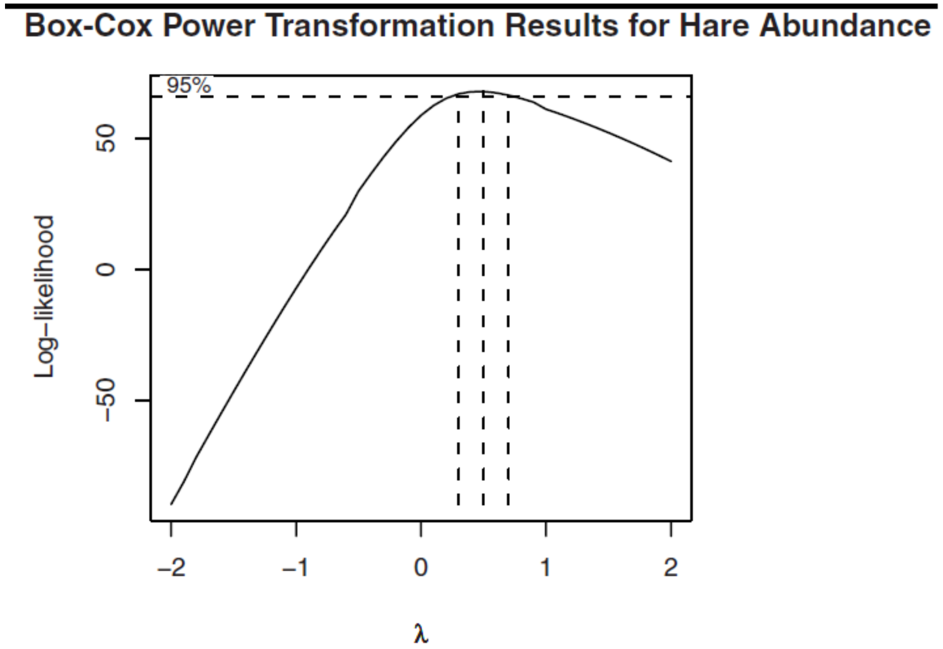
Sample Partial ACF for the Color Property Series



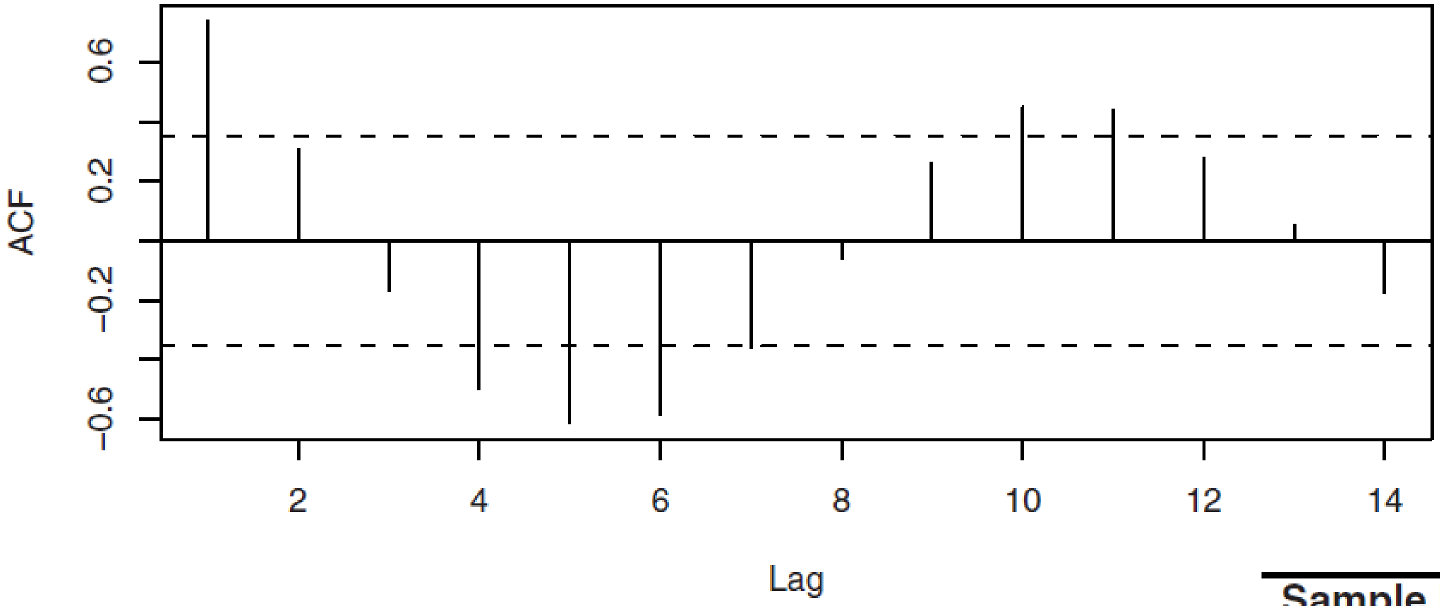
The Annual Abundance of Canadian Hare Series


