Times Series and Forecasting (II)

Chapter 2. Fundamental Concepts

Jianhua Hu

School of Statistics and Management
Shanghai University of Finance and Economics

Spring 2013

2.1. Discrete stochastic processes and times series

- The sequence of random variables $\{Y_t : t = 0, 1, 2, \cdots,\}$, or $\{Y_t\}$, is called a discrete stochastic process (DSP).
- It is a collection of random variables indexed by time t, i.e.

$$\begin{aligned} Y_0 &= \text{value of the process at time } t = 0 \\ &\cdots \\ Y_n &= \text{value of the process at time } t = n. \end{aligned}$$

- A DSP can be described as "a statistical phenomenon that evolves through time according to a set of probabilistic laws".
- An equally spaced time DSP is a time series process. Any realization of a time series process is a times series.

2.1. DSP and times series (continued)

- A complete probabilistic time series model for $\{Y_t\}$ would specify all of the joint distributions of random vectors $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_n)$, for all $n = 1, 2, \cdots$, or, equivalently, specify the joint probabilities

$$P\left(Y_1 \leq y_1, Y_2 \leq y_2, \cdots, Y_n \leq y_n\right),$$

for all $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ and $n = 1, 2, \cdots$.

- In practice, it is hard to get joint distributions for all n.
- In this course, we specify only the first and second-order moments; namely, $\mathsf{E}(Y_t)$ and $\mathsf{E}(Y_t Y_{t-k})$, for $k, t = 0, 1, 2, \cdots$.

2.2. Means and autocovariance

- For the stochastic process $\{Y_t : t = 0, 1, 2, \cdots,\}$, the mean function is defined as

$$\mu_t = \mathsf{E}(Y_t),$$

for $t = 0, 1, 2, \cdots$.

- The autocovariance function (variances) is defined as

$$\gamma_{t,s} = \mathsf{Cov}(Y_t, Y_s),$$

for $t, s = 0, 1, 2, \cdots$, with

$$\mathsf{Cov}(Y_t, Y_s) = \mathsf{E}(Y_t Y_s) - \mathsf{E}(Y_t)\mathsf{E}(Y_s).$$

Autocorrelation

The autocorrelation function (ACF) is given by

$$\rho_{t,s} = \text{corr}(Y_t, Y_s) = \frac{\text{Cov}(Y_t, Y_s)}{\sqrt{\text{var}(Y_t)\text{var}(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}.$$

- $\rho_{s,t}$ is close to $\pm 1 \Longrightarrow$ strong linear dependence between Y_t and Y_s.
- $\rho_{s,t}$ is close to $0 \Longrightarrow$ weak linear dependence between Y_t and Y_s.
- $\rho_{s,t} = 0 \Longrightarrow Y_t$ and Y_s are uncorrelated.

2.3.1. White noise process

A process $\{e_t : t = 0, 1, 2, \cdots, \}$ is called a white noise process if it is a sequence of uncorrelated (independent) and identically distributed random variables with

$$\mathrm{E}(e_t) = \mu_e \text{ and } \mathrm{var}(e_t) = \sigma_e^2.$$

- μ_e and σ_e^2 are constant, free of t. Write $\{e_t\} \sim WN(\mu_e, \sigma_e^2)$.
- Assume that $\{e_t\} \sim WN(0, \sigma_e^2)$ throughout our course.
- If $e_t \sim N(0, \sigma_e^2)$ for all t, or $\{e_t\} \sim$ normal WN$(0, \sigma_e^2)$, $\{e_t\}$ is a sequence of iid (for linear time series models).

