Chapter 1

Time Series Concepts

Updated: January 22, 2015.

This chapter reviews some basic times series concepts that are important for describing and modeling financial time series.

1.1 Stochastic Processes

A *stochastic process*

$$\{\ldots, Y_1, Y_2, \ldots, Y_t, Y_{t+1}, \ldots\} = \{Y_t\}_{t=-\infty}^{\infty},$$

is a sequence of random variables indexed by time t[1]. In most applications, the time index is a regularly spaced index representing calendar time (e.g., days, months, years, etc.). In modeling time series data, the ordering imposed by the time index is important because we often would like to capture the temporal relationships, if any, between the random variables in the stochastic process. In random sampling from a population, the ordering of the random variables representing the sample does not matter because they are independent.

A realization of a stochastic process with T observations is the sequence of observed data

$$\{Y_1 = y_1, Y_2 = y_2, \ldots, Y_T = y_T\} = \{y_t\}_{t=1}^{T}.$$

[1] To conserve on notation, we will often represent the stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ simply as $\{Y_t\}$.

The goal of time series modeling is to describe the probabilistic behavior of the underlying stochastic process that is believed to have generated the observed data in a concise way. In addition, we want to be able to use the observed sample to estimate important characteristics of a time series model such as measures of time dependence. In order to do this, we need to make a number of assumptions regarding the joint behavior of the random variables in the stochastic process such that we may treat the stochastic process in much the same way as we treat a random sample from a given population.

1.1.1 Stationary Stochastic Processes

We often describe random sampling from a population as a sequence of independent, and identically distributed (iid) random variables $X_1, X_2 \ldots$ such that each X_i is described by the same probability distribution F_X, and write $X_i \sim F_X$. With time series data, we would like to preserve the identical distribution assumption but we do not want to impose the restriction that each random variable in the sequence is independent of all of the other variables. In many contexts, we would expect some dependence between random variables close together in time (e.g, X_1, and X_2) but little or no dependence between random variables far apart in time (e.g., X_1 and X_{100}). We can allow for this type of behavior using the concepts of *stationarity* and *ergodicity*.

We start with the definition of strict stationarity.

Definition 1 *Strict stationarity*

A stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is *strictly stationary* if, for any given finite integer r and for any set of subscripts $t_1, t_2, \ldots, t_r$ the joint distribution of $(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_r})$ depends only on $t_1 - t, t_2 - t, \ldots, t_r - t$ but not on t. ■

In simple terms, the joint distribution of random variables in a strictly stationary stochastic process is time invariant. For example, the joint distribution of (Y_1, Y_5, Y_7) is the same as the distribution of (Y_{12}, Y_{16}, Y_{18}). Just like in an iid sample, in a strictly stationary process all of the random variables Y_t $(t = -\infty, \ldots, \infty)$ have the same marginal distribution F_Y. This means they all have the same mean, variance etc., assuming these quantities exist. However, assuming strict stationarity does not make any assumption about the correlations between $Y_t, Y_{t_1}, \ldots, Y_{t_r}$ other than that the correlation between Y_t and Y_{t_r} only depends on $t - t_r$ (the time between Y_t and Y_{t_r}) and not on t. That is, strict stationarity allows for general temporal dependence between the random variables in the stochastic process.

A useful property of strict stationarity is that it is preserved under general transformations, as summarized in the following proposition.

Proposition 2 *let $\{Y_t\}_{t=-\infty}^{\infty}$ be strictly stationary and let $g(\cdot)$ be any function of the elements in $\{Y_t\}_{t=-\infty}^{\infty}$. Then $\{g(Y_t)\}_{t=-\infty}^{\infty}$, is also strictly stationary.*

For example, if $\{Y_t\}_{t=-\infty}^{\infty}$ is strictly stationary then $\{Y_t^2\}_{t=-\infty}^{\infty}$ and $\{Y_tY_{t-1}\}_{t=-\infty}^{\infty}$ are also strictly stationary.

Example 3 *iid sequence*

If $\{Y_t\}_{t=-\infty}^{\infty}$ is an iid sequence, then it is strictly stationary. ■

Example 4 *Non iid sequence*

Let $\{Y_t\}_{t=-\infty}^{\infty}$ be an iid sequence and let $X \sim N(0,1)$ independent of $\{Y_t\}_{t=-\infty}^{\infty}$. Define $Z_t = Y_t + X$. The sequence $\{Z_t\}_{t=-\infty}^{\infty}$ is not an independent sequence (because of the common X) but is an identically distributed sequence and is strictly stationary. ■

If we assume that the stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is strictly stationary and that $E[Y_t]$, $\mathrm{var}(Y_t)$, and all pairwise covariances exist, then we say that $\{Y_t\}_{t=-\infty}^{\infty}$ is a *covariance stationary* stochastic process.

