Lecture 6

Forecasting example

Figure 4.13 shows the time series plot of 197 observations.

From the plot it is seen that

- Data is positively correlated (high values is followed by high value and low values by low value). Possibly non-stationary.
- ACF in Figure 4.15 does not die out even for large values.
- PACF is significant for first 2 lags, and close to the boundary for some higher lags.

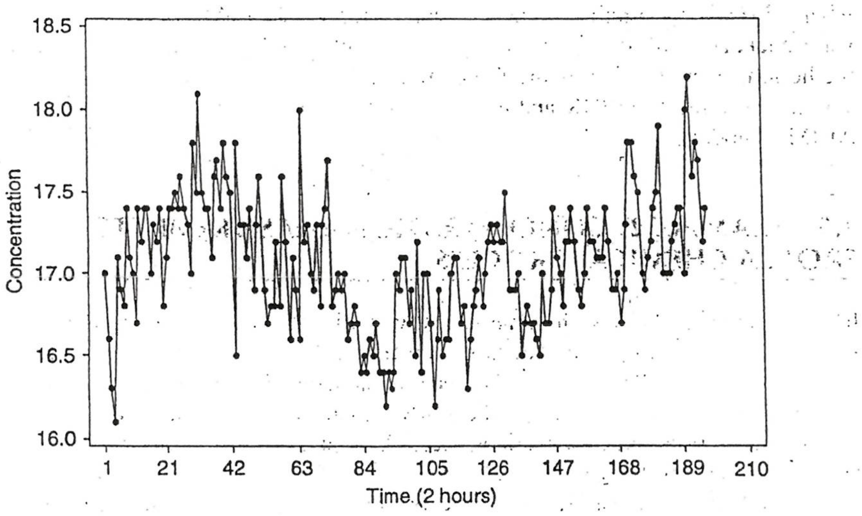
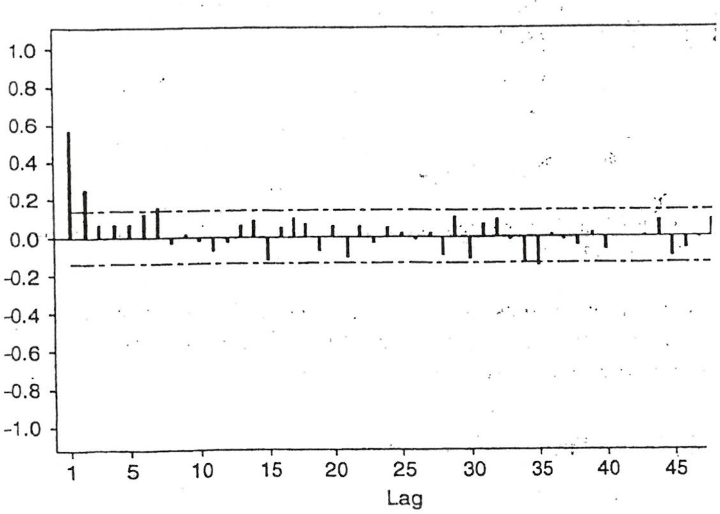


Figure 4.13 Time series plot of chemical process concentration readings sampled every 2 h: series A from BJR.



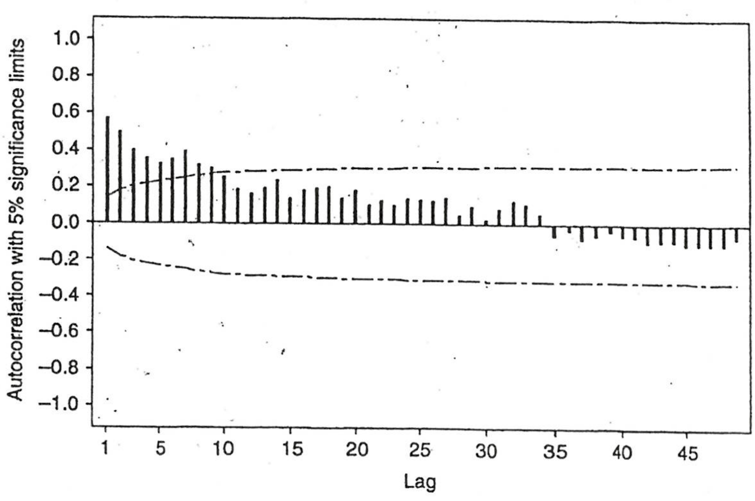


Figure 4.15 ACF of the chemical process concentration: series A from BJR.

Differences series. Next, we analyze the differenced series $z_t = X_t - X_{t-1}$.