Figure 2.1. Four simulated realizations of $e_t \sim N(0, 1)$

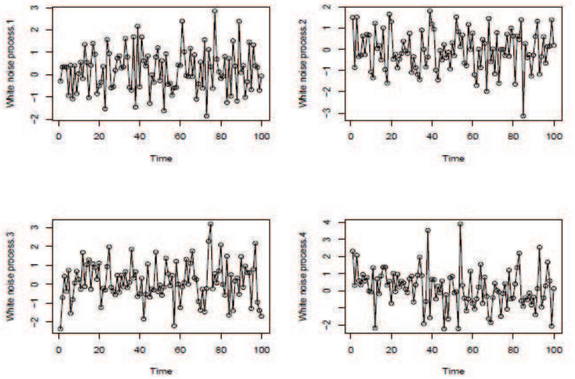


Figure 2.1: *Four white noise processes* $e_t \sim$ *iid* $\mathcal{N}(0, \sigma_e^2)$, *where* $n = 100$ *and* $\sigma_e^2 = 1$.

SAS programming for generating data of white noise N(0,1)

```
Data figure21;
do t=0 to 100;
y1=normal(0);
output;
end;
run;
Proc gplot data= figure21; symbol i= v=circle h=1.5;
title 'Scatter plot of the 1st simulated realization of WN N(0,1)';
plot y1 * t;
run;
symbol i=spline v=circle h=1.5;
title 'The first simulated realization of WN N(0,1)';
plot y1 * t;
run;
```

Mean and autocovariances of white noise process

- For $t \neq s$, the independence of $\{e_t\}$ gives

$$\text{Cov}(e_t, e_s) = 0.$$

- Thus, the autocovariance function of $\{e_t\}$ is

$$\gamma_{t,s} = \begin{cases} \sigma_e^2, & t = s \\ 0, & t \neq s. \end{cases}$$

- Thus, the autocorrelation function of $\{e_t\}$ is

$$\rho_{t,s} = \begin{cases} 1, & t = s \\ 0, & t \neq s. \end{cases}$$

Remarks for white noise processes

- A white noise is the simplest time series model.
- A white noise process, itself, is generally not interesting!
- However, white noise processes play a crucial role in the analysis of time series data! Why? Many important time series models can be constructed from white noise.
- Engineers and physicists use the term "white noise" to describe a random signal of every frequency in the audio or visual spectrum, all of which have an average uniform power level.

Remarks for white noise processes (continued)

- To understand the crucial role in the analysis of time series data, let us look at analysis approaches to time series processes $\{Y_t\}$.
- Time series process $\{Y_t\}$ contain two different types of variation:
  - systematic variation (that we would like to capture and model; e.g., trends, seasonal components, etc.)
  - random variation (that is just inherent background noise in the process).

Remarks for white noise processes (continued)

- Our goal as data analysts is to extract the systematic part of the variation in the data (and incorporate this into our model).

- If we do an adequate job of extracting the systematic part, then the only variation "left over" should be just random noise, which can be modeled by a white noise.

2.3.2. Random walk

- Define

$$
\begin{aligned}
Y_1 &= e_1 \\
Y_2 &= e_1 + e_2 \\
&\vdots \\
Y_n &= e_1 + e_2 + \cdots + e_n.
\end{aligned}
$$

- By this definition, note that we can write, for $t > 1$,

$$Y_t = Y_{t-1} + e_t.$$

- The process $\{Y_t\}$ is called a random walk process.

Mean and autocorrelation of a random walk process

- The mean of Y_t is $\mu_t = \mathsf{E}(Y_t) = \mathsf{E}(e_1 + \cdots + e_t) = 0$.
- The autocovariance of Y_t and Y_s is

$$\gamma_{t,s} = \mathsf{Cov}(e_1 + \cdots + e_t, e_1 + \cdots + e_s) = \min\{s, t\}\sigma_e^2.$$

- The autocorrelation function is

$$\rho_{t,s} = \mathsf{corr}(Y_t, Y_s) = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \frac{\min\{s, t\}}{\sqrt{st}}.$$

Mean and autocorrelation of a random walk process

- $\rho_{t,s}$ is closer to 1 when t is close to s. That is, observations that are close together in time are more positively correlated than observations which are far apart.
- For t fixed, the correlation becomes smaller as $d(t,s)$ increases. In fact, for t fixed, it is easy to see that

$$\lim_{s\to\infty}\rho_{t,s} = 0.$$

- Random walk processes often are used to model stock prices, movements of molecules in gases and liquids, wild animal locations, etc.

