Lecture 2: Fundamental concepts

Introduction and recalls...

Alain Latour[1][2]

[1]Département d'informatique, mathématiques et sciences sociales
Université Pierre-Mendès-France
Grenoble, France

[2]Département de mathématiques
Université du Québec à Montréal
1 Montréal, Canada

Grenoble – January 2015

LABORATOIRE JEAN-KUNTZMANN

Grenoble INP Ensimag

INFÉRENCE PROCESSUS STOCHASTIQUES

Outline of the lecture

Grenoble INP Ensimag

Notation

Observation of a process

- The sequence of random variables $\{Y_t\}_{t\in\mathbb{Z}} = \{Y_t : t = 0, \pm 1, \ \pm 2, \pm 3, \ldots\}$ is called a stochastic process
- $X(t)$: observation at time t.
- Observations at regular times:

$$
\begin{array}{cccc}
t_0 & t_0 + 1h & \cdots & t_0 + (N-1)h \\
\downarrow & \downarrow & & \downarrow \\
X(t_0) & X(t_0 + h) & \cdots & X(t_0 + (N-1)h) \\
\downarrow & \downarrow & & \downarrow \\
X_1 & X_2 & \cdots & X_N
\end{array}
$$

Grenoble INP Ensimag

Observations

A discrete time series can be obtained in two ways:

- By sampling at regular time intervals of a continuous series (value of shares on the stock exchange, concentration of a pollutant measured every hour)
- Accumulation of a variable over a period of time (number of accidents per week, total rainfall per month, the weekly average concentration of a pollutant).

Grenoble INP Ensimag

- The sequence of random variables $\{Y_t\}_{t\in\mathbb{Z}} = \{Y_t : t = 0, \pm 1, \ \pm 2, \pm 3, \dots\}$ is called a stochastic process
- It serves as a model for an observed time series.
- The complete probabilistic structure of such a process is determined by the set of distributions of all finite collections of the Y's.
- In practice, we do not have to deal explicitly with these multivariate distributions.
- Much of the information can be described in terms of means, variances, and covariances.

Outline of the lecture

Means, Variances, and Covariances

- For a stochastic process $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \dots\}$, the mean function is defined by

$$\mu_t = E(Y_t) \quad \text{for } t = 0, \pm 1, \pm 2, \dots \tag{1}$$

  - μ_t is just the expected value of the process at time t.
  - μ_t can be different at each time point t.
- The autocovariance function, $\gamma_{t,s}$, is defined as

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

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

Means, Variances, and Covariances

Correlations

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

$$\rho_{t,s} = Corr(Y_t, Y_s) \quad \text{for } t, \ s = 0, \pm 1, \pm 2, \dots \tag{3}$$

where

$$Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} \tag{4}$$

- Both covariance and correlation are measures of the (linear) dependence between random variables.

Means, Variances, and Covariances

Correlations

- Unitless correlation is somewhat easier to interpret.
- The following important properties follow from known results and our definitions:

$$\left.\begin{array}{ll} \gamma_{t,t} = Var(Y_t) & \rho_{t,t} = 1 \\ \gamma_{t,s} = \gamma_{s,t} & \rho_{t,s} = \rho_{s,t} \\ |\gamma_{t,s}| \leqslant \sqrt{\gamma_{t,t}\gamma_{s,s}} & |\rho_{t,s}| \leqslant 1 \end{array}\right\} \tag{5}$$

-
  - Values of $\rho_{t,s}$ near ± 1 indicate strong (linear) dependence
  - Values near zero indicate weak (linear) dependence.
  - If $\rho_{t,s} = 0$, we say that Y_t and Y_s are uncorrelated.