- The plot in Figure 4.18 looks stationary.
- ACF shows large correlation at lag 1, and no correlation at higher lags. So we could fit AR(1) model.
- PACF is significant for a few lags, but dies off fast
- Estimation results in Table 4.4 suggest that we can fit MA(1) model

$$z_t = \varepsilon_t + \theta \varepsilon_{t-1}.$$

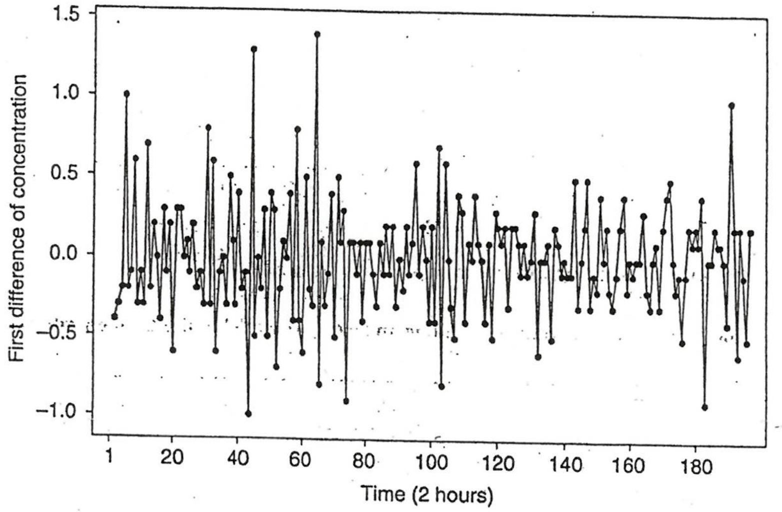
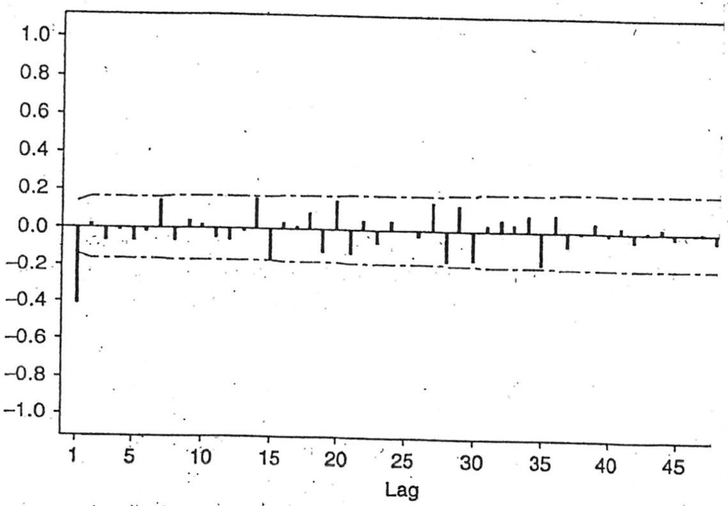


Figure 4.18 First difference of chemical process concentration: series A from BJR.



ure 4.19 ACF of the first difference of series A from BJR.

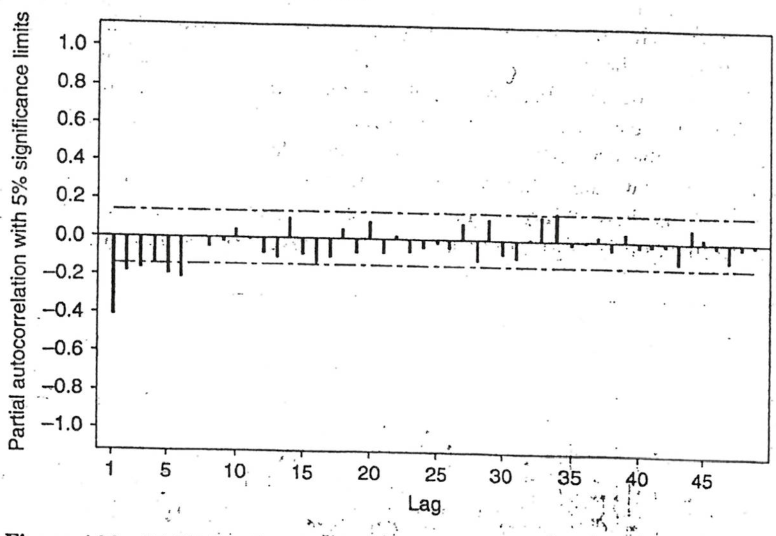


Figure 4.20 PACF for the first difference of series A from BJR.