Abundance of Canadian Hare

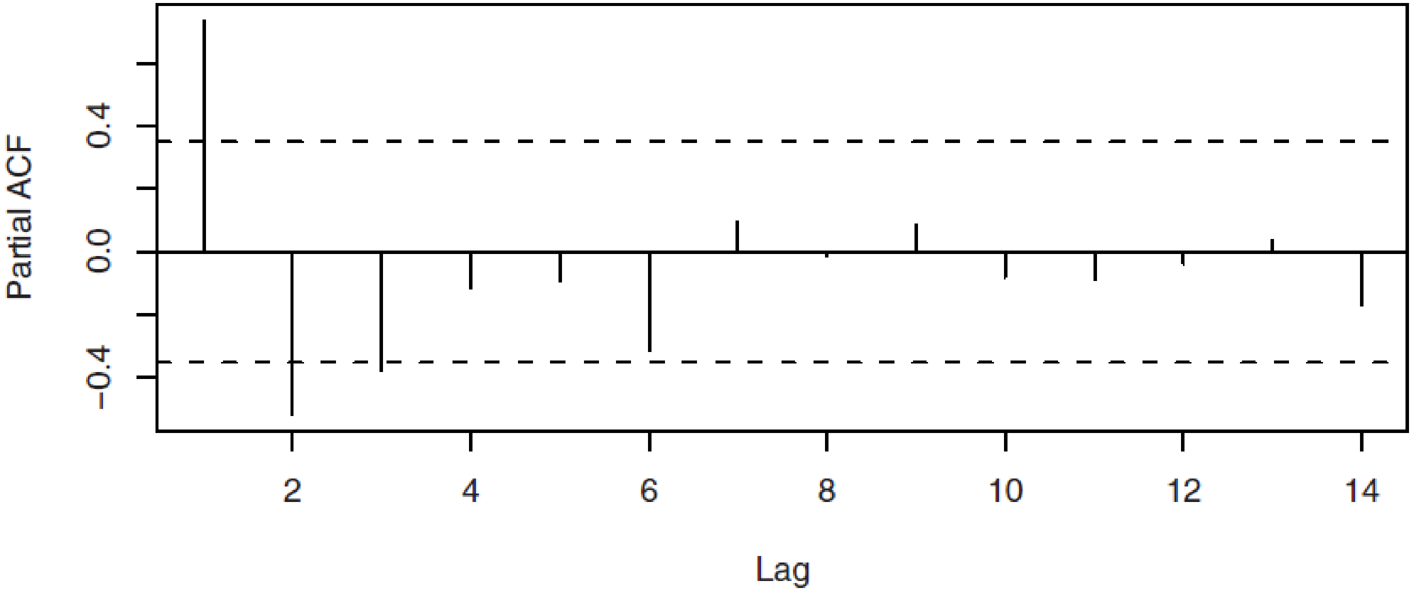
λ	Transformed Data
-2	y^{-2}
-1	y^{-1}
-0.5	1/√y
0	ln(y)
0.5	√y
1	y
2	y^2