Computations with covariance

Covariance between two linear combinations

The following results will be used repeatedly:

- If $c_1, \ldots, c_m$ and $d_1, \ldots, d_n$ are constants and $t_1, \ldots, t_m$ and $s_1, \ldots, s_n$ are time points, then

$$Cov\left[\sum_{i=1}^{m} c_i Y_{t_i}, \sum_{j=1}^{n} d_j Y_{t_j}\right] = \sum_{i=1}^{m}\sum_{j=1}^{n} c_i d_j Cov(Y_{t_i}, Y_{t_j}) \tag{6}$$

- Variance as a special case:

$$Var\left[\sum_{i=1}^{n} c_i Y_{t_i}\right] = \sum_{i=1}^{n} c_i^2 Var(Y_{t_i}) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} c_i c_j Cov(Y_{t_i}, Y_{t_j}) \tag{7}$$

An example

The random walk

- Let $\{e_t\}$ be a sequence of independent, identically distributed random variables each with zero mean and variance σ_e^2. The observed time series, $\{Y_t : t = 1, 2, \ldots\}$, is constructed as follows:

$$\begin{aligned} Y_1 &= e_1 \\ Y_2 &= e_1 + e_2 \\ &\vdots \\ Y_t &= e_1 + e_2 + \cdots + e_t \end{aligned} \tag{8}$$

Alternatively, we can write

$$Y_t = Y_{t-1} + e_t \tag{9}$$

with "initial condition" $Y_1 = e_1$.

An example

The random walk

- From Equation (8), we obtain the mean function

$$\begin{aligned} \mu_t &= E(Y_t) = E(e_1 + e_2 + \cdots + e_t) \\ &= E(e_1) + E(e_2) + \cdots + E(e_t) \\ &= 0 + 0 + \cdots + 0 \\ \mu_t &= 0 \quad \text{for all } t \end{aligned} \tag{10}$$

- Also

$$\begin{aligned} Var(Y_t) &= Var(e_1 + e_2 + \cdots + e_t) \\ &= Var(e_1) + Var(e_2) + \cdots + Var(e_t) \\ &= \sigma_e^2 + \sigma_e^2 + \cdots + \sigma_e^2 \\ Var(Y_t) &= t\sigma_e^2 \end{aligned}$$

An example
The random walk

- To investigate the covariance function, suppose that $1 \leqslant t \leqslant s$.

$$\begin{aligned}\gamma_{t,s} &= Cov(Y_t, Y_s)\\ &= Cov(e_1 + e_2 + \cdots + e_t, e_1 + e_2 + \cdots + e_t + e_{t+1} + \cdots + e_s)\end{aligned}$$

From Equation (6), we have

$$\gamma_{t,s} = \sum_{i=1}^{s}\sum_{j=1}^{t} Cov(e_i, e_j) = t\sigma_e^2$$

Covariances are zero unless $i = j$, in which case $Var(e_i) = \sigma_e^2$.

An example
The random walk

- Since $\gamma_{t,s} = \gamma_{s,t}$, this specifies the autocovariance function for all time points t and s and

$$\gamma_{t,s} = t\sigma_e^2 \quad \text{for } 1 \leqslant t \leqslant s \tag{11}$$

- The autocorrelation function for the random walk is

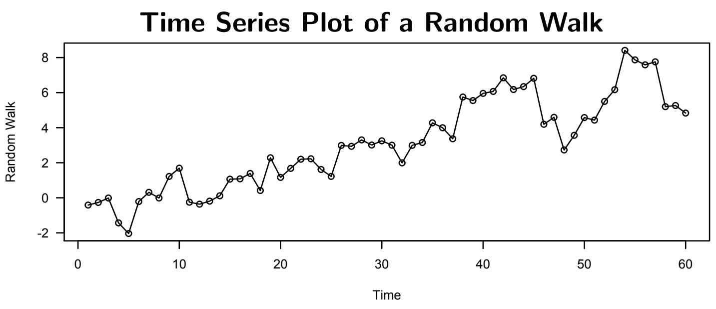
$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \sqrt{\frac{t}{s}} \tag{12}$$

- Numerical values:

$$\rho_{1,2} = \sqrt{\frac{1}{2}} = 0,707 \qquad \rho_{8,9} = \sqrt{\frac{8}{9}} = 0,943$$

$$\rho_{24,25} = \sqrt{\frac{24}{25}} = 0,980 \qquad \rho_{1,25} = \sqrt{\frac{1}{25}} = 0,200$$

An example
The random walk




```
> quartz(width=6, height=2.75, pointsize=7.5, title="Exhibit 2.1")
> data(rwalk) # rwalk contains a simulated random walk
> plot(rwalk, type='o', ylab='Random Walk', las=1)
```


Another example
A moving average

- Suppose that $\{Y_t\}$ is constructed as

$$Y_t = \frac{e_t + e_{t-1}}{2} \tag{13}$$

where the e's are assumed to be independent and identically distributed with zero mean and variance σ_e^2.