TABLE 4.4 Summary of Information from Fitting an IMA(1, 1) Model to the Concentration Data

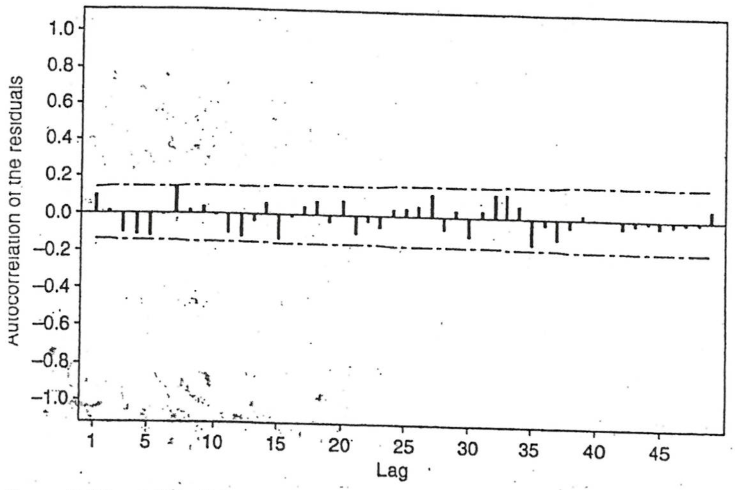
Term	Estimate	Standard error of coefficient	t	p
θ	0.7050	0.0507	13.90	0.000
Residual Sum of Squares (SS)	19.6707	$df = 195$		
Mean Squares (MS)	0.1009			

Diagnostic check

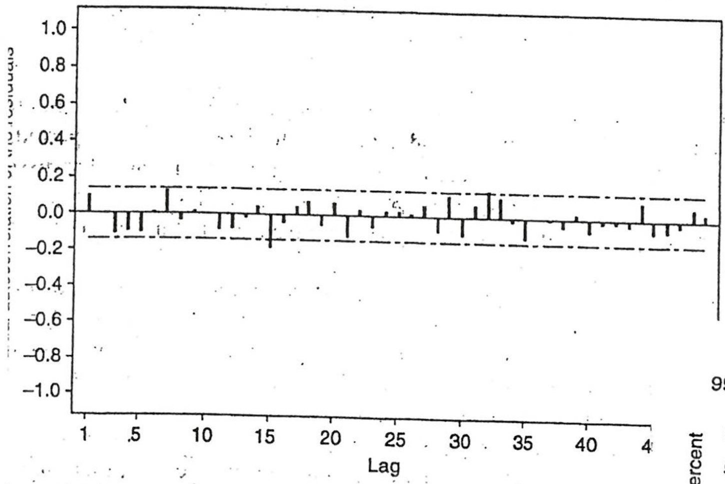
Figure 4.21 shows the ACF of residuals.
Figure 4.22 shows PACF.

Both ACF and PACF are <u>not significant</u> at any lag.

Hence, MA(1) models fits to the data z_t reasonably well.



igure 4.21 ACF of the residuals after fitting an IMA(1,1) model to the concentration
ata: series A from BJR.



igure 4.22. PACF of the residuals after fitting an IMA(1,1) model tc
ita: series A from BJR.

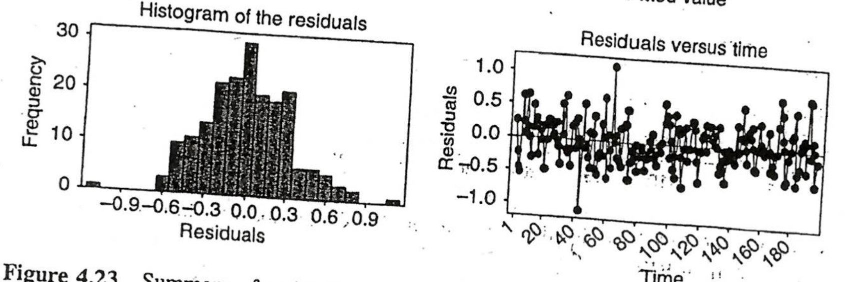


Figure 4.23 Summary of residual plots after fitting an IMA(1,1) model to the concentration data: series A from BJR.

Forecast.

Figure 4.25 shows the 1-step-ahead forecasts using MA(1) model. Forecasts follow closely the data set.

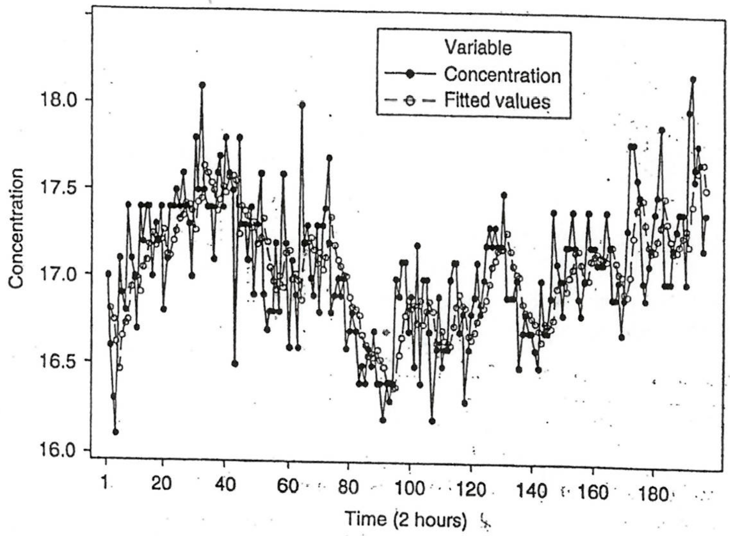


Figure 4.25 Time series plot of the concentration data (solid dots) with the fitted values from the IMA(1, 1) model superimposed (open dots).