Box-Cox Power Transformation Results for Hare Abundance

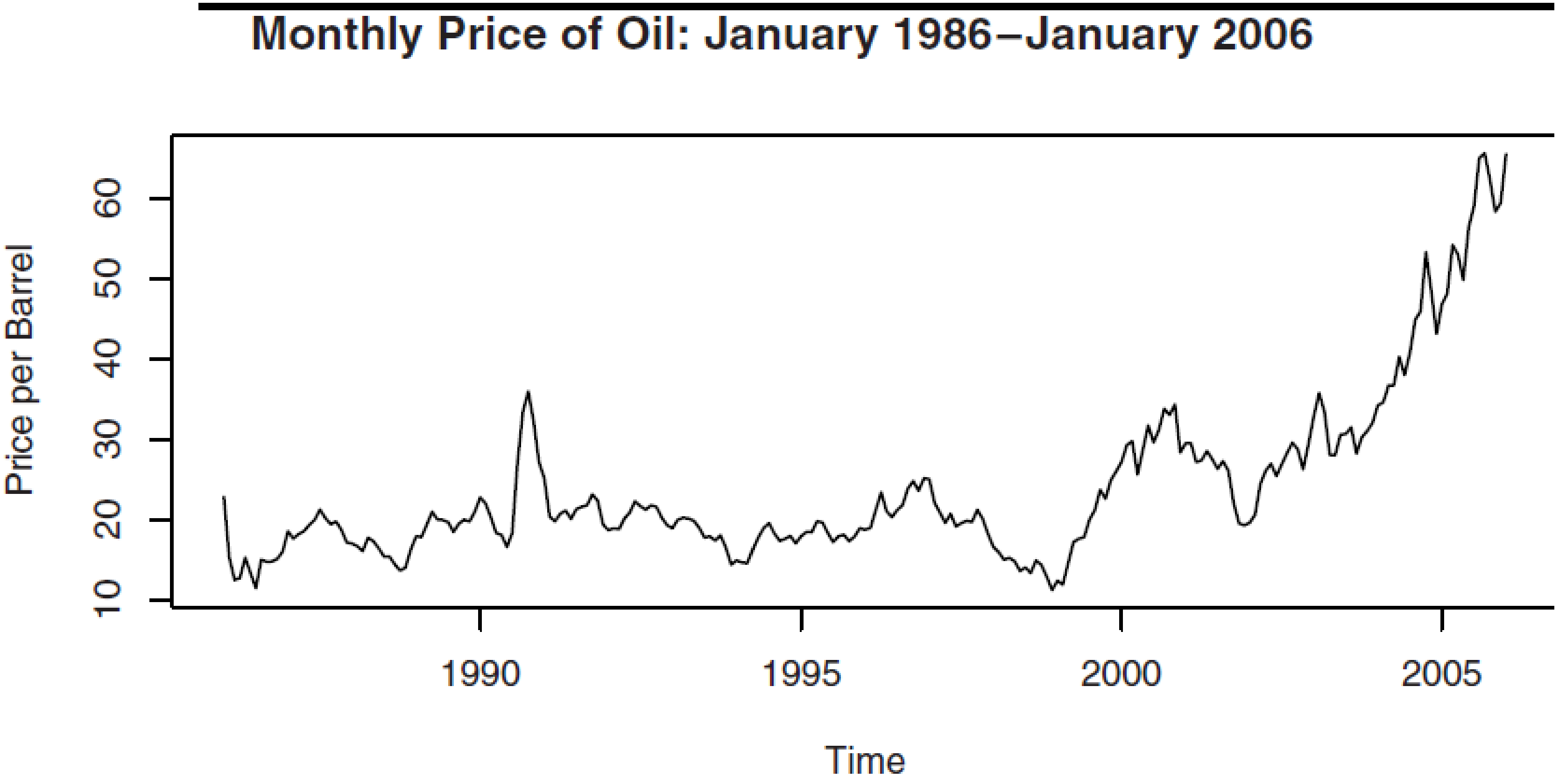
Sample ACF for Square Root of Hare Abundance