Definition 5 *Covariance stationarity*

A stochastic process $\{Y_t\}_{t=1}^{\infty}$ is *covariance stationary* if

1. $E[Y_t] = \mu$ does not depend on t

2. $\mathrm{var}(Y_t) = \sigma^2$ does not depend on t

3. $\mathrm{cov}(Y_t, Y_{t-j}) = \gamma_j$ exists, is finite, and depends only on j but not on t for $j = 0, 1, 2, \ldots$

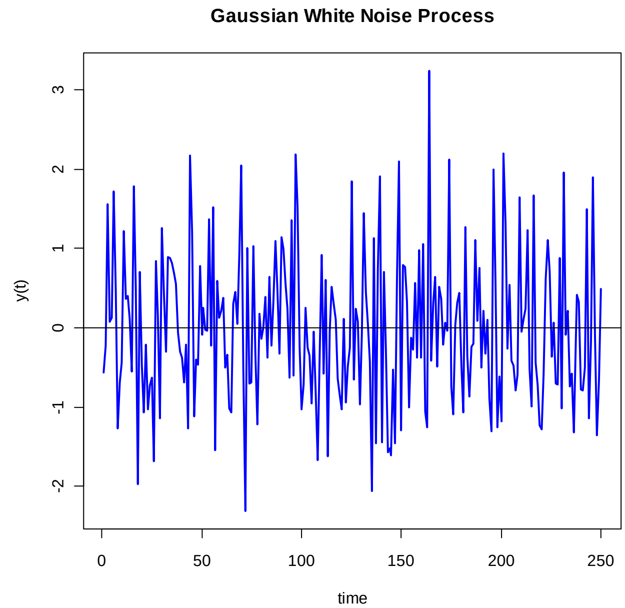

Figure 1.1: Realization of a GWN(0,1) process.

The term γ_j is called the jth order *autocovariance.* The jth order *autocorrelation* is defined as

$$\rho_j = \frac{\mathrm{cov}(Y_t, Y_{t-j})}{\sqrt{\mathrm{var}(Y_t)\mathrm{var}(Y_{t-j})}} = \frac{\gamma_j}{\sigma^2}. \tag{1.1}$$

The autocovariances, γ_j, measure the direction of linear dependence between Y_t and Y_{t-j}. The autocorrelations, ρ_j, measure both the direction and strength of linear dependence between Y_t and Y_{t-j}.

The autocovariances and autocorrelations are measures of the linear temporal dependence in a covariance stationary stochastic process. A graphical summary of this temporal dependence is given by the plot of ρ_j against j, and is called the *autocorrelation function* (ACF).

Example 6 *Gaussian White Noise*

Let $Y_t \sim iid\ N(0, \sigma^2)$. Then $\{Y_t\}_{t=-\infty}^{\infty}$ is called a *Gaussian white noise* process and is denoted $Y_t \sim \text{GWN}(0, \sigma^2)$. Notice that

$$\begin{aligned} E[Y_t] &= 0 \text{ independent of } t, \\ \text{var}(Y_t) &= \sigma^2 \text{ independent of } t, \\ \text{cov}(Y_t, Y_{t-j}) &= 0 \text{ (for } j > 0) \text{ independent of } t \text{ for all } j, \end{aligned}$$

so that $\{Y_t\}_{t=-\infty}^{\infty}$ is covariance stationary. Figure 1.1 shows a realization of a GWN(0,1) process. The defining characteristic of a GWN process is the lack of any predictable pattern over time in the realized values of the process as illustrated in 1.1. In the electrical engineering literature, white noise represents the absence of any signal.