Figure 2.2. Four simulated realizations of a N(0,1) walk

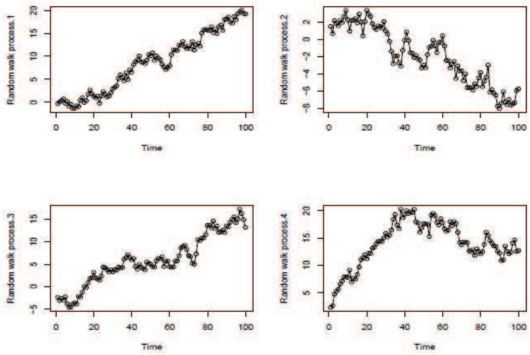


Figure 2.2: *Four random walk processes* $Y_t = Y_{t-1} + e_t$, *where* $e_t \sim iid\ \mathcal{N}(0, \sigma_e^2)$, $n = 100$, *and* $\sigma_e^2 = 1$. *These were constructed from the white noise processes in Figure 2.1.*

SAS programming for random walk with the normal WN(0,1)

```
Data figure22;
y1=0;
do t=0 to 100;
y1=y1+normal(0);
output;
end;
run;
Proc gplot data=figure22;
symbol i= v=circle h=1;
title 'Scatter plot of the 1st simulated realization of a N(0,1) walk';
plot y1 * t;
run; symbol i=spline v=circle h=1.5;
title 'The 1st simulated realization of a N(0,1) walk';
plot y1 * t;
run;
```

2.3.3. Random walk with drift

- Define

$$\begin{aligned} Y_1 &= \theta_0 + e_1 \\ Y_2 &= 2\theta_0 + e_1 + e_2 \\ &\vdots \\ Y_n &= n\theta_0 + e_1 + e_2 + \cdots + e_n. \end{aligned}$$

- By this definition, note that we can write, for $t > 1$,

$$Y_t = \theta_0 + Y_{t-1} + e_t.$$

- The process $\{Y_t\}$ is called a random walk process with drift.
- The constant θ_0 is called the drift parameter. Note that if $\theta_0 = 0$, then this process becomes a random walk.

Mean and autocorrelations

- The mean of Y_t is

$$\mu_t = \mathsf{E}(Y_t) = \mathsf{E}(t\theta_0 + e_1 + \cdots + e_t) = t\theta_0.$$

  Thus, the mean function μ_t changes with time (compare this to the random walk, where the mean function is zero for all t).

- For the random walk process with drift, the autocovariance function is

$$\gamma_{t,s} = \mathsf{Cov}(Y_t, Y_s) = \min\{s,t\}\sigma_e^2.$$

- The autocorrelation function is

$$\rho_{t,s} = \mathsf{corr}(Y_t, Y_s) = \frac{\min\{s,t\}}{\sqrt{ts}}.$$

Figure 2.3. Four realizations of a random walks with drift

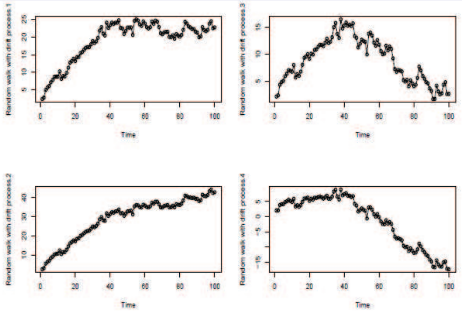


Figure 2.3: *The top two processes are the same. Each is the fourth random walk process in Figure 2.2. The middle processes are the drift versions of the top process with* $\theta_0 = 0.1$ *(left) and* $\theta_0 = -0.1$ *(right). The bottom processes are the drift versions of the top process with* $\theta_0 = 0.3$ *(left) and* $\theta_0 = -0.3$ *(right).*

SAS programming for a random walk with drift

```
Data figure23;
y=0; theta=0.1;
do t=1 to 100;
y=theta+y+normal(0);
output;
end;
run;
Proc gplot data= figure23;
symbol i=spline v=circle h=1.5;
title '1st simulated realization of a N(0,1) random walk with drift 0.1';
plot y * t;
run;
```

2.3.4. A moving average model

- Define
$$Y_t = \frac{1}{3}\left(e_t + e_{t-1} + e_{t-2}\right),$$
that is, Y_t is a running (or moving average of the white noise process (averaged across the most recent 3 time periods).

