Lecture 3

3.5 Testing for correlation

Definition 5 *We say that a covariance stationary series (X_t) is not serially correlated if and only if $\rho_j = 0$ for all $j > 0$.*

Important example of uncorrelated time series, is i.i.d. (independent indentically distributed) series $\varepsilon_t \; iid(0, \sigma_\varepsilon^2)$.

Definition A simplest example of stationary sequence is a sequence, ε_t, of *independent identically distributed variables* with zero mean and variance σ_ε^2.

Notice that by assumption of independence,

$$Cov(\varepsilon_t, \varepsilon_s) = E(\varepsilon_t - E\varepsilon_t)(\varepsilon_s - E\varepsilon_s) = E[\varepsilon_t \varepsilon_s] = E[\varepsilon_t]E[\varepsilon_s] = 0, \qquad if \quad t \neq s.$$

White noise

Another important second-order stationary process is so-called white noise time series.

Definition 6 *A process, ε_t, is called a white noise if*

$$E[\varepsilon_t] = 0, \qquad E[\varepsilon_t^2] = \sigma_\varepsilon^2$$
$$E[\varepsilon_t \varepsilon_{t-j}] = 0 \quad if \; j \neq 0.$$

White noise time series is zero mean, constant variance, and serially uncorrelated.

- "uncorrelated" implies "independent" only if ε_t has normal distribution.
- We shall use a white noise and an i.i.d. sequences, ε_t, as the building blocks to construct new models of dependent (correlated) time series.

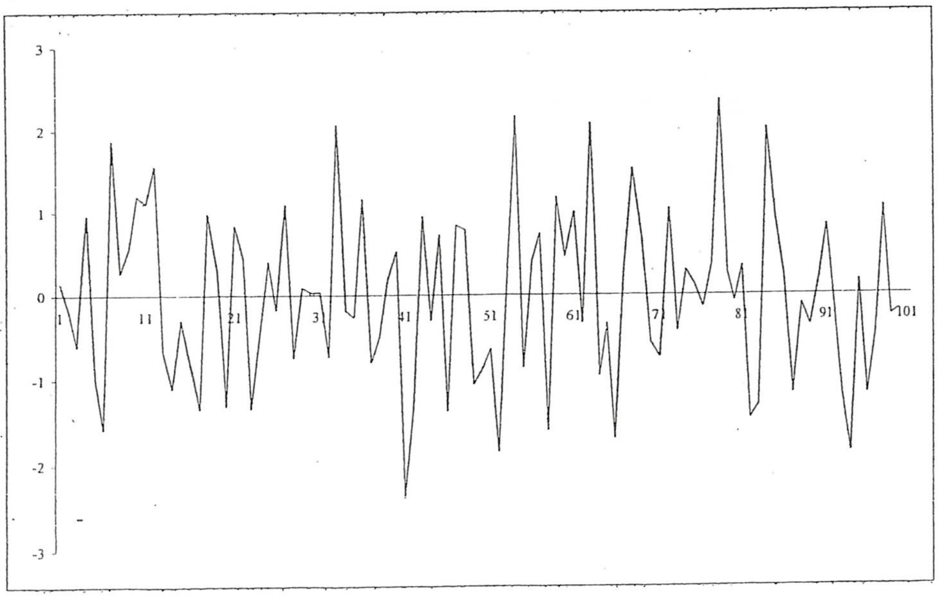


Figure 1.4: Simulated data from a white noise sequence with mean 0 and variance 1.

Testing for correlation.

Why? If time series variables are uncorrelated, then there is no structure in the data, and we do not fit a time series model. If variable are correlated, then we can fit a model, and use it for forecasting.

Fact. If X_t are i.i.d. (independent identically distributed) variables then the statistic

$$t = \sqrt{N}\widehat{\rho}_1 \sim N(0,1);$$
$$t = \sqrt{N}\widehat{\rho}_k \sim N(0,1) \quad \text{for any fixed } k \geq 1$$

have asymptotical standard normal distribution.

To test the null-hypothesis

$$H_0 : \rho_1 = 0 \text{ against alternative } H_1 : \rho_1 \neq 0$$

we can use the rule similar as in testing for skewness:

Rule: reject H_0 at significance level 5% if

$$|t| > 2, \quad or \quad |\hat{\rho}_1| > \frac{2}{\sqrt{n}}.$$

The same rule applies for any lag $k \geq 1$:

Reject

$$H_0 : \rho_k = 0 \text{ against alternative } H_1 : \rho_k \neq 0$$

if

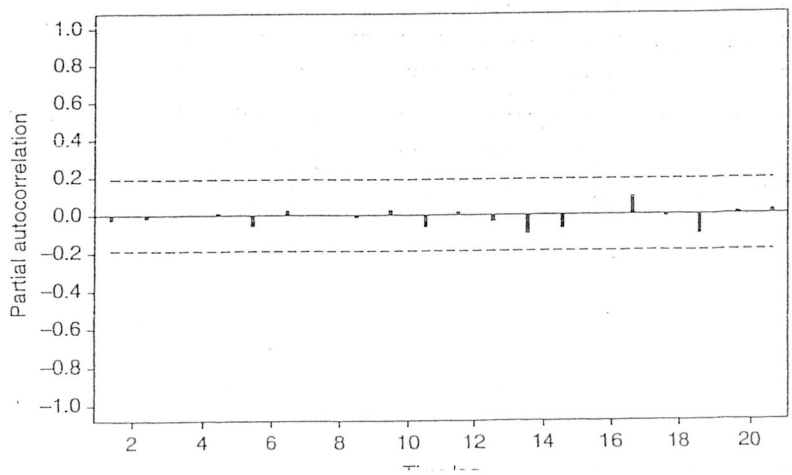
$$|\hat{\rho}_k| > \frac{2}{\sqrt{n}}.$$

Note: plotting sample autocorrelation in e-views, it will be give the 95% confidence band for ρ_k:

$$[-\frac{2}{\sqrt{n}}, \frac{2}{\sqrt{n}}].$$

Then

- if for lag k, sample correlation $\hat{\rho}_k$ is outside the band, then correlation at lag k is significant: $\rho_k \neq 0$.
- If $\hat{\rho}_k$ is inside the band, then correlation at lag k is not significant, that is $\rho_k = 0$.