To be or not to be stationary?

The choice between stationary and non-stationary time series is not easily made when sample is small.

In the present example one could argue that the stationary model also may fit well, see Figure 4.13.

Fitting ARMA(1,1) model $z_t = \phi_1 z_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$, we get the following estimates of parameters:

TABLE 4.5 Estimated Parameters and Summary Statistics from Fitting a Stationary ARIMA (1, 0, 1) Model to the Concentration Data, BJR Series A

Model term	Coefficient	Standard error	t	p
AR: ϕ_1	0.9151	0.0433	21.11	0.000
MA: θ_1	0.5828	0.0849	6.87	0.000
Constant	1.44897	0.0094	154.38	0.000

[a]Residual SS = 19.188; MS = 0.0989; df = 194.

They show that $\phi_1 = 0.91$ with a standard error 0.0433 is close to 1, which may be indication of nonstationarity.

One should ask question: is it reasonable that the uncontrolled chemical process generating this data is stationary?

When left alone, it is not realistic that such process will be stationary. So we are inclined to assume, that the process is non-stationary and fit to its differences $z_t = X_t - X_{t-1}$ a stationary process MA(1) model, as we did above.

5 Unit root nonstationarity

Until now we discussed stationary time series. A number of economical series and financial series, for example, interest rates, foreign exchange rates, price series of an asset tend to be nonstationary.

In time series literature, such series are called unit-root nonstationary time series.

The best known example of nonstationary time series is random walk model.

5.1 Random walk

Definition 7 *A time series X_t is called a random walk if*

$$X_t = X_{t-1} + \varepsilon_t, \qquad t = 0, 1, 2, \cdots$$

where ε_t is a zero mean white noise.

We shall denote by X_0 the starting value of the random walk. For example, if X_t is the log price of a stock at time t, then X_0 is the initial offering.

Note: if ε_t has a symmetric distribution around zero, then p_t has a 50-50 chance going up or down around X_{t-1}.

Analogy to AR(1) model: if we treat a random walk as a AR model

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

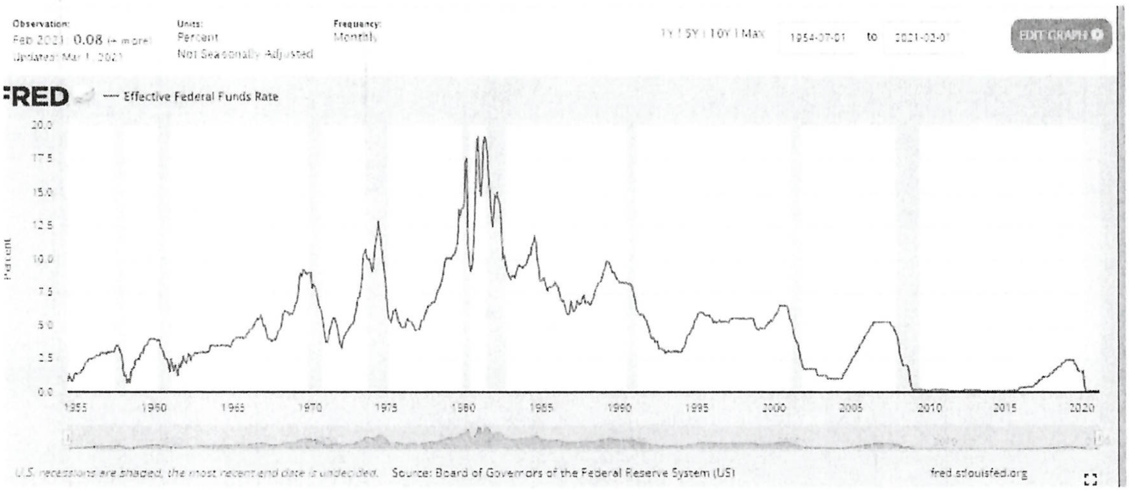
then the coefficient $\phi = 1$ at Y_{t-1} equals to 1. Such model does not satisfy condition of covariance stationarity of an AR(1) model.