- On pp 14-15 in textbook CC, consider
$$Y_t = \frac{1}{2}\left(e_t + e_{t-1}\right),$$

Mean and autocovariances

- The mean of Y_t is $\mu_t = \mathsf{E}(Y_t) = \mathsf{E}(e_t + e_{t-1} + e_{t-2})/3 = 0.$
- If $s > t + 2$, then $\gamma_{t,s} = 0$ because Y_t and Y_s are uncorrected.
- And

$$\begin{aligned}
\gamma_{t,t} &= \mathsf{Cov}(Y_t, Y_t) = \mathsf{var}(Y_t) = \frac{1}{3}\sigma_e^2;\\
\gamma_{t,t+1} &= \mathsf{Cov}(Y_t, Y_{t+1}) = \frac{1}{9}\mathsf{Cov}\left(\sum_{i=0}^{2} e_{t-i}, \sum_{j=-1}^{1} e_{t+j}\right) = \frac{2}{9}\sigma_e^2;\\
\gamma_{t,t+2} &= \frac{1}{9}\mathsf{Cov}(Y_t, Y_{t+2}) = \mathsf{Cov}\left(\sum_{i=0}^{2} e_{t-i}, \sum_{j=0}^{2} e_{t+j}\right) = \frac{1}{9}\sigma_e^2.
\end{aligned}$$

Autocovariance and autocorrelation

- Thus, the autocovariance function can be written as

$$\gamma_{t,s} = \begin{cases} \sigma_e^2/3, & |t-s| = 0 \\ 2\sigma_e^2/9, & |t-s| = 1 \\ \sigma_e^2/9, & |t-s| = 2 \\ 0, & |t-s| > 2. \end{cases}$$

- The autocorrelation function for this moving average process is

$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \begin{cases} 1, & |t-s| = 0 \\ 2/3, & |t-s| = 1 \\ 1/3, & |t-s| = 2 \\ 0, & |t-s| > 2. \end{cases}$$

Figure 2.4. Four realizations of the moving average process

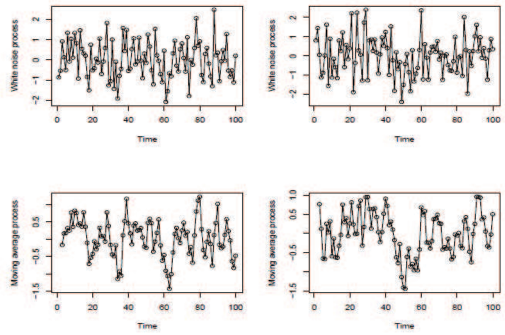


Figure 2.4: *The top plots are two different normal zero-mean white noise processes with* $\sigma_e^2 = 1$*. The bottom plots are the moving average* $Y_t = \frac{1}{3}(e_t + e_{t-1} + e_{t-2})$ *versions of the top plots.*

SAS programming for MA $Y_t = (e_t + e_{t-1} + e_{t-2})/3$

```
Data figure24_1;
do t=1 to 100; y_t = normal(0); output; end; run;
Data figure24_2;
set figure24_1;
if t=1 then delete; y_1 = y_t; t = t - 1; drop y_t; run;
Data figure24_3; set figure24_1;
if t < 3 then delete; y_2 = y_t; t = t - 2; drop y_t; run;
Data figure24;
merge figure24_1 figure24_2 figure24_3;
by t; y = (y_t + y_1 + y_2)/3; t = t + 2; run;
Proc gplot data=figure24; ...; run;
```

2.3.5. An autoregressive model

- Consider the stochastic process defined by

$$Y_t = 0.7Y_{t-1} + e_t.$$

- Y_t is related to the (downweighted) previous value of Y_{t-1} and e_t (a "shock" or "innovation" that occurs at time t).
- This is called an autoregressive model. Autoregression means "regression on itself". Essentially, we can envision "regressing" Y_t on Y_{t-1}.
- The usual calculations for this autoregressive model are postponed to Chapter 4.