Ljung-Box test for serial correlation. The Ljung-Box statistic can be used to test for correlation not at one lag, but at few lags simultaneously. We choose the number m of lags, and test the null hypothesis:

$$H_0 : \quad \rho_1 = \rho_2 = \cdots = \rho_m = 0$$

against alternative

$$H_1 : \quad \rho_j \neq 0 \quad \text{for some } 1 \leq j \leq m.$$

Test uses Ljung-Box statistic

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

where N is the number of observations, and $\hat{\rho}_k$ is the sample ACF at the lag k. Under H_0 it will have not-normal distribution.

E-views will give p-value.

We reject hypothesis H_0 at significance level 5% if the p-value is less than 0.05.

Note: What m should we use? If we choose m to large, the test will have low ability to detect that H_0 is not true.

- Use: $m = \sqrt{N}$ where N is the number of observations.
- E-views will call this test Q test and give $p-$ values for all $m = 1, 2, \cdots$. Take a look at $m \leq \sqrt{N}$ to see if there is any p value less than 0.05. If you find such you may reject hypothesis of no serial correlation, i.e. of white noise.

Example of application of ACF

Sample autocorrelations $\hat{\rho}_1, \hat{\rho}_2, \ldots$ play important role in linear time series analysis. They can capture the linear dynamic of the data.

Figure 2.1 shows the sample ACF of monthly simple and log-returns of IBM stock from January 1926 to 1997. We observe that:

- two sample ACF are very close to each other,
- they show that serial correlation of IBM stock returns are very small, practically zero.
- the ACF's are within two standard-error limits, indicating that they are not significantly different from 0 at 5% level.

Figure 2.2 shows ACF's of monthly returns of the value-weighted index of US markets. There are some significant correlations at 5% level for both return series.

Comment: Testing for zero correlations has been used in practice as a tool to check the efficient market assumption, which means that series of returns should be uncorrelated.

However, the way how stock prices are determined and index returns are calculated might introduce some autocorrelations in observed return series.

In practice, if all sample ACF's are close to zero, then the series is a white noise. Based on Figures 2.1 and 2.2, the monthly returns of IBM stock are close to the white noise, but those of the value-weighted index are not.

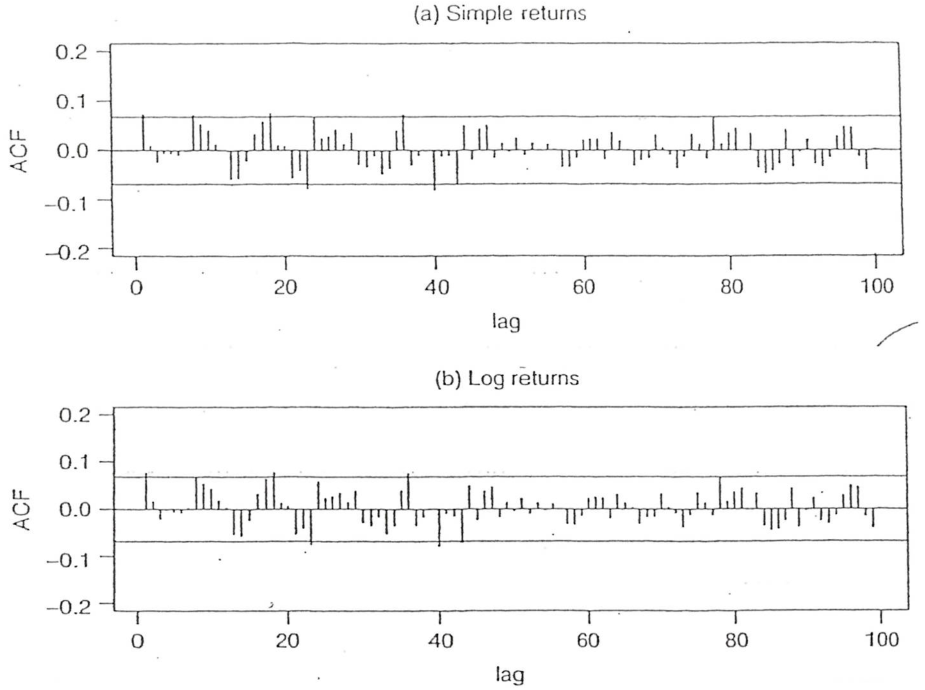


Figure 2.1. Sample autocorrelation functions of monthly (a) simple returns and (b) log returns of IBM stock from January 1926 to December 1997. In each plot, the two horizontal lines denote two standard-error limits of the sample ACF.