- Mean and variance

$$\mu_t = \mathrm{E}[Y_t] = \mathrm{E}\left[\frac{e_t + e_{t-1}}{2}\right] = \frac{\mathrm{E}[e_t] + \mathrm{E}[e_{t-1}]}{2} = 0$$

$$\mathrm{Var}[Y_t] = \mathrm{Var}\left[\frac{e_t + e_{t-1}}{2}\right] = \frac{\mathrm{Var}[e_t] + \mathrm{Var}[e_{t-1}]}{4} = \frac{1}{2}\sigma_e^2$$

Another example

A moving average

- Covariance at lag 1

$$
\begin{aligned}
&\mathrm{Cov}[Y_t, Y_{t-1}] \\
&= \mathrm{Cov}\left[\frac{e_t + e_{t-1}}{2}; \frac{e_{t-1} + e_{t-2}}{2}\right] \\
&= \frac{\mathrm{Cov}[e_t, e_{t-1}] + \mathrm{Cov}[e_t, e_{t-2}] + \mathrm{Cov}[e_{t-1}, e_{t-1}] + \mathrm{Cov}[e_{t-1}, e_{t-2}]}{4} \\
&= \frac{\mathrm{Cov}[e_{t-1}, e_{t-1}]}{4} = \frac{1}{4}\sigma_e^2
\end{aligned}
$$

or

$$\gamma_{t,t-1} = \frac{1}{4}\sigma_e^2, \quad \text{for all } t \tag{14}$$

Another example

A moving average

- Covariance at lag 2

$$
\begin{aligned}
\mathrm{Cov}[Y_t, Y_{t-2}] &= \mathrm{Cov}\left[\frac{e_t + e_{t-1}}{2}; \frac{e_{t-2} + e_{t-3}}{2}\right] \\
&= 0 \qquad \text{since the } e\text{'s are independent}
\end{aligned}
$$

- Covariance at lag k: $\mathrm{Cov}[Y_t, Y_{t-k}] = 0$ for $k > 1$,
- In general

$$\gamma_{t,s} = \begin{cases} \frac{1}{2}\sigma_e^2 & \text{for } |t-s| = 0 \\ \frac{1}{4}\sigma_e^2 & \text{for } |t-s| = 1 \\ 0 & \text{for } |t-s| > 1 \end{cases}$$

Another example

A moving average

- Obvioulsy for the autocorrelation function, we have

$$\rho_{t,s} = \begin{cases} 1 & \text{for } |t-s| = 0 \\ \frac{1}{2} & \text{for } |t-s| = 1 \\ 0 & \text{for } |t-s| > 1 \end{cases} \tag{15}$$

  since $\frac{1}{4}\sigma_e^2 / \frac{1}{2}\sigma_e^2 = \frac{1}{2}$.
- Notice:
  - Notice that $\rho_{2,1} = \rho_{3,2} = \rho_{4,3} = \rho_{9,8} = \frac{1}{2}$.
  - Values of Y precisely one time unit apart have exactly the same correlation no matter where they occur in time.
  - Furthermore, $\rho_{t,t-k}$ is the same for all values of t.

Outline of the lecture

Notions of stationarity

Strict stationarity

- The basic idea of stationarity is that the probability laws that govern the behavior of the process do not change over time.
- From a mathematical point of view:
  - Let $n \in \mathbb{N}$
  - Let $t_1, \ldots, t_n \in \mathbb{Z}$
  - Let $k \in \mathbb{Z}$
- $\{Y_t\}$ is **strictly stationary** if

$$(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}) \stackrel{Law}{=} (Y_{t_1-k}, Y_{t_2-k}, \ldots, Y_{t_n-k})$$

  for all choices of time points $t_1, \ldots, t_n$ and all choices of time lag k.