Figure 2.5. Four realizations of the autoregressive model

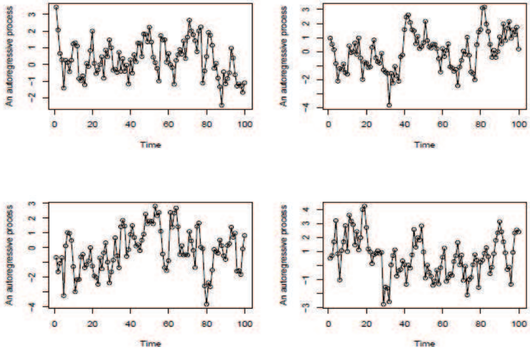


Figure 2.5: *Four different realizations of the autoregressive model* $Y_t = 0.7Y_{t-1} + e_t$ *with* $n = 100$ *and* $\sigma_e^2 = 1$.

SAS programming for the AR $Y_t = 0.7Y_{t-1} + e_t$

```
Data figure25;
y=0; a=0.7; do t=1 to 100;
y=a*y+normal(0); output; end; run;

Proc gplot data=figure25;
symbol i=spline v=circle h=1.5;
plot y * t;
title '1st simulated realization of the AR model Y_t = 0.7Y_{t-1} + e_t';
run;
```

2.3.6. A sinusoidal model

- Many time series processes in practice exhibit seasonal patterns that correspond to different weeks, months, years, etc.
- One way to capture these seasonal patterns is to use models with deterministic parts which are trigonometric in nature.
- Consider the process defined by

$$Y_t = a\mathsf{sin}(2\pi\omega t + \phi) + e_t,$$

  where a is the amplitude, ω is the frequency of oscillation, and $\phi/2\pi\omega$ is the phase shift.
- With $a = 2$, $\omega = 1/50$ (one cycle/50 time points), and $\phi = 0.6\pi$, we have

$$\mathsf{E}(Y_t) = 2\mathsf{sin}(2\pi t/50 + 0.6\pi) \text{ and } \mathsf{var}(Y_t) = \sigma_e^2.$$

Figure 2.6. Sinusoidal model illustration and realizations

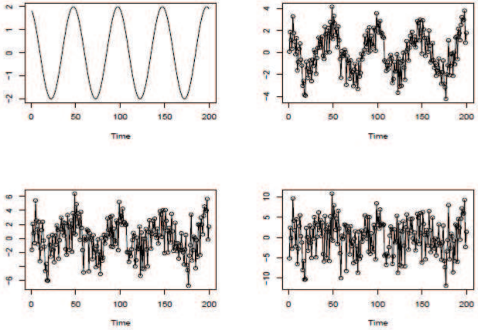


Figure 2.6: *Sinusoidal model illustration. Top left:* $E(Y_t) = 2\sin(2\pi t/50 + 0.6\pi)$. *The other plots are simulated realizations of this process* ($n = 200$) *with* $\sigma_e^2 = 1$ *(top right),* $\sigma_e^2 = 4$ *(bottom left), and* $\sigma_e^2 = 16$ *(bottom right).*

SAS programming for a sinusoidal model

Data figure26_1;
do t=1 to 200;
$y = 2 * sin(2 * 3.14159 * t/50 + 0.6 * 3.14159);$
output; end; run;

Proc gplot data= figure26_1;
symbol i=spline v=circle h=1.5;
plot y * t;
title 'the curve of $2sin(2\pi t/50 + 0.6\pi)$';
run;

2.4. Stationarity

Stationarity is a very important concept in the analysis of time series data.

Broadly speaking, a time series is said to be stationary

- if there is no systematic change in mean (no trend),
- if there is no systematic change in variance, and
- if strictly periodic variations have been removed.

In other words, the properties of one section of the data are much like those of any other section.