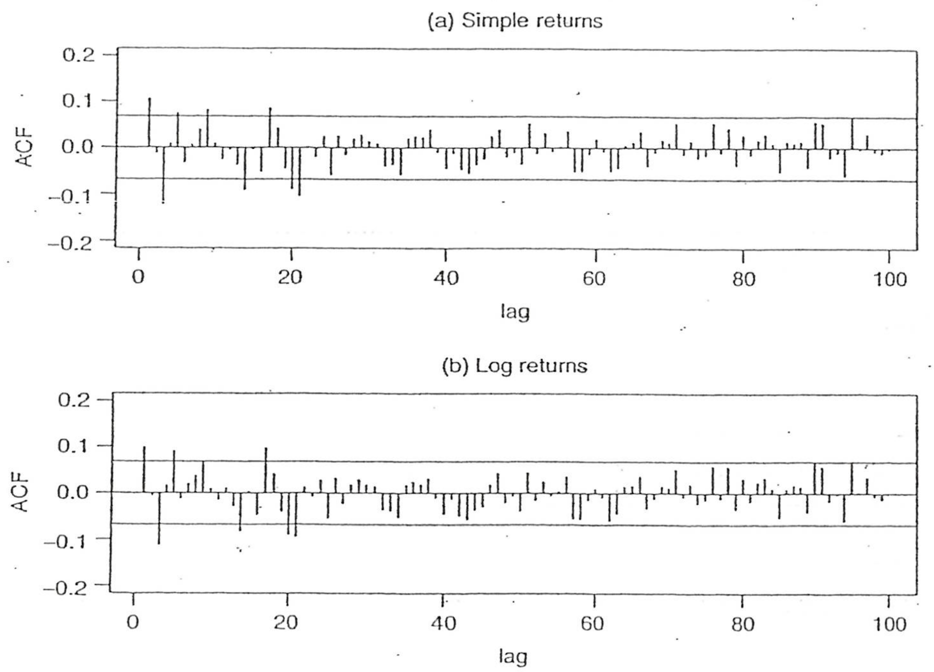


Figure 2.2. Sample autocorrelation functions of monthly (a) simple returns and (b) log returns of the value-weighted index of U.S. markets from January 1926 to December 1997. In each plot, the two horizontal lines denote two standard-error limits of the sample ACF.

Autoregressive AR(p) model

Now we consider what to do when the time series has significant correlations, that is it is not white noise.

To autocorrelated time series data can be try to fit a linear autoregressive moving average model AR(p).

Q: Should such model fit to the date? Not necessarily. We will discuss how to check if the model is fitting well.

If it does not fit well, then we can try to fit MA (moving average) or ARMA models, we will discuss below.

What is autoregressive model AR(p) of order p?

- In autoregressive model AR(1) of order 1:

  $$X_t = \mu + \phi X_{t-1} + \varepsilon_t.$$

  In this model $\{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2)$ is assumed to be a white noise. The model has 3 parameters: μ, ϕ and σ_ε^2.

- In autoregressive model AR(2) of order 2:

  $$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t.$$

  We regress X_t on the past two values X_{t-1} and X_{t-2}. What remains is the white noise $\{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2)$. The model is defined by parameters: μ, ϕ_1, ϕ_2 and σ_ε^2.

- In autoregressive model AR(p) of order p:

  $$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots \phi_p X_{t-p} + \varepsilon_t.$$

  Here we regress X_t on the past p values $X_{t-1}, \cdots, X_{t-p}$. What remains is the white noise $\{\varepsilon_t\} \sim WN(0, \sigma_\varepsilon^2)$.

3.6 Stationarity of AR(1) model

The AR(1) model is

$$X_t = \mu + \phi X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim WN(0, \sigma_u^2).$$

It has a stationary solution if $|\phi| < 1$.

If Figure 3.10 we see four realizations of AR(1) model with $\mu = 0$:

- $\phi = 0$, then $X_t = \varepsilon_t$ is white noise.
- $\phi = 0.75$: series exhibits short sequences of up and down but always returns back to equilibrium.
- $\phi = 1$: is unit root model which is nonstationary
- $\phi > 1.25$ is explosive model which is non-stationary

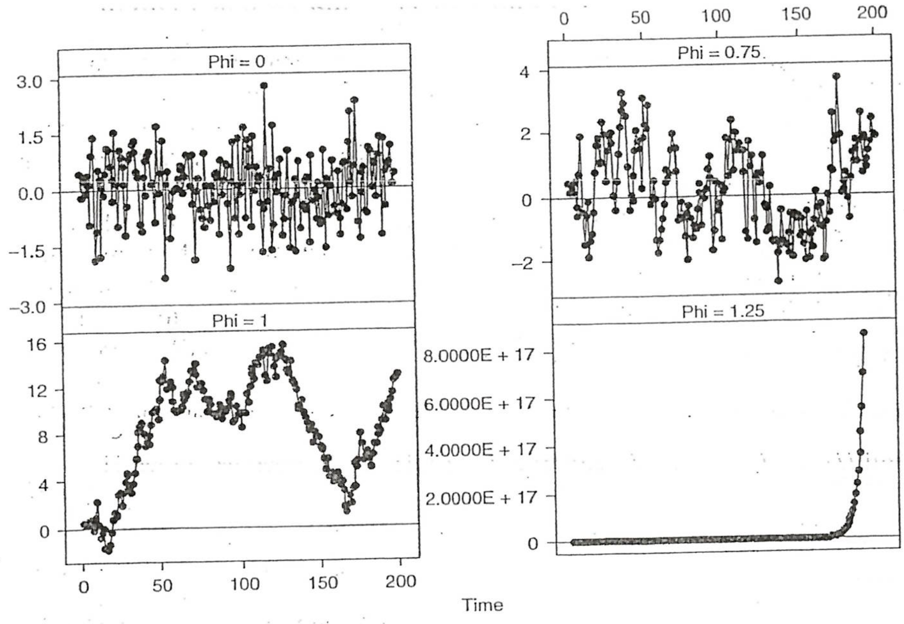


Figure 3.10 Four realizations of the AR(1) process with $\phi = 0, 0.75, 1$ and 1.25.