Notions of stationarity

Strict stationarity + second moments

- When $n = 1$

$$Y_t \stackrel{Law}{=} Y_{t-k}, \quad \forall t \text{ and } \forall k.$$

$$\Longrightarrow E(Y_t) = E(Y_{t-k})$$

  for all t and k
- The mean function is constant for all time.
- Additionally, $Var(Y_t) = Var(Y_{t-k})$ for all t and k so that the variance is also constant over time.

Notions of stationarity

Strict stationarity + second moments

- When $n = 2$

$$(Y_t, Y_s) \stackrel{Law}{=} (Y_{t-k}, Y_{s-k}), \quad \forall t, \forall s \text{ and } \forall k.$$

$$\Longrightarrow \mathrm{Cov}[Y_t, Y_s] = \mathrm{Cov}[Y_{t-k}, Y_{s-k}]$$

  for all t, s and k
- Putting $k = s$ and then $k = t$, we obtain

$$\begin{aligned}\gamma_{t,s} &= \mathrm{Cov}[Y_{t-s}, Y_0] \\ &= \mathrm{Cov}[Y_0, Y_{t-s}] \\ &= \mathrm{Cov}\left[Y_0, Y_{|t-s|}\right] \\ &= \gamma_{0,|t-s|}\end{aligned}$$

- The covariance between Y_t and Y_s depends on the time difference $|t-s|$ and not otherwise on the actual times t and s.

Notions of stationarity

Strict stationarity + second moments

- Thus, for a stationary process, we can simplify our notation and write

$$\gamma_k = \mathrm{Cov}[Y_t, Y_{t-k}] \qquad \text{and} \qquad \rho_k = \mathrm{Corr}[Y_t, Y_{t-k}]$$

  Note also that

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

  The general properties given in Equation (5) now become

$$\left.\begin{array}{ll} \gamma_0 = Var(Y_t) & \rho_0 = 1 \\ \gamma_k = \gamma_{-k} & \rho_k = \rho_{-k} \\ |\gamma_k| \leqslant \gamma_0 & |\rho_k| \leqslant 1 \end{array}\right\} \tag{16}$$

Notions of stationarity

Weak stationarity

- A stochastic process $\{Y_t\}$ is said to be **weakly** (or **second-order**) **stationary** if
  1. The mean function is constant over time, and
  2. $\gamma_{t,t-k} = \gamma_{0,k}$ for all time t and lag k.
- For stationary processes, we usually only consider $k \geqslant 0$.

White Noise

A second order stationary process

- A strong **white noise** process is a sequence of independent, identically distributed random variables $\{e_t\}$ with mean 0 and variance σ_e^2. *Such a process is a strictly stationary process.*
- A weak **white noise** process is a sequence of uncorrelated, random variables $\{e_t\}$ with mean 0 and variance σ_e^2. *Such a process is at least a weakly stationary process.*
- We have

$$\gamma_k = \begin{cases} \mathrm{Var}[e_t] = \sigma_e^2 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

Alternatively, we can write

$$\rho_k = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

Another example

Moving average process

- The moving average $Y_t = (e_t + e_{t-1})/2$, is another example of a stationary process constructed from white noise.
- We have for the moving average process that

$$\rho_k = \begin{cases} 1 & k = 0 \\ 0.5 & |k| = 1 \\ 0 & |k| \geqslant 2 \end{cases}$$

Random Cosine Wave

As a somewhat different example

- Consider the process defined as follows:

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

where Φ is selected (once) from a uniform distribution on the interval from 0 to 1.
- This is a stationary process. Amazing?