Importance of Stationarity

IMPORTANCE: Much of the theory of time series is concerned with stationary time series.

- For this reason, time series analysis often requires one to transform a non-stationary time series into a stationary one as to use this theory.

- For example, it may be of interest to remove the trend and seasonal variation from a set of data and then try to model the variation in the residuals by means of a stationary stochastic process.

2.4.1. Strict stationarity

- The stochastic process $\{Y_t : t = 0, 1, 2, \cdots, n\}$ is said to be strictly stationary if the joint distribution of

$$Y_{t_1}, Y_{t_2}, \cdots, Y_{t_n}$$

is the same as

$$Y_{t_1-k}, Y_{t_2-k}, \cdots, Y_{t_n-k}$$

for all time points $t_1, t_2, \cdots, t_n$ and for all time lags k.

- In other words, shifting the time origin by an amount k has no effect on the joint distributions, which must therefore depend only on the intervals between $t_1, t_2, \cdots, t_n$. This is a very strong condition.

Implication of strict stationarity

- When $n = 1$, this implies Y_t and Y_{t-k} have the same marginal distribution for all t and k.
- For all t and k,

$$\mathsf{E}(Y_t) = \mathsf{E}(Y_{t-k})$$
$$\mathsf{var}(Y_t) = \mathsf{var}(Y_{t-k}).$$

- Therefore, for a strictly stationary process, both $\mu_t = \mathsf{E}(Y_t)$ and $\gamma_{t,t} = \mathsf{var}(Y_t)$ are constant over time.

Implication of strict stationarity

- When $n = 2$, this implies (Y_t, Y_s) and (Y_{t-k}, Y_{s-k}) have the same joint distribution for all t, s, and k.
- For all t, s and k,

$$\mathrm{Cov}(Y_t, Y_s) = \mathrm{Cov}(Y_{t-k}, Y_{s-k}).$$

- For any s, t,

$$\gamma_{t,s} = \mathrm{Cov}(Y_t, Y_s) = \mathrm{Cov}(Y_0, Y_{|t-s|}) = \gamma_{0,|t-s|}.$$

  This means that the covariance between Y_t and Y_s does not depend on the actual values of t and s; it only depends on the time difference $|t - s|$.

New notation

- For a (strictly) stationary process, the covariance $\gamma_{t,s}$ depends only on the time difference $|t-s|$.
- The covariance between Y_t and any observation $k=|t-s|$ time points from it only depends on the lag k.
- Therefore, we write

$$\gamma_k = \mathsf{Cov}(Y_t, Y_{t-k})$$
$$\rho_k = \mathsf{corr}(Y_t, Y_{t-k}).$$

- We use this simpler notation only when we refer to a process which is stationary. Thay are

$$\gamma_0 = \mathsf{Cov}(Y_t, Y_t) = \mathsf{var}(Y_t),$$
$$\rho_k = \mathsf{corr}(Y_t, Y_{t-k}) = \frac{\gamma_k}{\gamma_0}.$$

Necessary conditions of strict stationarity

For a (strictly) stationary process,

1. the mean function $\mu_t = \mathsf{E}(Y_t)$ is constant throughout time; i.e., μ_t is free of t.

2. the covariance between any two observations depends only the time lag between them; i.e., $\gamma_{t,t-k} \equiv \gamma_k$ depends only on k.

2.4.2. Weak stationarity

A weak form of stationarity:

Definition: The stochastic process $\{Y_t : t = 0, 1, 2, \cdots, n\}$ is said to be weakly stationary (or **second-order stationary**) if

1. The mean function $\mu_t = \mathrm{E}(Y_t)$ is constant throughout time; i.e., μ_t is free of t.
2. The covariance between any two observations depends only the time lag between them; i.e., $\gamma_{t,t-k}$ depends only on k.

Remarks for weak stationarity

- Nothing is assumed about the collection of joint distributions of the process. Instead, we only are specifying the characteristics of the first two moments of the process.

- Strict stationarity implies weak stationarity. It is also clear that the converse to statement is not true.