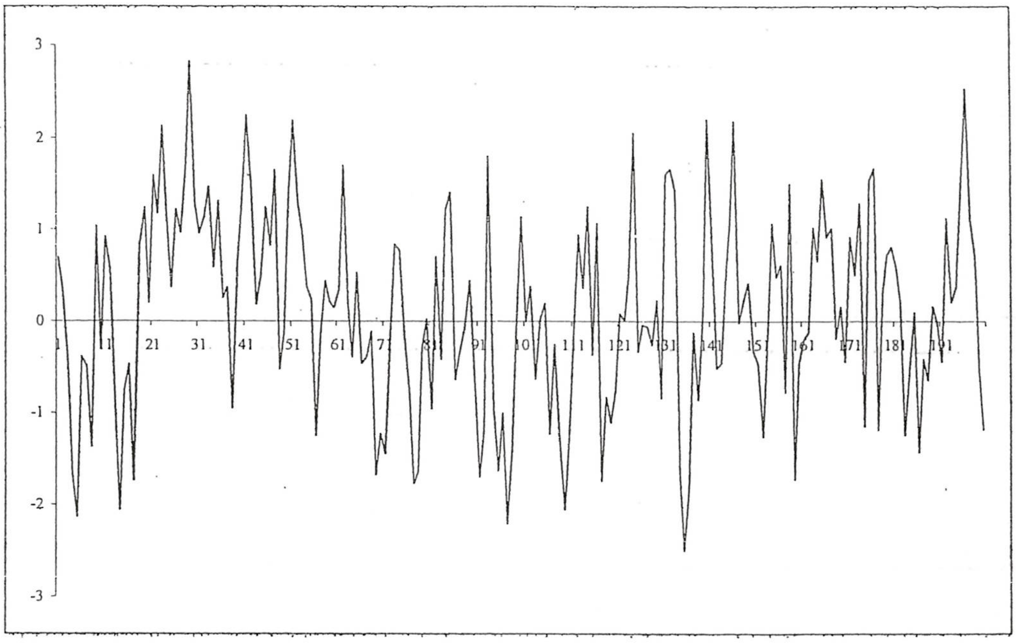


Figure 2.1: Simulated data of an $AR(1)$ model with $a_0 = 0$, $a_1 = 0.5$ and $\sigma_\varepsilon^2 = 1$.

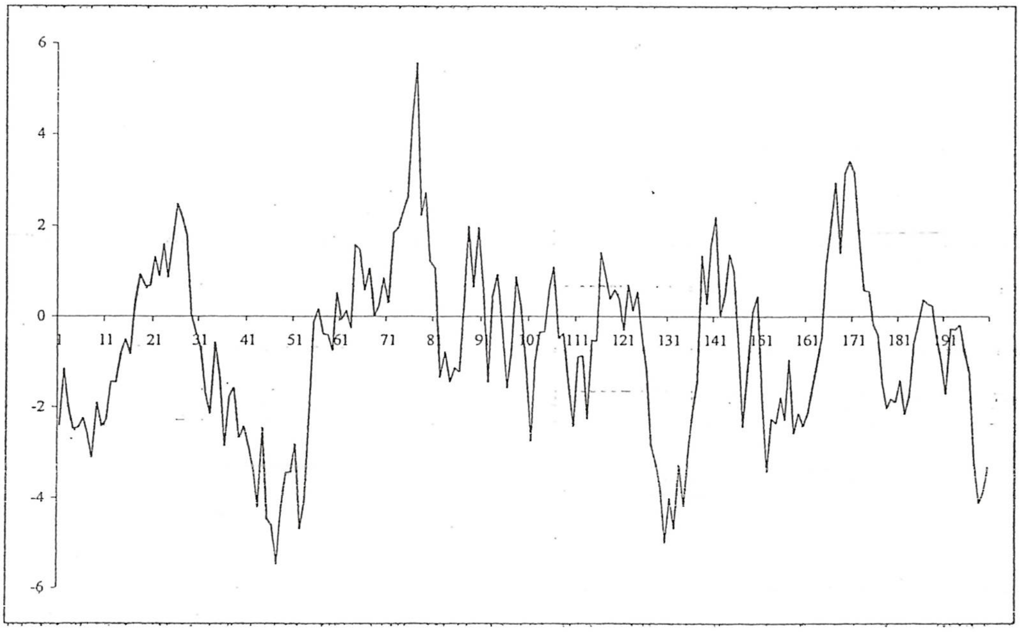


Figure 2.2: Simulated data of an $AR(1)$ model with $a_0 = 0$, $a_1 = 0.9$ and $\sigma_\varepsilon^2 = 1$.