Figure 1.1was created with the R commands:

```
> set.seed(123)
> y = rnorm(250)
> ts.plot(y, main="Gaussian White Noise Process", xlab="time",
+         ylab="y(t)", col="blue", lwd=2)
> abline(h=0)
```

The simulated iid N(0,1) values are generated using the `rnorm()` function. The command `set.seed(123)` initializes R's internal random number generator using the seed 123. Everytime the random number generator seed is set to a particular value, the random number generator produces the same set of random numbers. This allows different people to create the same set of random numbers so that results are reproducible. The function `ts.plot()` creates a time series line plot with a dummy time index. An equivalent plot can be created using the generic `plot()` function:

```
> plot(y, main="Gaussian White Noise Process", type="l", xlab="time",
+    ylab="y(t)", col="blue", lwd=2)
> abline(h=0)
```

■

Example 7 *Gaussian White Noise Model for Continuously Compounded Returns*

Let r_t denote the continuously compounded monthly return on Microsoft stock and assume that $r_t \sim \text{GWN}(0.01, (0.05)^2)$. 60 Simulated values of $\{r_t\}$ are computed using

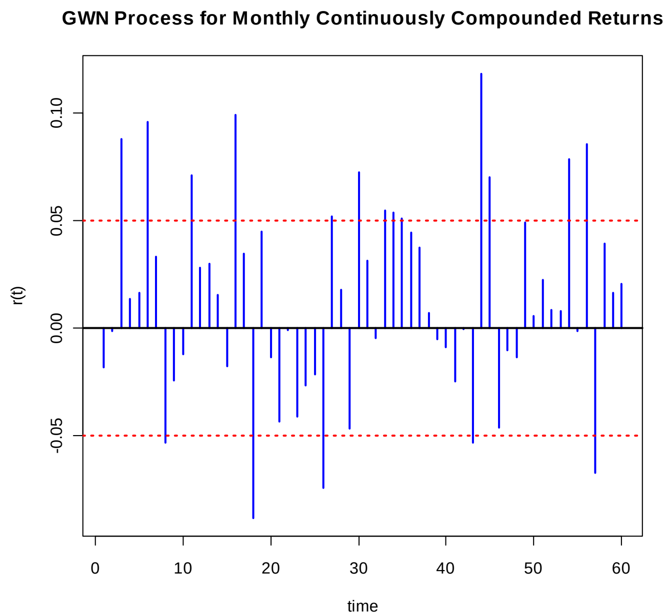

Figure 1.2: Simulated returns from GWN(0.01,(0.05)^2).

```
> set.seed(123)
> y = rnorm(60, mean=0.01, sd=0.05)
> ts.plot(y,main="GWN Process for Monthly Continuously Compounded Returns",
+         xlab="time",ylab="r(t)", col="blue", lwd=2, type="p")
> abline(h=c(0,-0.05,0.05), lwd=2, lty=c("solid","dotted","dotted"),
+        col=c("black", "red", "red"))
```

and are illustrated in Figure 1.2. Notice that the returns fluctuate around the mean value of 0.01, and the size of a typically deviation from the mean is about 0.05. An implication of the GWN assumption for monthly returns is that non-overlapping multiperiod returns are also GWN. For example, consider the two-month return $r_t(2) = r_t + r_{t-1}$. The non-overlapping process $\{r_t(2)\} = \{..., r_{t-2}(2), r_t(2), r_{t+2}(2), ...\}$ is GWN with mean $E[r_t(2)] = 2 \cdot (0.01) = 0.02$ and variance $\text{var}(r_t(2)) = 2 \cdot (0.05)^2$.

Example 8 *Independent White Noise*

Let $Y_t \sim$ iid $(0, \sigma^2)$. Then $\{Y_t\}_{t=-\infty}^{\infty}$ is called an independent white noise process and is denoted $Y_t \sim \text{IWN}(0, \sigma^2)$. The difference between GWN and IWN is that with IWN we don't specify that all random variables are normally distributed. The random variables can have any distribution with mean zero and variance σ^2.■

Example 9 *Weak White Noise*

Let $\{Y_t\}_{t=-\infty}^{\infty}$ be a sequence of uncorrelated random variables each with mean zero and variance σ^2. Then $\{Y_t\}_{t=-\infty}^{\infty}$ is called a *weak white noise* process and is denoted $Y_t \sim \text{WN}(0, \sigma^2)$. With a weak white noise process, the random variables are not independent, only uncorrelated. This allows for potential non-linear temporal dependence between the random variables in the process. ■

1.1.2 Non-Stationary Processes

In a covariance stationary stochastic process it is assumed that the means, variances and autocovariances are independent of time. In a non-stationary process, one or more of these assumptions is not true.