- The additional assumption of multivariate normality (for the Y_t process) is given, then

$$\text{weak stationarity} + \text{multivariate normality} \Longrightarrow \text{strict stationarity.}$$

- Convention: When the term "stationary process" is used in this courese, this is understood to mean that the process is weakly stationary.

A white noise process is stationary

- Suppose that $\{e_t\}$ is a white noise process with $\mathsf{E}(e_t) = \mu_e$ and $\mathsf{var}(e_t) = \sigma_e^2$, both constant (free of t).
- Clearly, the mean process $\mu_t = \mathsf{E}(e_t)$ is constant over time.
- In addition, the autocovariance function $\gamma_k = \mathsf{cov}(Y_t, Y_{t-k})$ is given by

$$\gamma_k = \begin{cases} \sigma_e^2, & k = 0 \\ 0, & k \neq 0, \end{cases}$$

  which is free of time t (i.e., γ_k depends only on k).
- Thus, a white noise process is stationary.

A random walk process is not stationary

- Suppose that $\{Y_t\}$ is a random walk process. That is,

$$Y_t = Y_{t-1} + e_t.$$

- $\mu_t = \mathsf{E}(Y_t) = 0$, for all t, which is constant over time.
- However,

  $$\mathsf{Cov}(Y_t, Y_{t-k}) = \mathsf{Cov}(e_1 + \cdots + e_t, e_1 + \cdots + e_{t-k}) = (t-k)\sigma_e^2,$$

  which clearly depends on time t.
- Thus, a random walk process is not stationary.

A random walk with drift process is not stationary

- Suppose that $\{Y_t\}$ is a random walk with drift process; that is,

$$Y_t = \theta_0 + Y_{t-1} + e_t.$$

- $\mu_t = \mathsf{E}(Y_t) = t\theta_0$, which clearly is not free of time t.
- Additionally, $\mathsf{cov}(Y_t, Y_{t-k}) = (t-k)\sigma_e^2$ remains unchanged.
- Thus, a random walk with drift process is not stationary.

A moving average process is stationary

- Suppose that $\{Y_t\}$ is a moving average process given by

$$Y_t = \frac{1}{3}\left(e_t + e_{t-1} + e_{t-2}\right).$$

- We calculated $\mu_t = \mathsf{E}(Y_t) = 0$ (which is free of t) and $\gamma_k = \mathsf{cov}(Y_t, Y_{t-k})$ to be

$$\gamma_k = \begin{cases} \sigma_e^2/3, & k = 0 \\ 2\sigma_e^2/9, & k = 1 \\ \sigma_e^2/9, & k = 2 \\ 0, & k > 2. \end{cases}$$

- Because $\mathsf{Cov}(Y_t, Y_{t-k})$ is free of time t, this moving average process is stationary.

An autoregressive process is stationary

- Suppose that $\{Y_t\}$ is the autoregressive process

$$Y_t = 0.7Y_{t-1} + e_t.$$

- We avoided the calculation of $\mu_t = \mathsf{E}(Y_t)$ and $\mathsf{Cov}(Y_t, Y_{t-k})$ for this process, so we will not make a definite determination here.

- It turns out that if e_t is independent of $Y_{t-1}, Y_{t-2}, \cdots$, and if $\sigma_e^2 > 0$, then this autoregressive process is stationary (details coming later).

A sinusoidal process is not stationary

- Suppose that $\{Y_t\}$ is the sinusoidal process defined by

$$Y_t = a\,\mathsf{sin}(2\pi\omega t + \phi) + e_t.$$

- Clearly $\mu_t = \mathsf{E}(Y_t) = a\,\mathsf{sin}(2\pi\omega t + \phi)$ is not free of t, so this sinusoidal process is not stationary.