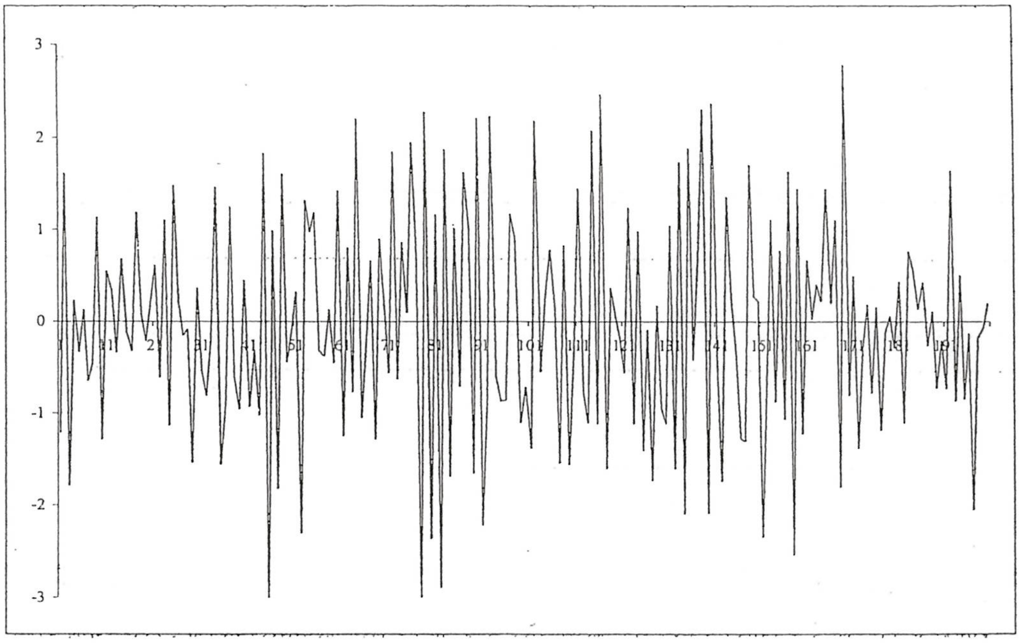


Figure 2.3: Simulated data of an $AR(1)$ model with $a_0 = 0$, $a_1 = -0.5$ and $\sigma_\varepsilon^2 = 1$.

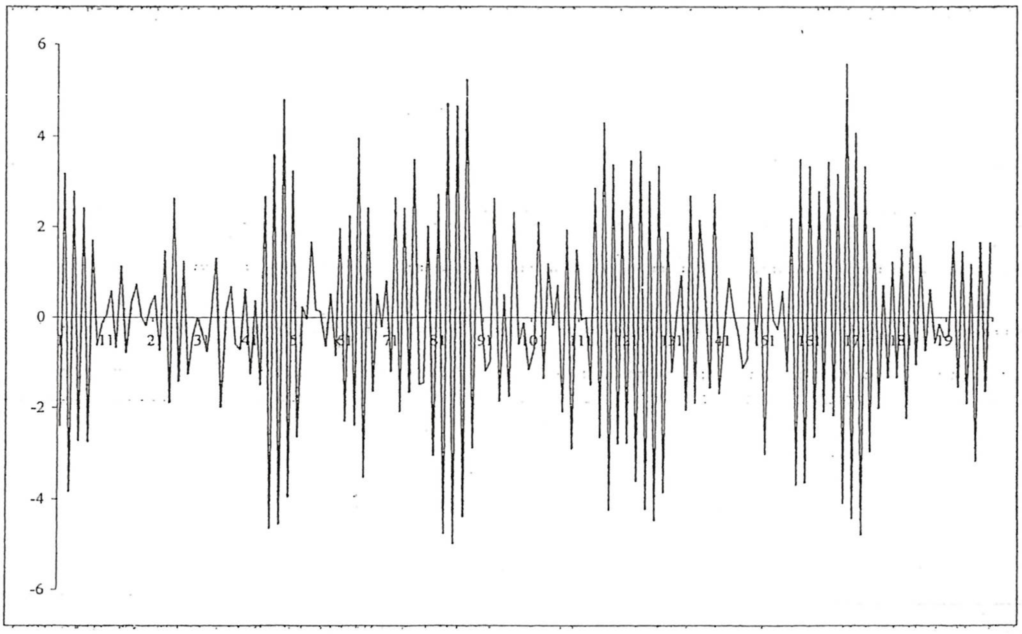


Figure 2.4: Simulated data of an $AR(1)$ model with $a_0 = 0$, $a_1 = 0.5$ and $\sigma_\varepsilon^2 = 1$.

Causal time series: we say that a stationary time series X_t is causal, if it does not depend on future shocks. That is easy to see for AR(1) model:

$$\begin{aligned} X_t &= \phi X_{t-1} + \varepsilon_t \\ &= \phi(\phi X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \phi^2 X_{t-2} + \phi \varepsilon_{t-1} + \varepsilon_t. \end{aligned}$$

We can show continuing as above and replacing X_{t-2} by X_{t-3} and so on that

$$X_t = \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \phi^3\varepsilon_{t-3} + \cdots$$

is a linear combination of past shocks. Since $|\phi| < 1$, this series is converging, and therefore is causal.

Later we compute EX_t, $Var(X_t)$ and autocorrelation function ρ_k of a stationary AR(1) model.

Identification of the order

The identification which ARMA model to use often can be determined by looking at the ACF and PACF (partial autocorrelation function).

Selection of order (p) when we fit $AR(p)$ model. Assume that we want to fit to the data AR(p) model. Then first we need to select p. For that we use PACF function.

Note: Partial autocorrelation function PACF is computed at lags $k = 1, 2, \cdots,$. It has nothing to do with correlation (ACF). The only its use is to determine the order of AR(p) model.

Meaning of PACF: The PACF at lag k is the last regression coefficient ϕ_{kk} when we fit regression equations for $k = 0, 1, 2, \cdots$:

$$X_t = \mu + \phi_{k1}X_{t-1} + \cdots \phi_{kk}X_{t-k} + \varepsilon_t.$$

Assume that data was generated by AR(p) model. When we fit too many lags, then PACF ϕ_{kk} will be approximately 0. The last non-zero PACF will be at lag "p" which is the order of AR(p) model.
In other words: PACF cuts off after lag p.

For example, if we fit an AR(k) model to the data that truly follows AR(2) model, the PACF coefficients at lags $3, 4, 5, \ldots$ will be zero.

Since we can compute only sample PACF, we need to check at which lag it cuts to zero, i.e. becomes not significantly different from zero.