Examples of plots of non-stationary time series

1. The **S&P 500** Index (Standard & Poor's **500** Index):

market-capitalization-weighted index of the **500** largest U.S. publicly traded companies

2. Effective Federal Funds Rate



US. recessions are shaded

Source: https://fred.stlouisfed.org/series/FEDFUNDS

A random-walk is covariance non-stationary model.

Variance

Assume that $X_0 = 0$. Then

$$X_t = X_{t-1} + \varepsilon_t = X_{t-2} + \varepsilon_{t-1} + \varepsilon_t = \dots = \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t.$$

Therefore

$$EX_t = E[\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t] = 0,$$

and

$$\begin{aligned} Var(X_t) &= E(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t)^2 \\ &= E[\varepsilon_1^2] + E[\varepsilon_2^2] + \dots + E[\varepsilon_t^2] \\ &= \sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \dots + \sigma_\varepsilon^2 = t\sigma_\varepsilon^2. \end{aligned}$$

The variance

$$Var(X_t) = t\sigma_\varepsilon^2$$

depends on time t and increases with t. That indicates that X_t is a non-stationary time series.

Unit root model is often used as a statistical model for the movement of log prices:

$$X_t = \log P_t, \qquad P_t \quad \text{price at time } t.$$

Properties of a unit-root model:

- Unit-root process X_t is not predictable
- not mean-reverting.

Consequently, log prices not predictable and not mean-reverting.

First we show that X_t is not-predictable. Note that $X_{t+1} = X_t + \varepsilon_{t+1}$.

Then 1-set ahead forecast of the process X_t at the origin t is:

$$\begin{aligned} \hat{X}_t(1) &= E[X_{t+1}|F_t] = [X_{t+1}] \\ &= [X_t + \varepsilon_{t+1}] = [X_t] + [\varepsilon_{t+1}] \\ &= X_t. \end{aligned}$$

In terms of log prices: if X_t is the log price of the stock at time t, the the forecast is the last observation.

To find the 2-step ahead forecast, notice that

$$X_{t+2} = X_{t+1} + \varepsilon_{t+2}.$$

So,

$$\begin{aligned} \hat{X}_t(2) &= E[X_{t+2}|F_t] = [X_{t+2}] = [X_{t+1} + \varepsilon_{t+2}] \\ &= [X_{t+1}] + [\varepsilon_{t+2}] = [X_{t+1}] \\ &= X_t, \end{aligned}$$

which is again the last observation X_t.

In general, for any k-step ahead forecast,

$$\hat{X}_t(k) = X_t.$$

So, the forecast of a random-walk model is last observation X_t.

Therefore: forecast is not mean reverting.

Forecast error. We showed that

$$X_t = \varepsilon_t + \varepsilon_{t-1} + ... + \varepsilon_1.$$

The k-step ahead forecasting error:

$$e_t(k) \equiv X_{t+k} - \hat{X}_t(k) = \{\varepsilon_{t+k} + \varepsilon_{t-1} + ... + \varepsilon_1\} - \{\varepsilon_t + ... + \varepsilon_1\} = \varepsilon_{t+k} + \cdots, \varepsilon_{t+1}.$$

The variance of forecast error increases with k:

$$Var(e_t(k)) = E[(\varepsilon_{t+k} + \varepsilon_{t+k-1} + ... + \varepsilon_{t+1})^2] = k\sigma_\varepsilon^2 \to \infty.$$

This shows again, that model in not predictable in the long-run. So, X_t (e.g. log prices) can take any real value for large t.

Property: a unit-root time series has sample autocorrelation function (ACF), which approaches to 1 as the sample size n increases.

Market indexes need a different model.

5.2 Random walk with a drift

Empirical examples we discussed, show that the log return series

$$r_t = \log(\frac{P_t}{P_{t-1}}) = \log P_t - \log P_{t-1}$$

of the market index tend to have a small positive mean.

Model: for the log price $Y_t = \log P_t$ can be written as a random walk with a drift:

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

where

$$\mu = E[X_t - X_{t-1}]$$

and ε_t is a white noise.

The constant parameter μ in financial time series represents the time trend or the drift of (the log price) X_t.

Assume that the initial value is X_0. Then

$$\begin{aligned}
X_1 &= \mu + X_0 + \varepsilon_1, \\
X_2 &= \mu + X_1 + \varepsilon_2 = 2\mu + X_0 + \varepsilon_2 + \varepsilon_1, \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
X_t &= \mu + X_{t-1} + \varepsilon_t = t\mu + X_0 + \varepsilon_t + \dots + \varepsilon_2 + \varepsilon_1.
\end{aligned}$$

This equation shows that the log price Y_t contains:

- a time trend μt
- a pure random walk $\varepsilon_t + ... + \varepsilon_2 + \varepsilon_1$.