A random cosine wave process is stationary

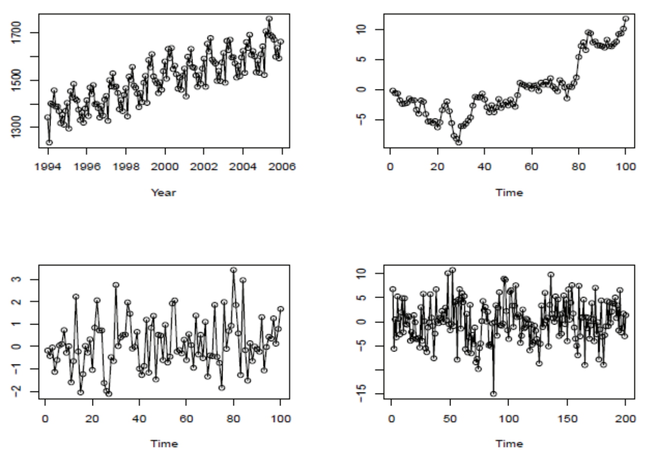
- Consider the random cosine wave process

$$Y_t = \cos\left[2\pi\left(\frac{t}{12} + \Phi\right)\right],$$

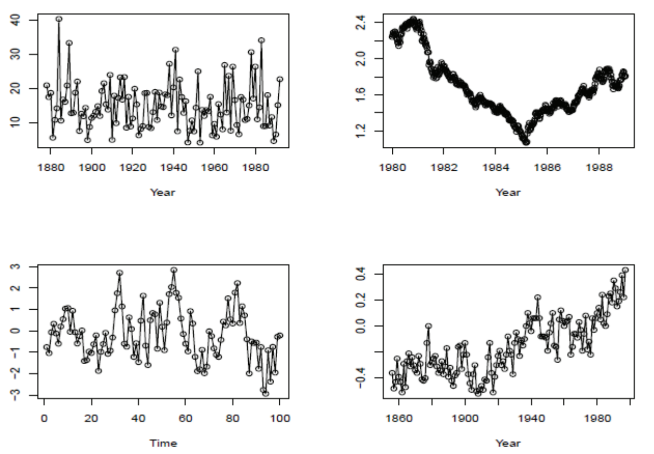
where Φ is a uniform random variable from 0 to 1; i.e., $\Phi \sim U(0,1)$.

- The calculations on pp 18-19 (CC) show that this process is (perhaps unexpectedly) stationary.

Which plots appear stationary?



Which plots appear stationary?



Remarks for stationary processes

Importance for stationary processes:

- In order to start thinking about viable stationary time series models for real data, we need to have a stationary process.
- However, as we have just seen, many data sets exhibit non-stationary behavior.
- A simple, but effective, technique to convert a non-stationary process into a stationary one is to examine data differences.

Transforming a non-stationary process to a stationary one

- Definition: Consider the process $\{Y_t : t = 0, 1, 2, \cdots, n\}$. The (first) difference process of $\{Y_t\}$ is defined by

$$\nabla Y_t = Y_t - Y_{t-1}, \text{ for } t = 1, 2, \cdots, n.$$

- In many realistic examples, a non-stationary process can be transformed into a stationary process by taking (first-order) differences.

- For example, the random walk $Y_t = Y_{t-1} + e_t$ is not stationary. However, it follows immediately that $\nabla Y_t = Y_t - Y_{t-1}$ is stationary. For a visual depiction, see Figure 2.7.

Figure 2.7. non-stationary RW and its stationary difference

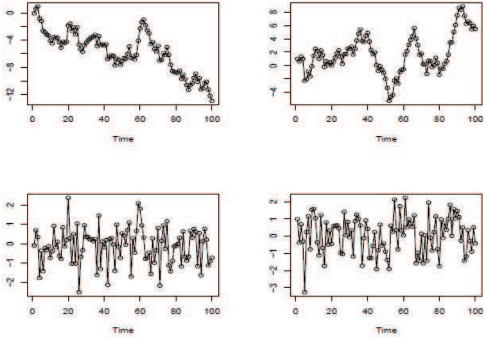


Figure 2.7: *The top two processes are each random walks* $Y_t = Y_{t-1} + e_t$, *where* $e_t \sim$ *iid* $\mathcal{N}(0, 1)$. *The bottom processes are the difference processes, respectively.*

Thank you for your attention !