The rule: the same as in testing for correlation at lag k:

- $PACF_k$ is significant, if $|PACF_k| > 2/\sqrt{N}$, where N is the number of observations. at 5% significance level
- If $|PACF_k| \leq 2/\sqrt{N}$, we assume the the PACF is not significant.

  In this case

$$PACF_k \in [-\frac{2}{\sqrt{N}}, \frac{2}{\sqrt{N}}] \quad \text{lies in 95\% confidence interval for 0.}$$

Example The PACF if Figure 3.6 shows that the first two partial autocorrelations are significant, since they are outside the 2SD confidence band.
This suggest, that we could fit AR(2) model.

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t.$$

Rule: To fit AR(p) model, select p the largest lag such that $PACF_k$ is significant

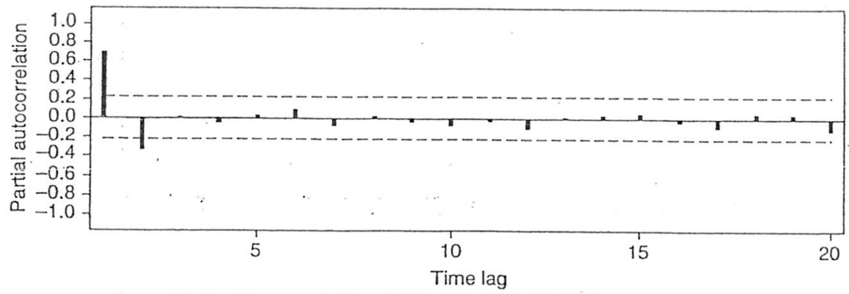


Figure 3.6 The partial autocorrelation function for the furnace data.

$AR(2)$ model. For AR(2) model

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t.$$

to be stationary it is required that coefficients satisfy the the relations:

$$\phi_2 + \phi_1 < 1$$
$$\phi_2 - \phi_1 < 1$$
$$-1 < \phi_2 < 1.$$

Example. In Figure 3.6 we see PACF of furnace temperature time series. In Table 3.1 we have output of estimated coefficients:

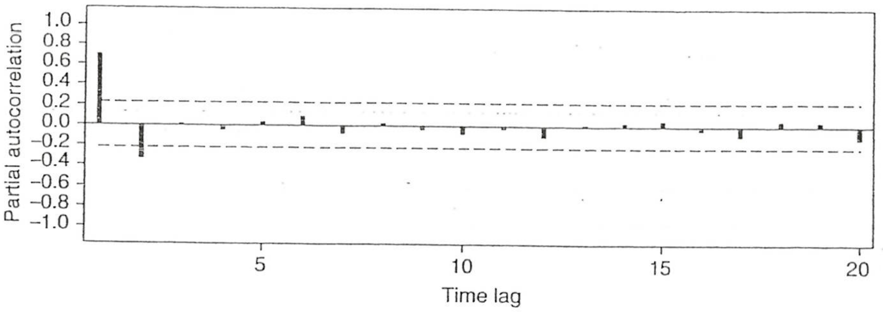


Figure 3.6 The partial autocorrelation function for the furnace data.

TABLE 3.1 Estimated Coefficients for an AR(2) Process

Coefficient	Estimate	Standard error	t-value	p-value
$\hat{\phi}_1$	0.9824	0.1062	9.25	0.000
$\hat{\phi}_2$	−0.3722	0.1066	−3.49	0.001
Constant	615.836	0.042	—	—
$\hat{\mu}$	1579.79	0.11	—	—
$\hat{\sigma}_a^2$	0.1403	—	—	—

Using them we can write the model:

$$X_t = 61.5.836 + 0.9824X_{t-1} - 0.3722X_{t-2} + \varepsilon_t, \qquad \sigma_\varepsilon^2 = 0.1403.$$

1. To verify the fit of the model we first check, that it is stationary: from

$$\hat{\phi}_2 + \hat{\phi}_1 = -0.3722 + 0.9824 = 0.61,$$
$$\hat{\phi}_2 - \hat{\phi}_1 = -0.3722 - 0.9824 = -1.35, \qquad \hat{\phi}_2 = -0.3722.$$

So, we conclude that the model is indeed stationary.

The mean $E[X_t]$ of AR(2) model. It is easy to compute the mean as follows:

Assume that X_t is stationary. Then $EX_t = \mu_X$ for all t. Taking expectation of AR(2) equation we get

$$\begin{aligned} \mu_X = EX_t &= E[\mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t] \\ &= \mu + \phi_1 E[X_{t-1}] + \phi_2 E[X_{t-2}] + E\varepsilon_t] \\ &= \mu + \phi_1 \mu_X + \phi_2 E \mu_X. \end{aligned}$$

So, $\mu_X(1 - \phi_1 - \phi_2) = \mu$, and

$$\mu_X = \frac{\mu}{1 - \phi_1 - \phi_2}.$$

So, in our example

$$\mu_X = \frac{\hat{\mu}}{1 - \hat{\phi}_2 - \hat{\phi}_1} = \frac{61.5.836}{1 - 0.61} = 1537.$$

Residuals. Next we have to check if this model fits to the data i.e. if residuals $\hat{\varepsilon}_t$ are uncorrelated. We compute residuals:

$$\hat{\varepsilon}_t = X_t - \hat{\mu} - \hat{\phi}_1 X_{t-1} - \hat{\phi}_2 X_{t-2}$$

and then compute ACF and PACF functions, see Figure 3.7 and 3.8. They show that ACF and PACF are not significant at any lag, which means that residuals $\hat{\varepsilon}_t$ is a white noise process. Otherwise, significant correlation would mean that we are fitting wrong model. Then we could try e.g. AR(3) model and check if residuals become uncorrelated.