Example 10 *Deterministically trending process*

Suppose $\{Y_t\}_{t=0}^{\infty}$ is generated according to the deterministically trending process

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 t + \varepsilon_t, \ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2), \\ t &= 0, 1, 2, \ldots \end{aligned}$$

Then $\{Y_t\}_{t=0}^{\infty}$ is nonstationary because the mean of Y_t depends on t :

$$E[Y_t] = \beta_0 + \beta_1 t \ \text{ depends on } t.$$

Figure 1.3 shows a realization of this process with $\beta_0 = 0, \beta_1 = 0.1$ and $\sigma_\varepsilon^2 = 1$ created using the R commands:

```
> set.seed(123)
> e = rnorm(250)
> y.dt = 0.1*seq(1,250) + e
> ts.plot(y.dt, lwd=2, col="blue", main="Deterministic Trend + Noise")
> abline(a=0, b=0.1)
```

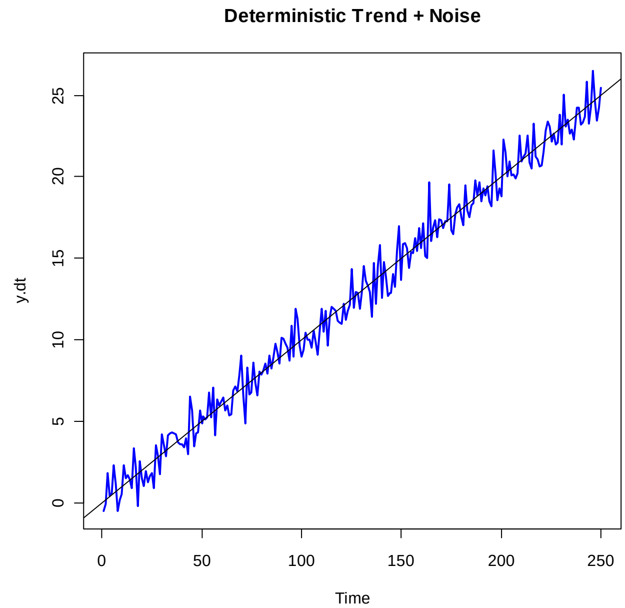

Figure 1.3: Deterministically trending nonstationary process $Y_t = 0.1t + \varepsilon_t, \varepsilon_t \sim N(0,1)$

Here the non-stationarity is created by the deterministic trend $\beta_0 + \beta_1 t$ in the data. The non-stationary process $\{Y_t\}_{t=0}^{\infty}$ can be transformed into a stationary process by simply subtracting off the trend:

$$X_t = Y_t - \beta_0 - \beta_1 t = \varepsilon_t \sim \mathrm{WN}(0, \sigma_\varepsilon^2).$$

The detrended process $X_t \sim$WN$(0, \sigma_\varepsilon^2)$. ■

Example 11 *Random walk*

A random walk (RW) process $\{Y_t\}_{t=1}^{\infty}$ is defined as

$$\begin{aligned} Y_t &= Y_{t-1} + \varepsilon_t,\ \varepsilon_t \sim \mathrm{IWN}(0, \sigma_\varepsilon^2), \\ & Y_0 \text{ is fixed (non-random).} \end{aligned}$$

By recursive substitution starting at $t = 1$, we have

$$\begin{aligned} Y_1 &= Y_0 + \varepsilon_1, \\ Y_2 &= Y_1 + \varepsilon_2 = Y_0 + \varepsilon_1 + \varepsilon_2, \\ &\vdots \\ Y_t &= Y_0 + \varepsilon_1 + \cdots + \varepsilon_t \\ &= Y_0 + \sum_{j=1}^{t} \varepsilon_j. \end{aligned}$$

Now, $E[Y_t] = Y_0$ which is independent of t. However,

$$\text{var}(Y_t) = \text{var}\left(\sum_{j=1}^{t} \varepsilon_j\right) = \sum_{j=1}^{t} \sigma_\varepsilon^2 = \sigma_\varepsilon^2 \times t$$

which depends on t,and so $\{Y_t\}_{t=1}^{\infty}$ is not stationary.