It shows that

$$E[X_t] = \mu t + E[X_0].$$

Observation: since $Var(\varepsilon_t + ... + \varepsilon_2 + \varepsilon_1) = t\sigma_\varepsilon^2$, the SD $\sqrt{t}\sigma_\varepsilon$ of X_t is grows slower than the mean $E[X_t] = \mu t$:

$$\sqrt{t}\sigma_\varepsilon << \mu t.$$

Consequence: the linear trend μt dominates! Graphing X_t we obtain a time trend with slope μ.

- a positive $\mu > 0$ implies that the log price X_t increases.
- a negative $\mu < 0$ implies that the log price X_t decreases

Therefore not surprising that the log returns of the equal-weighted index have a small, but statistically significant mean $\mu > 0$.

Example. This example illustrates effect of the drift on the price series of monthly log stock returns of the 3M Company for the period February 1946-December 1997.

The series of log returns r_t has no significant correlation and follows the model

$$r_t = 0.00115 + \varepsilon_t, \qquad \hat{\sigma}_\varepsilon^2 = 0.0000639$$

where 0.00115 is the sample mean of r_t,
and $0.00026 = \sigma_\varepsilon$ is the standard error.
Note: sample mean is significant.

By definition $r_t = \log p_t - \log p_{t-1}$. Therefore we can write

$$\log p_t = 0.00115 + \log p_{t-1} + \varepsilon_t$$

where p_t is price at the moment t.

Figure 2.10 shows time plots of $\log p_t$, and the straight line $y_t = 0.0115t$.

The impact if the constant 0.00115 is evident: the slope of the upward trend of $\log p_t$ is about 0.00115

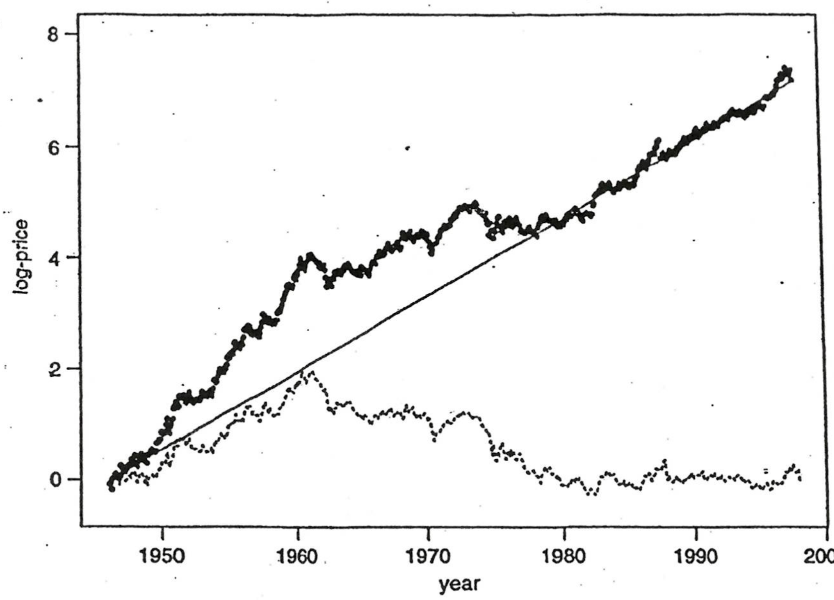


Figure 2.10. Time plots of log prices for 3M stock from February 1946 to December 1997, assuming that the log price of January 1946 was zero. The dashed line is for log price without time trend. The straight line is $y_t = 0.0115 \times t$.

5.2.1 Nonstationary ARIMA unit root models

A random walk model is a separate case of a nonstationary ARIMA model, or unit root.

It can be transformed to a stationary model by differencing.

Definition. A time series, X_t is said to be an ARIMA(p,1,q) model if

$$X_t - X_{t-1} = Y_t, \qquad t = 0, 1, 2, \cdots$$

where Y_t is a stationary ARMA(p,q) process.

Example. Prices P_t is commonly believed to be non-stationary
Log return series $r_t = \log P_t - \log P_{t-1}$ is stationary.
Log prices $\log P_t$ follow an ARIMA(p,1,q) model, often with $p = q = 0$.

5.3 Testing for Unit Root

In practice often we need to test:

- If time series X_t follows a random walk
- unit root model
- a random walk (unit root) with a drift.

For testing we use the models

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \varepsilon_t, & (1) \\ X_t &= \phi_0 + \phi_1 X_{t-1} + \varepsilon_t & (2) \end{aligned}$$

where ε_t denotes an error term.
Here: ε_t is white noise in case of a random walk;
ε_t stationary ARMA process in case of unit root

Dickey and Fuller test. Here ε_t is white noise
We test the null hypothesis

H_0 **(random walk/unit root)**: $\phi_1 = 1$

against alternative

H_1: $|\phi_1| < 1$ (Y_t **is a stationary series.**)

This is Dickey and Fuller (1979) test.