Simulation experiment I

In R

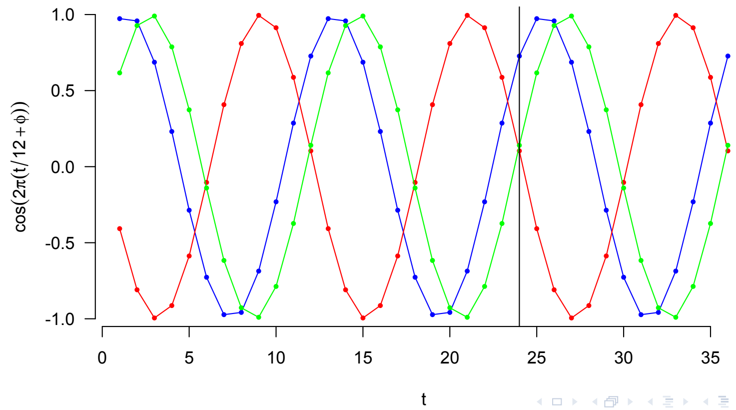
```
t <- 1:36
phi <- runif(1)
y <- cos(2*pi*(t/12+phi))
plot(t,y,type="o",pch=19,cex=0.5,col="blue",las=1,ylab=expression(cos(2*pi*(t/12+phi)))↘
→,frame=FALSE)

phi <- runif(1)
y <- cos(2*pi*(t/12+phi))
lines(t,y,col="red")
points(t,y,pch=19,cex=0.5,col="red")

phi <- runif(1)
y <- cos(2*pi*(t/12+phi))
lines(t,y,col="green")
points(t,y,pch=19,cex=0.5,col="green")
abline(v=24)
```


Simulation experiment I

In R



Random Cosine Wave

Moments

- Expected value:

$$
\begin{aligned}
E(Y_t) &= E\left\{\cos\left[2\pi\left(\frac{t}{12}+\Phi\right)\right]\right\} \\
&= \int_0^1 \cos\left[2\pi\left(\frac{t}{12}+\phi\right)\right]d\phi \\
&= \frac{1}{2\pi}\sin\left[2\pi\left(\frac{t}{12}+\phi\right)\right]\Bigg|_{\phi=0}^{1} \\
&= \frac{1}{2\pi}\left[\sin\left(2\pi\frac{t}{12}+2\pi\right)-\sin\left(2\pi\frac{t}{12}\right)\right]
\end{aligned}
$$

But this is zero since the sines must agree. So $\mu_t = 0$ for all t.

Random Cosine Wave

Moments

- Variance:

$$
\begin{aligned}
\gamma_{t,s} &= E\left\{\cos\left[2\pi\left(\frac{t}{12}+\Phi\right)\right]\cos\left[2\pi\left(\frac{s}{12}+\Phi\right)\right]\right\} \\
&= \int_0^1 \cos\left[2\pi\left(\frac{t}{12}+\phi\right)\right]\cos\left[2\pi\left(\frac{s}{12}+\phi\right)\right]d\phi
\end{aligned}
$$

using $\cos(A-B)+\cos(A+B) = 2\cos(A)\cos(B)$

$$
\begin{aligned}
&= \frac{1}{2}\int_0^1\left\{\cos\left[2\pi\left(\frac{t-s}{12}\right)\right]+\cos\left[2\pi\left(\frac{t+s}{12}+2\phi\right)\right]\right\}d\phi \\
&= \frac{1}{2}\left\{\cos\left[2\pi\left(\frac{t-s}{12}\right)\right]+\frac{1}{4\pi}\sin\left[2\pi\left(\frac{t+s}{12}+2\phi\right)\right]\Bigg|_{\phi=0}^{1}\right\} \\
&= \frac{1}{2}\cos\left[2\pi\left(\frac{|t-s|}{12}\right)\right]
\end{aligned}
$$

Random Cosine Wave

Moments

- So the process is stationary with autocorrelation function

$$\rho_k = \cos\left(2\pi \frac{k}{12}\right), \qquad \text{for } k = \pm 1, \ \pm 2, \ \dots$$