Figure 1.4 shows a realization of the RW process with $Y_0 = 0$ and $\sigma_\varepsilon^2 = 1$ created using the R commands:

```
> set.seed(321)
> e = rnorm(250)
> y.rw = cumsum(e)
> ts.plot(y.rw, lwd=2, col="blue", main="Random Walk")
> abline(h=0)
```

Although $\{Y_t\}_{t=1}^{\infty}$ is non-stationary, a simple first-differencing transformation, however, yields a covariance stationary process:

$$X_t = Y_t - Y_{t-1} = \varepsilon_t \sim \text{IWN}(0, \sigma_\varepsilon^2).$$

■

Example 12 *Random Walk Model for log Stock Prices*

Let r_t denote the continuously compounded monthly return on Microsoft stock and assume that $r_t \sim \text{GWN}(0.01, (0.05)^2)$. Since $r_t = \ln(P_t/P_{t-1})$ it follows that $\ln P_t = \ln P_{t-1} + r_t$ and so $\ln P_t$ follows a random walk process. Prices, however, do not follow a random walk since $P_t = e^{\ln P_t} = e^{\ln P_{t-1} + r_t}$.

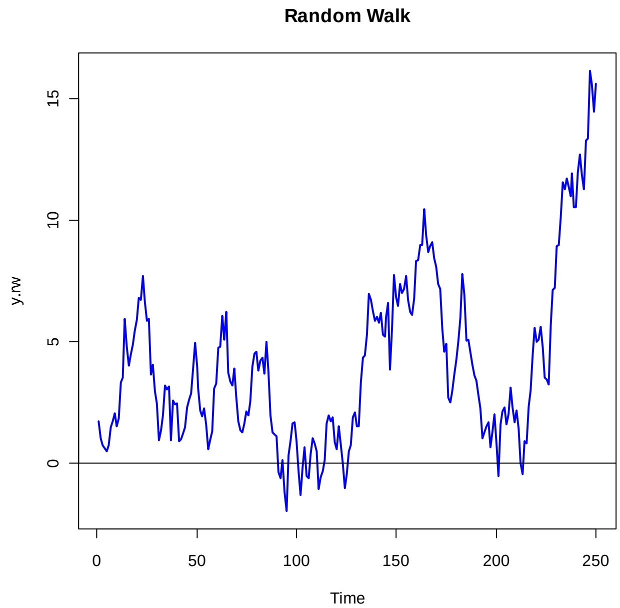

Figure 1.4: Random walk process: $Y_t = Y_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, 1)$.

1.1.3 Ergodicity

In a strictly stationary or covariance stationary stochastic process no assumption is made about the strength of dependence between random variables in the sequence. For example, in a covariance stationary stochastic process it is possible that $\rho_1 = \mathrm{cor}(Y_t, Y_{t-1}) = \rho_{100} = \mathrm{cor}(Y_t, Y_{t-100}) = 0.5$, say. However, in many contexts it is reasonable to assume that the strength of dependence between random variables in a stochastic process diminishes the farther apart they become. That is, $\rho_1 > \rho_2 \cdots$ and that eventually $\rho_j = 0$ for j large enough. This diminishing dependence assumption is captured by the concept of *ergodicity*.

Definition 13 *Ergodicity (intuitive definition)*

Intuitively, a stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is *ergodic* if any two collections of random variables partitioned far apart in the sequence are essentially independent. ■

The formal definition of ergodicity is highly technical and requires advanced concepts in probability theory. However, the intuitive definition captures the essence of the concept. The stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is ergodic if Y_t and Y_{t-j} are essentially independent if j is large enough.

If a stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is covariance stationary and ergodic then strong restrictions are placed on the joint behavior of the elements in the sequence and on the type of temporal dependence allowed.

Example 14 *White noise processes*

If $\{Y_t\}_{t=-\infty}^{\infty}$ is GWN or IWN then it is both covariance stationary and ergodic.

Example 15 *Covariance stationary but not ergodic process (White 1984, pp. xxx)*

Let $Y_t \sim \text{GWN}(0,1)$ and let $X \sim N(0,1)$ independent of $\{Y_t\}_{t=-\infty}^{\infty}$. Define $Z_t = Y_t + X$. Then $\{Z_t\}_{t=-\infty}^{\infty}$ is covariance stationary but not ergodic. To see why $\{Z_t\}_{t=-\infty}^{\infty}$ is not ergodic, note that for all $j > 0$

$$\begin{aligned}
\text{var}(Z_t) &= \text{var}(Y_t + X) = 1 + 1 = 2, \\
\gamma_j &= \text{cov}(Y_t + X, Y_{t-j} + X) = \text{cov}(Y_t, Y_{t-j}) + \text{cov}(Y_t, X) + \text{cov}(Y_{t-j}, X) + \text{cov}(X, X) \\
&= \text{cov}(X, X) = \text{var}(X) = 1, \\
\rho_j &= \frac{1}{2} \text{ for all } j.
\end{aligned}$$