It uses least squares estimates $\hat{\phi}_1$ and $\hat{\sigma}_\varepsilon^2$:

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{N} Y_t Y_{t-1}}{\sum_{t=1}^{N} Y_{t-1}^2}, \qquad \hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=1}^{N}(Y_t - \hat{\phi}_1 Y_{t-1})^2}{N-1}$$

where $Y_0 = 0$ and N is the sample size.

Testing statistic: the t-ratio statistic is

$$t_{ratio} = \frac{\hat{\phi} - 1}{\mathrm{std}(\hat{\phi}_1)} = \frac{\sum_{t=1}^{N} Y_{t-1}\varepsilon_t}{\hat{\sigma}_\varepsilon \sqrt{\sum_{t=1}^{N} Y_{t-1}^2}}.$$

Applying it, in E-views, you need to select to include drift, i.e. $\phi_0 \neq 0$, or no drift, i.e. $\phi_0 = 0$.

Augmented Dickey-Fuller test

Note:

- In D-F test we assume that ε_t is white noise
- in ADF test we assume that ε_t is AR(p) stationary time series. Here we need to select p

Applying ADF unit root test, we fit to the data the AR(p) model:

$$X_t = \alpha + \mu t + \beta X_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta X_{t-i} + \varepsilon_t$$

where $\Delta X_j = X_j - X_{j-1}$ is the differenced series X_j.

To verify the presence of a unit root in X_t we test hull hypothesis

H_0: $\beta = 1$

against alternative

H_1: $\beta < 1$.

In practice,

- we can set $\alpha + \mu t = 0$ to be zero or constant
- include a linear trend $\alpha + \mu t$.

Augmented Dickey-Fuller test:

We reject unit root if test statistic $t <$"critical value", or if p-value < 0.05

Alternative form of ADF test: We rewrite the model as

$$\Delta X_t = c_t + \beta_c X_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta X_{t-i} + \varepsilon_t,$$

where $\beta_c = \beta - 1$. Then, testing for unit root is equivalent of testing the following hypotheses

H_0: $\beta_c = 0$

against alternative

H_1: $\beta_c < 0$.

Example 1

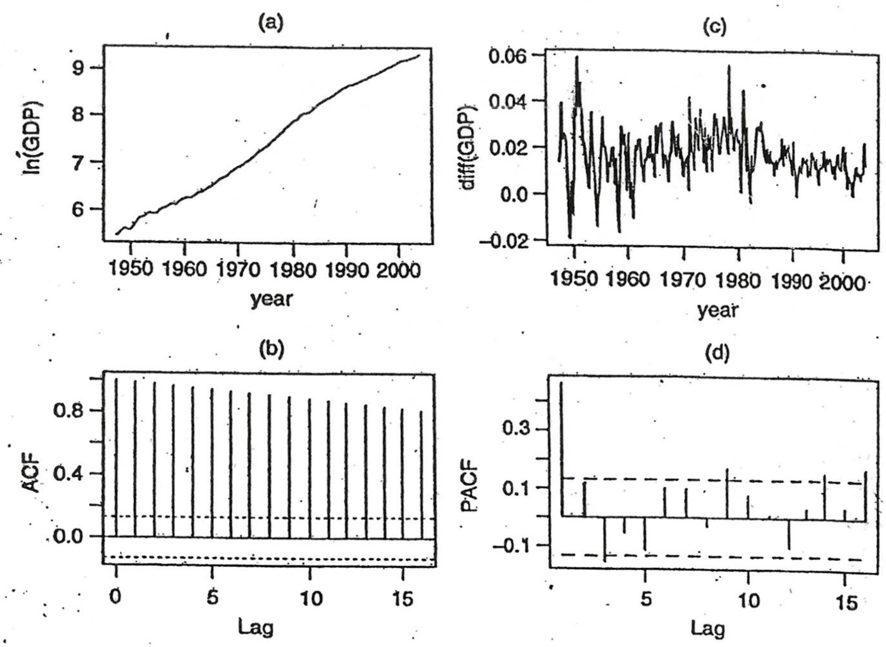


Figure 2.11. Log series of U.S. quarterly GDP from 1947.I to 2003.IV: (a) time plot of the logged GDP series, (b) sample ACF of the log GDP data, (c) time plot of the first differenced series, and (d) sample PACF of the differenced series.

Figure 2.11 (a) presents the log series of the US quarterly gross domestic product (GDP), from 1947 to 1923.

- The upward trend shows the growth of US economy.
- Picture (b) shows that log series has high serial correlation.
- picture (c) shows the differenced log series is the growth rate r_t. It is varying around fixed mean, and variability seems to be smaller in recent years.
- Picture (d) shows the PACF of the growth rate r_t.