Simulation experiment II

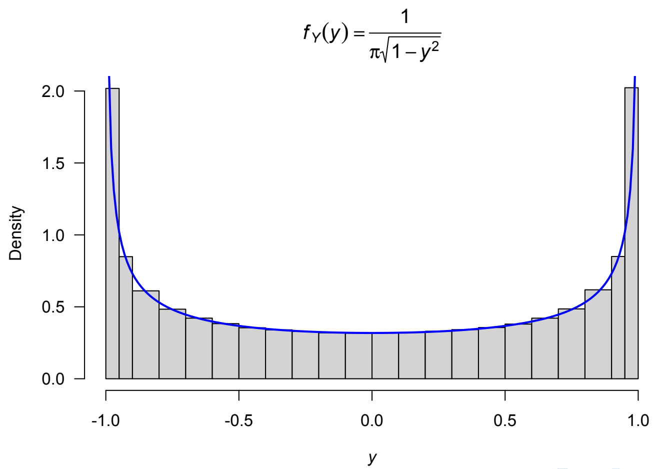
In R

```r
rm(list=ls())
t <- 1:36
nsim <- 1000000
y <- rep(NA, nsim)
for(i in 1:nsim){
  phi <- runif(1)
  y[i] <- cos(2*pi*(t/12+phi))[24]
}

hy <- hist(y, prob=TRUE,
        breaks=c(-1,-0.95,seq(-0.9,0.9,.1),0.95,1),
        col="lightgrey",
        main=expression(italic(f[Y](y)==frac(1,pi*sqrt(1-y^2)))),
        xlab=expression(italic(y)),
        las=1
        )
a <- seq(-1,1,0.01)
z <- 1/(sqrt(1-a^2)*pi)
lines(a,z, col="blue", lwd=2)
```

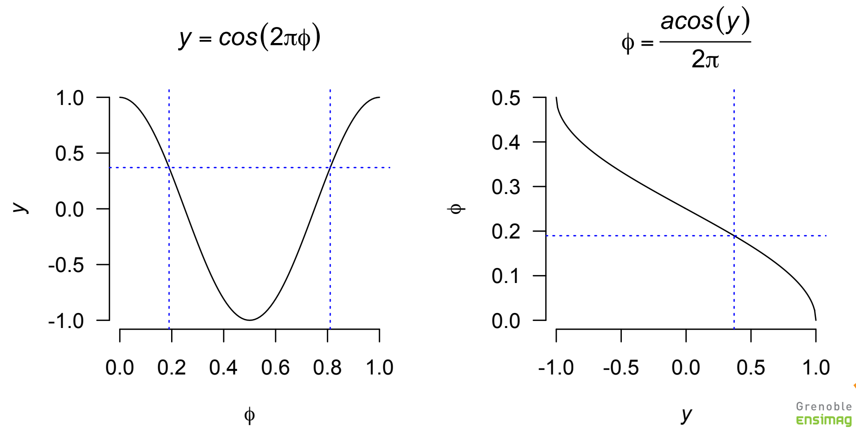

Simulation experiment II

In R



Distribution of Y_t

Suppose $t = 0$ to simplify the computations



Distribution of Y_t

Suppose $t = 0$ to simplify the computations

- We have:

$$
\begin{aligned}
F_Y(y) &= \Pr(Y_0 \leqslant y) \\
&= \Pr\left(\cos(2\pi\Phi) \leqslant y\right) \\
&= 2\Pr\left(\frac{\arccos(y)}{2\pi} \leqslant \Phi \leqslant \frac{1}{2}\right) \\
&= \frac{1}{2} - \frac{\arccos(y)}{2\pi} \\
f_Y(y) &= \frac{1}{2\,\pi\,\sqrt{1-y^2}}
\end{aligned}
$$

References

Cryer, J. D. and K.-S. Chan (2008). *Time Series Analysis: With Applications in R* (2nd edition ed.). Springer texts in statistics. New York: Springer.

Lafaye de Micheaux, P., R. Drouilhet, and B. Liquet (2011). *Le logiciel R: Maîtriser le langage - Effectuer des analyses statistiques (Statistique et probabilités appliquées) (French Edition)* (1 ed.). Springer.

Shumway, R. H. and D. S. Stoffer (2010). *Time Series Analysis and Its Applications: With R Examples* (3rd ed. edition ed.). Springer texts in statistics. New York: Springer.

Tsay, R. (2010). *Analysis of Financial Time Series* (3 ed.). John Wiley Sons Inc.