Hence, the correlation between random variables separated far apart does not eventually go to zero and so $\{Z_t\}_{t=-\infty}^{\infty}$ cannot be ergodic. ■

The different flavors of white noise processes are not very interesting because they do not allow any linear dependence between the observations in the series. The following sections describe some simple covariance stationary and ergodic time series models that allow for different patterns of time dependence captured by autocorrelations.

1.2 Moving Average Processes

Moving average models are simple covariance stationary and ergodic time series models that can capture a wide variety of autocorrelation patterns.

1.2.1 MA(1) Model

Suppose you want to create a covariance stationary and ergodic stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ in which Y_t and Y_{t-1} are correlated but Y_t and Y_{t-j} are not correlated for $j > 1$. That is, the time dependence in the process only lasts for one period. Such a process can be created using the *first order moving average* (MA(1)) model:

$$\begin{aligned} Y_t &= \mu + \varepsilon_t + \theta\varepsilon_{t-1}, \quad -1 < \theta < 1, \\ &\varepsilon_t \sim iid\ N(0, \sigma_\varepsilon^2). \end{aligned} \tag{1.2}$$

The moving average parameter θ determines the sign and magnitude of the correlation between Y_t and Y_{t-1}. Clearly, if $\theta = 0$ then $Y_t = \mu + \varepsilon_t$ so that $\{Y_t\}_{t=-\infty}^{\infty}$ exhibits no time dependence.

To verify that (1.2) process is a covariance stationary process we must show that the mean, variance and autocovariances are time invariant. For the mean, we have

$$E[Y_t] = \mu + E[\varepsilon_t] + \theta E[\varepsilon_{t-1}] = \mu,$$

because $E[\varepsilon_t] = E[\varepsilon_{t-1}] = 0$.

For the variance, we have

$$\begin{aligned} \mathrm{var}(Y_t) &= \sigma^2 = E[(Y_t - \mu)^2] = E[(\varepsilon_t + \theta\varepsilon_{t-1})^2] \\ &= E[\varepsilon_t^2] + 2\theta E[\varepsilon_t\varepsilon_{t-1}] + \theta^2 E[\varepsilon_{t-1}^2] \\ &= \sigma_\varepsilon^2 + 0 + \theta^2\sigma_\varepsilon^2 = \sigma_\varepsilon^2(1+\theta^2). \end{aligned}$$

The term $E[\varepsilon_t\varepsilon_{t-1}] = \mathrm{cov}(\varepsilon_t, \varepsilon_{t-1}) = 0$ because $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ is an independent process.

For $\gamma_1 = \mathrm{cov}(Y_t, Y_{t-1})$, we have

$$\begin{aligned} \mathrm{cov}(Y_t, Y_{t-1}) &= E[(Y_t - \mu)(Y_{t-1} - \mu)] \\ &= E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})] \\ &= E[\varepsilon_t\varepsilon_{t-1}] + \theta E[\varepsilon_t\varepsilon_{t-2}] \\ &\quad +\theta E[\varepsilon_{t-1}^2] + \theta^2 E[\varepsilon_{t-1}\varepsilon_{t-2}] \\ &= 0 + 0 + \theta\sigma_\varepsilon^2 + 0 = \theta\sigma_\varepsilon^2. \end{aligned}$$

Note that the sign of γ_1 is the same as the sign of θ. For $\rho_1 = \mathrm{cor}(Y_t, Y_{t-1})$ we have

$$\rho_1 = \frac{\gamma_1}{\sigma^2} = \frac{\theta\sigma_\varepsilon^2}{\sigma_\varepsilon^2(1+\theta^2)} = \frac{\theta}{(1+\theta^2)}.$$

Clearly, $\rho_1 = 0$ if $\theta = 0$; $\rho_1 > 0$ if $\theta > 0$; $\rho_1 < 0$ if $\theta < 0$. Also, the largest value for $|\rho_1|$ is 0.5 which occurs when $|\theta| = 1$.