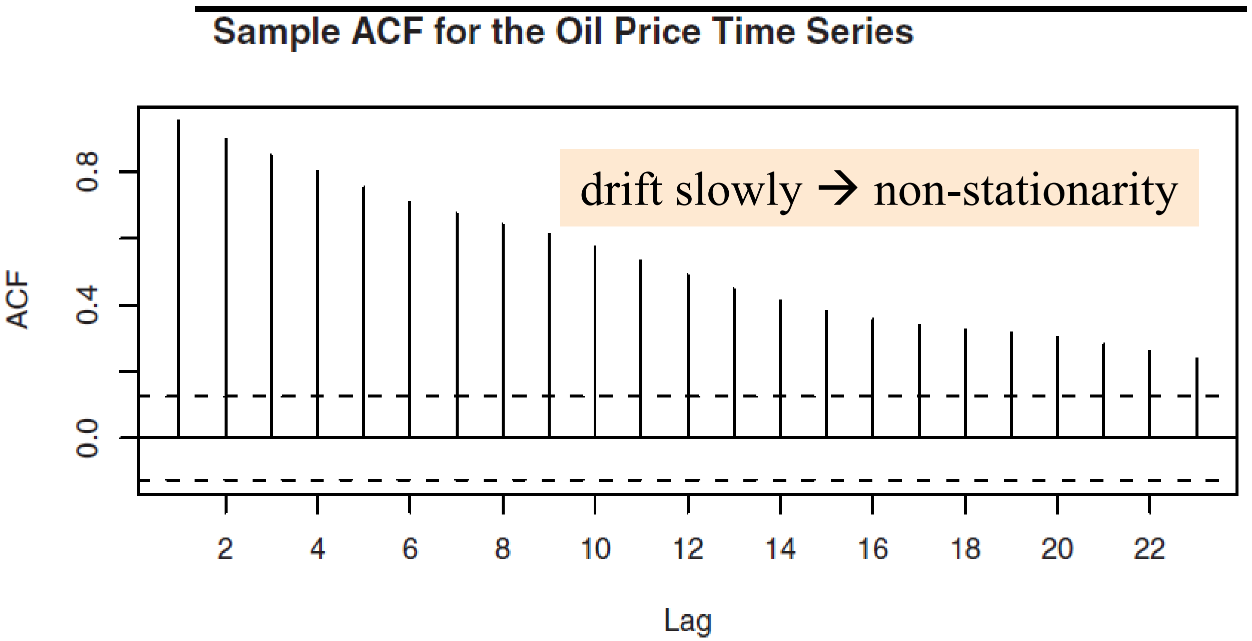


Sample Partial ACF for Square Root of Hare Abundance

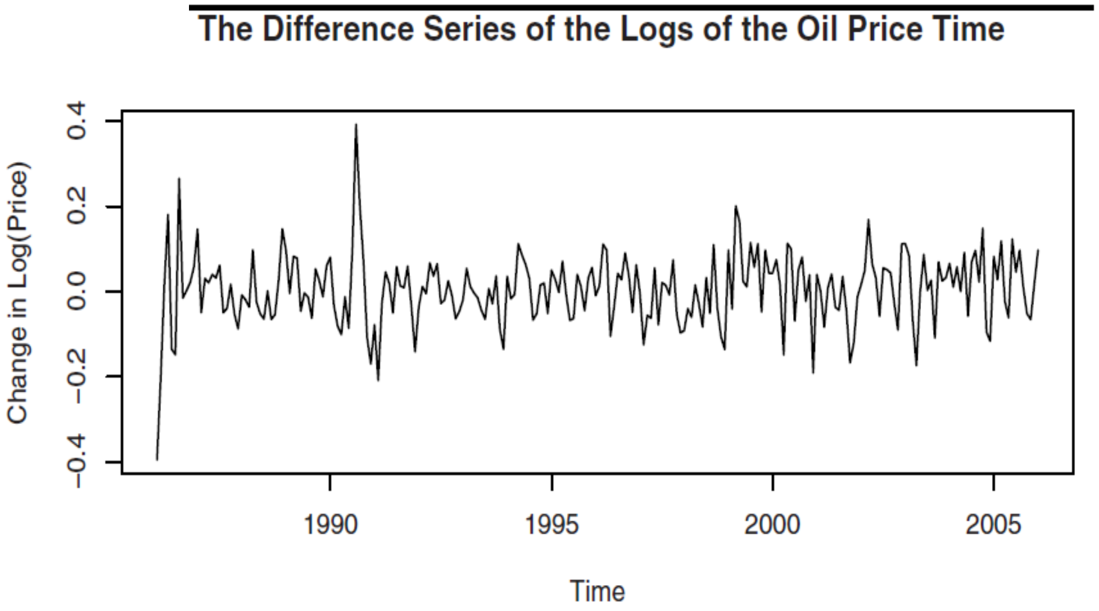


Monthly Price of Oil: January 1986-January 2006

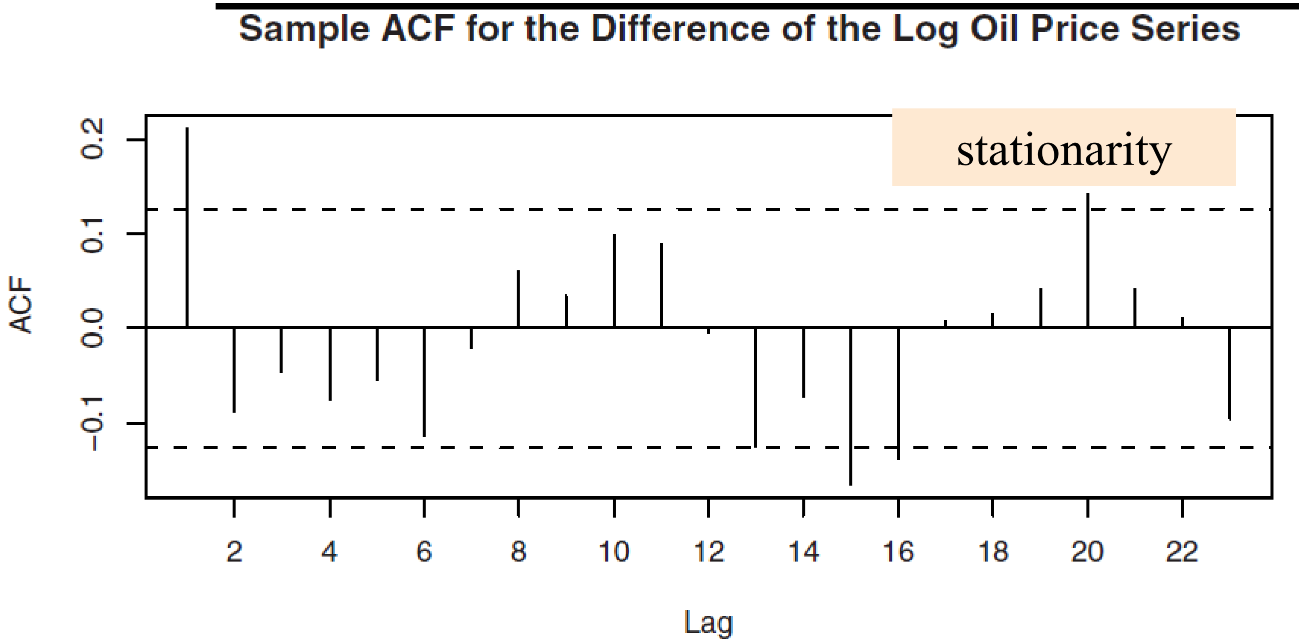




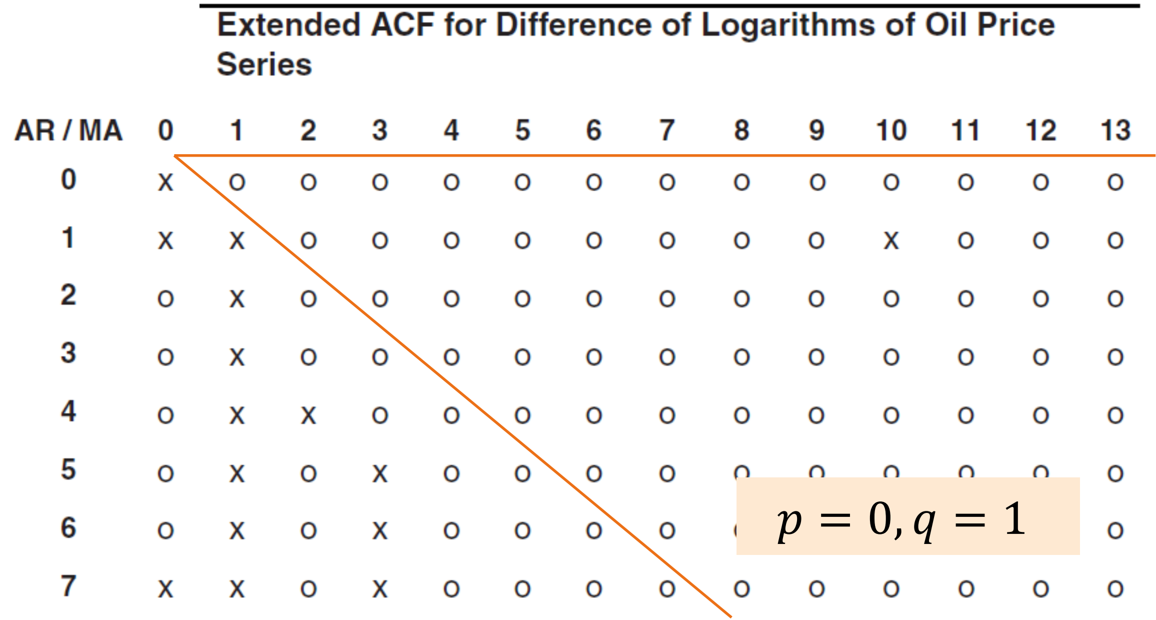
drift slowly → non-stationarity



Monthly Price of Oil → Box-cox transformation using log → differencing ($d = 1$)



stationarity



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	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	o
7	x	x	o	x	o	o	o	o	o	o	o	o	o	o

$p = 0, q = 1$

THANK
YOU