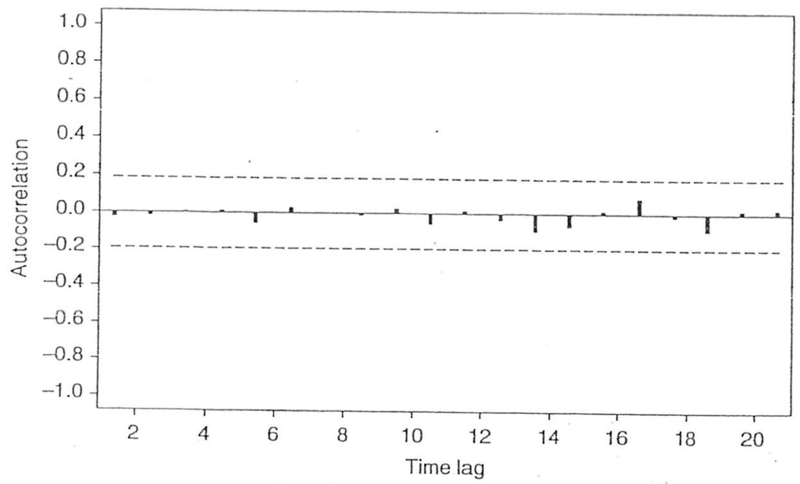


Figure 3.7 The ACF of the residuals after fitting an AR (2) model to the furnace data.

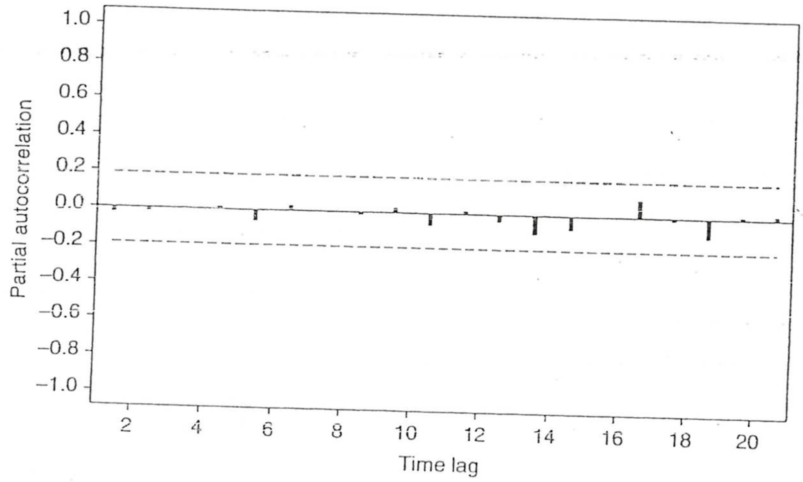


Figure 3.8 The PACF of the residuals after fitting an AR (2) model to the furnace data.

Alternative condition for stationarity. In the textbooks, for AR(2) model

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$$

you will find alternative condition for existence of a stationary solution, which is equivalent to that we had above. It is as follows.

After rearranging AR(2) equation, we get

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = \mu + \varepsilon_t.$$

Then we write the associate polynomial

$$1 - \phi_1 x - \phi_2 x^2 = 0,$$

which always has two solutions x_1 and x_2.

Rule: AR(2) model is stationary, if parameters ϕ_1 and ϕ_2 are such that $|x_1| > 1$ and $|x_2| > 1$.

Note: Similar condition for existence of a stationary solution exist also for AR(p) model, with $p \geq 3$. However, verification of this condition is more complicated than in case of the AR(1) and AR(2) models.

Moving average MA models.

Autoregressive models relate current observation X_t to previous observations $X_{t-1}, \cdots, X_{t-p}$.

Another large classes of model we can use to model stationary time series are MA (moving average) and ARMA models. They are defined as follows.

Moving average MA(q) model of order "q".

This model is "averaging" of the present and past noise terms ε_s:

$$X_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}.$$

It is defined by parameters $\theta_1, \theta_2, \cdots, \theta_q$ and σ_u^2.

Properties: Important properties of MA(q) models are:

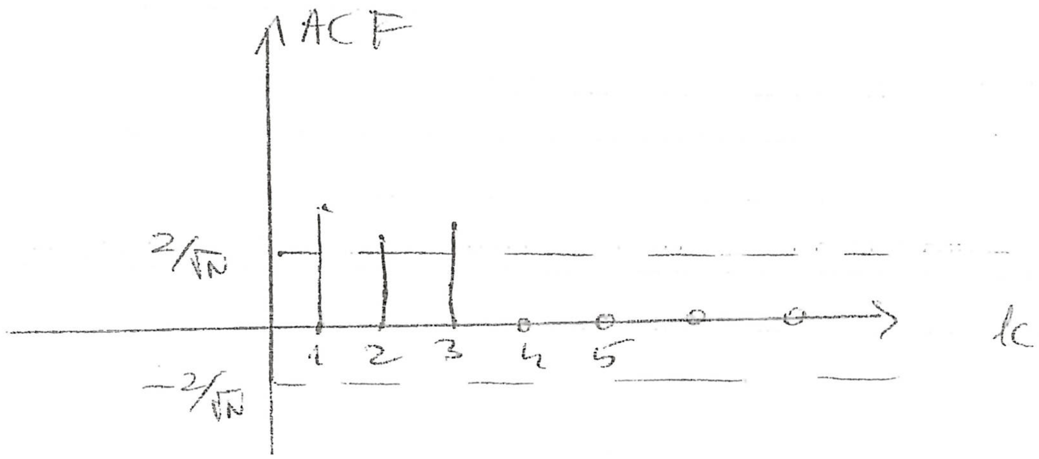
- MA(q) process X_t is always stationary.
- $E[X_t] = \mu$.
- The ACF of MA(q) model cuts off (equals to 0) after lag q.