For $\gamma_2 = \mathrm{cov}(Y_t, Y_{t-2})$, we have

$$\begin{aligned}
\mathrm{cov}(Y_t, Y_{t-2}) &= E[(Y_t - \mu)(Y_{t-2} - \mu)] \\
&= E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-2} + \theta\varepsilon_{t-3})] \\
&= E[\varepsilon_t\varepsilon_{t-2}] + \theta E[\varepsilon_t\varepsilon_{t-3}] \\
&\quad +\theta E[\varepsilon_{t-1}\varepsilon_{t-2}] + \theta^2 E[\varepsilon_{t-1}\varepsilon_{t-3}] \\
&= 0 + 0 + 0 + 0 = 0.
\end{aligned}$$

Similar calculations can be used to show that

$$\mathrm{cov}(Y_t, Y_{t-j}) = \gamma_j = 0 \text{ for } j > 1$$

Hence, for $j > 1$ we have $\rho_j = 0$ and there is only time dependence between Y_t and Y_{t-1} but no time dependence between Y_t and Y_{t-j} for $j > 1$. Because $\rho_j = 0$ for $j > 1$ the MA(1) process is ergodic.

Example 16 *Simulated values from MA(1) process*

The R function `arima.sim()` can be used to simulate observations from a MA(1) process[2]. For example, to simulate 250 observations from (1.2) with $\mu = 1$, $\theta = 0.9$ and $\sigma_\varepsilon = 1$ use

```
> ma1.model = list(ma=0.9)
> mu = 1
> set.seed(123)
> ma1.sim = mu + arima.sim(model=ma1.model,n=250)
```

The `ma` component of the list object `ma1.model` specifies the value of θ for the MA(1) model, and is used as an input to the function `arima.sim()`. By default, `arima.sim()` sets $\mu = 0$ and specifies $\varepsilon_t \sim \mathrm{GWN}(0, 1)$. Other distributions for ε_t can be specified using the optional `innov` argument. For example, to set $\varepsilon_t \sim \mathrm{GWN}(0, (0.1)^2)$ use

```
> ma1.sim = mu + arima.sim(model=ma1.model,n=250,
+                           innov=rnorm(250,mean=0,sd=0.1))
```

[2]The function `arima.sim()` can be used to simulate observations from the class of autoregressive integrated moving average (ARIMA) models, of which the MA(1) model is a special case.

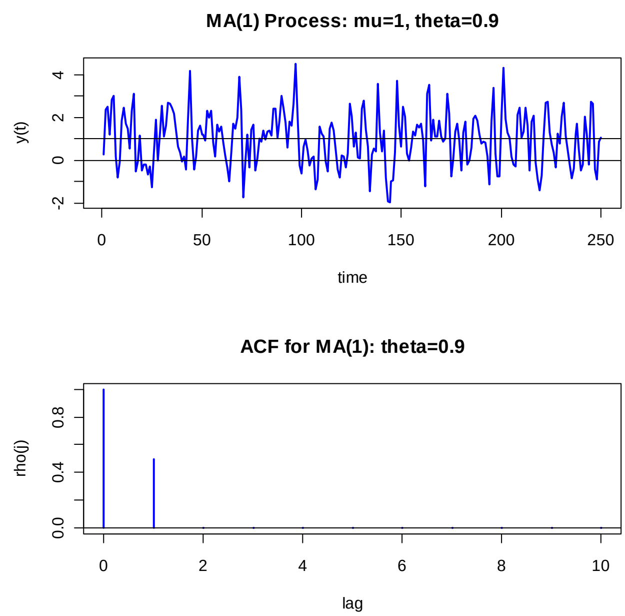

Figure 1.5: Simulated values and theoretical ACF from MA(1) process with $\mu = 1$, $\theta = 0.9$ and $\sigma_\varepsilon^2 = 1$.