To confirm unit root in $\log X_t$, we apply ADF test to the log series $\log X_t$. Based on the sample PACF, we choose AR(p) model with the "number of lags" $p = 10$.

With $p = 10$, the ADF test statistic value is -1.31, and the p- value is 0.7038. We do not reject unit root.

The output below shows that

$$\hat{\beta} = 1 + \hat{\beta}_c = 1 - 0.0006 = 0.9994.$$

S-Plus Demonstration
Output edited.

```
> adft=unitroot(gdp,trend='c',method='adf',lags=10)
> summary(adft)

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t test
 Test Statistic: -1.131
        P-value: 0.7038

Coefficients:
         Value Std. Error t value Pr(>|t|)
   lag1 -0.0006  0.0006   -1.1306  0.2595
   lag2  0.3797  0.0679    5.5946  0.0000

...

  lag10  0.1798  0.0656    2.7405  0.0067
constant 0.0123  0.0048    2.5654  0.0110

Regression Diagnostics:

         R-Squared 0.2831
Adjusted R-Squared 0.2485

Residual standard error: 0.009498 on 214 degrees of freedom.
```

Example 2. Figure 2.12 shows the daily log series X_t of the S&P index from Jan 1990 to Dec 2003 with 3532 observations.
Testing for a unit root in log series X_t is equivalent to verification of hypothesis that the X_t follows a random walk with a drift.

- We use ADF test with 13 lags: fitting $AR(13)$ obtained by AIC criterion applied to for the first differences of the log index.
- We include trend $\alpha + \mu t$ and apply the augmented Dickey-Fuller test
- The value of statistic is -0.9648 and p-value 0.9469. Hence the unit root hypothesis cannot be rejected.
- Estimation shows that trend parameters α and μ are not significant.

Conclusion: the log series of the index contains unit root, but no time trend.

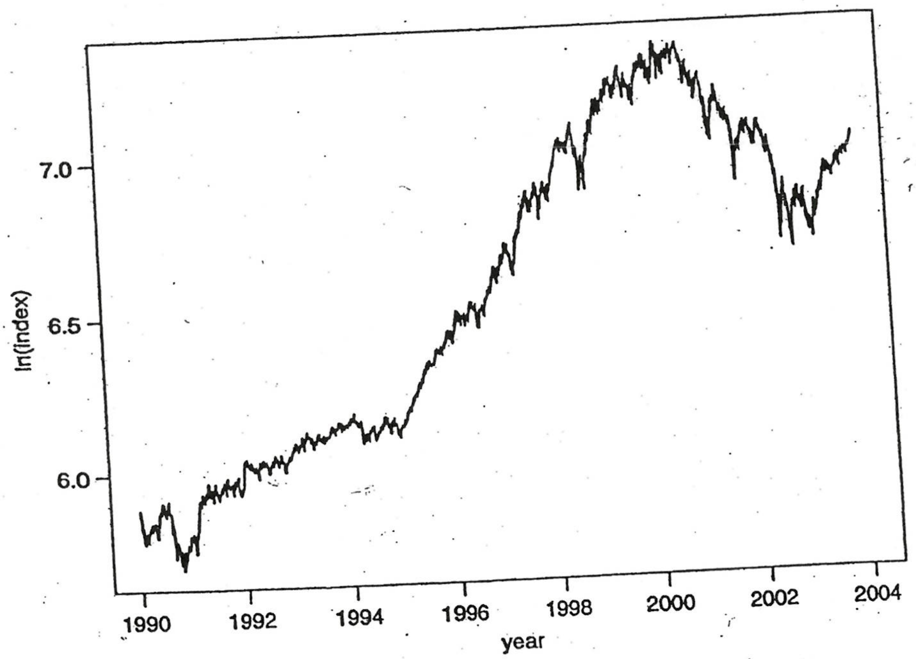


Figure 2.12. Time plot of the logarithm of daily S&P 500 index from January 2, 1990 to December 31, 2003.

S-Plus Demonstration
Output edited.

```
> adft=unitroot(sp,method='adf',trend='ct',lags=14)
> summary(adft)

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t test
 Test Statistic: -0.9648
        P-value: 0.9469

Coefficients:
           Value Std. Error t value Pr(>|t|)
    lag1 -0.0008  0.0008    -0.9648  0.3347
    ...
   lag14  0.0319  0.0169     1.8894  0.0589
constant  0.0056  0.0054     1.0316  0.3023
    time  0.0000  0.0000     0.4871  0.6262

Regression Diagnostics:

          R-Squared 0.0107
 Adjusted R-Squared 0.0065

Residual standard error: 0.01049 on 3514 degrees of freedom
```