  That means that the sample ACF will be significant and plot outside 95% confidence band given by $\pm 2/\sqrt{N}$ upto and including lag q.

  After lag q the sample ACF is expected to be within the confidence band, i.e. to be insignificant.

  This property is used to select order q when fitting MA(q) model.

Selection rule: select q as the largest lag where ACF is significant.



q = 3

Notice:

- To select the order of AR(p) model we use sample PACF function.
- To select the order of MA(q) model we use sample ACF.

ARMA(p,q) process.

X_t is defined as solution of equations

$$X_t = \mu + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}.$$

It is defined by parameters $\phi_1, \cdots, \phi_p$; $\theta_1, \theta_2, \cdots, \theta_q$ and σ_u^2.

- ARMA(p,q) model puts together AR(p) and MA(q) models. We can use it, when fitting AR(p) or MA(q) models requires large number of parameters.

  Usually, fitting ARMA model, ARMA(1,1), ARMA(1,2), ARMA(2,1) fit well: we do not need many parameters.

- Notice: ARMA(p,0) = AR(p) model, and ARMA(0,q) = MA(q) model.

The summary of behaviors of ACF and PACF for AR, MA and ARMA models is given in Table 3.2.

In Figure 3.9 we see simulated examples of ARMA models and their ACF's and PACF's.

- We can see from Figure 3.9, that ACF and PACF are excellent tools for identifying the order of MA and AR models, respectively.

  For ARMA model we do not have such simple rule using ACF and PACF. Its ACF and PACF do not cut off to zero.

- For ARMA model to select the order p and q we can use information criterions AIC and BIC.

- Using software packages we can always try to find a model between AR, MA and ARMA model, which gives uncorrelated residuals.

- We should seek for a model which has smallest number of parameters.

TABLE 3.2 Summary of Properties of Autoregressive (AR), Moving average (MA), and Mixed Autoregressive moving average (ARMA) processes

	AR(p)	MA(q)	ARMA(p, q)
Model	$w_t = \phi_1 w_{t-1} + \ldots + \phi_p w_{t-p} + a_t$	$w_t = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$	$w_t = \phi_1 w_{t-1} + \cdots + \phi_p w_{t-p} - \theta_1 a_{t-1} + \ldots - \theta_q a_{t-q} + a_t$
Autocorrelation function (ACF)	Infinite damped exponentials and/or damped sine waves; Tails off	Finite; cuts off after q lags	Infinite damped exponentials and/or damped sine waves; Tails off
Partial autocorrelation function (PACF)	Finite; cuts off after p lags	Infinite; damped exponentials and/or damped sine waves; Tails off	Infinite damped exponentials and/or damped sine waves; Tails off

Source: Adapted from BJR.

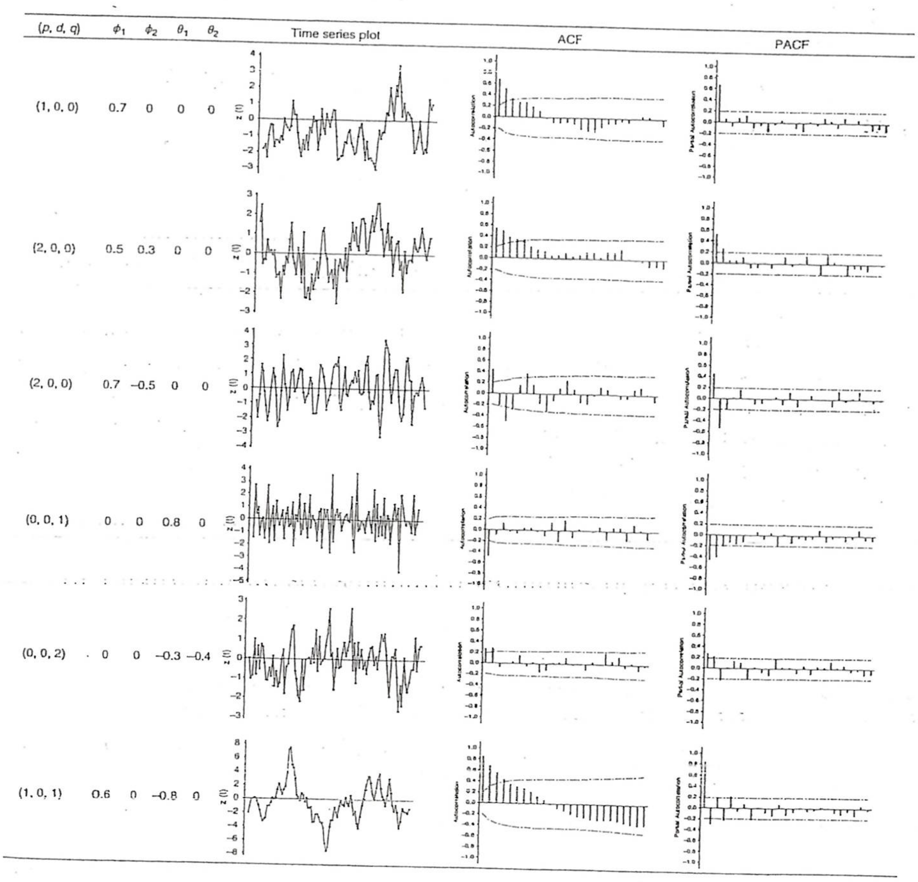


Figure 3.9 Various realizations of ARMA models.