The function `ARMAacf()` can be used to compute the theoretical autocorrelations, ρ_j, from the MA(1) model (recall, $\rho_1 = \theta/(1+\theta^2)$ and $\rho_j = 0$ for $j > 1$). For example, to compute ρ_j for $j = 1, \ldots, 10$ use

```
> ma1.acf = ARMAacf(ar=0, ma=0.9, lag.max=10)
> ma1.acf
     0      1      2      3      4      5      6      7      8      9
1.0000 0.4972 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
    10
0.0000
```

Figure 1.5 shows the simulated data and the theoretical ACF created using

```
> par(mfrow=c(2,1))
> ts.plot(ma1.sim,main="MA(1) Process: mu=1, theta=0.9",
+ xlab="time",ylab="y(t)", col="blue", lwd=2)
> abline(h=c(0,1))
```

```
> plot(0:10, ma1.acf,type="h", col="blue", lwd=2,
+ main="ACF for MA(1): theta=0.9",xlab="lag",ylab="rho(j)")
> abline(h=0)
> par(mfrow=c(1,1)
```

Compared to the GWN process in 1.1, the MA(1) process is a bit smoother in its appearance. This is due to the positive one-period time dependence $\rho_1 = 0.4972$.
■

Example 17 *MA(1) model for overlapping continuously compounded returns*

Let r_t denote the one-month continuously compounded return and assume that

$$r_t \sim \text{iid } N(\mu, \sigma^2).$$

Consider creating a monthly time series of two-month continuously compounded returns using

$$r_t(2) = r_t + r_{t-1}.$$

The time series of these two-month returns, observed monthly, overlap by one month:

$$\begin{aligned} r_t(2) &= r_t + r_{t-1}, \\ r_{t-1}(2) &= r_{t-1} + r_{t-2}, \\ r_{t-2}(2) &= r_{t-2} + r_{t-3}, \\ &\vdots \end{aligned}$$

The one-month overlap in the two-month returns implies that $\{r_t(2)\}$ follows an MA(1) process. To show this, we need to show that the autocovariances of $\{r_t(2)\}$ behave like the autocovariances of an MA(1) process.

To verify that $\{r_t(2)\}$ follows an MA(1) process, first we have

$$\begin{aligned} E[r_t(2)] &= E[r_t] + E[r_{t-1}] = 2\mu \\ \text{var}(r_t(2)) &= var(r_t + r_{t-1}) = 2\sigma^2 \end{aligned}$$

Next, we have

$$\text{cov}(r_t(2), r_{t-1}(2)) = \text{cov}(r_t + r_{t-1}, r_{t-1} + r_{t-2}) = \text{cov}(r_{t-1}, r_{t-1}) = \text{var}(r_{t-1}) = \sigma^2$$

and

$$\begin{aligned}
\mathrm{cov}(r_t(2), r_{t-2}(2)) &= \mathrm{cov}(r_t + r_{t-1}, r_{t-2} + r_{t-3}) = 0 \\
\mathrm{cov}(r_t(2), r_{t-j}(2)) &= 0 \text{ for } j > 1
\end{aligned}$$

Hence, the autocovariances of $\{r_t(2)\}$ are those of an MA(1) process. Notice that

$$\rho_1 = \frac{\sigma^2}{2\sigma^2} = \frac{1}{2}.$$

What MA(1) process describes $\{r_t(2)\}$? Because $\rho_1 = \frac{\theta}{1+\theta^2} = 0.5$ it follows that $\theta = 1$. Hence, the MA(1) process has mean 2μ and $\theta = 1$ and can be expressed as

$$\begin{aligned}
r_t(2) &= 2\mu + \varepsilon_t + \varepsilon_{t-1}, \\
&\quad \varepsilon_t \sim GWN(0, \sigma^2).
\end{aligned}$$

1.3 Autoregressive Processes

1.3.1 AR(1) Model

Suppose you want to create a covariance stationary and ergodic stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ in which Y_t and Y_{t-1} are correlated, Y_t and Y_{t-2} are slightly less correlated, Y_t and Y_{t-3} are even less correlated and eventually Y_t and Y_{t-j} are uncorrelated for j large enough. That is, the time dependence in the process decays to zero as the random variables in the process get farther and farther apart. Such a process can be created using the *first order autoregressive* (AR(1)) model:

$$\begin{aligned}
Y_t - \mu &= \phi(Y_{t-1} - \mu) + \varepsilon_t, \quad -1 < \phi < 1 \qquad (1.3) \\
&\quad \varepsilon_t \sim \text{iid } N(0, \sigma_\varepsilon^2)
\end{aligned}$$

It can be shown that the AR(1) model is covariance stationary and ergodic provided $-1 < \phi < 1$. We will show that the AR(1) process has the following properties:

$$\begin{aligned}
E[Y_t] &= \mu, & (1.4)\\
\mathrm{var}(Y_t) &= \sigma^2 = \sigma_\varepsilon^2/(1-\phi^2), & (1.5)\\
\mathrm{cov}(Y_t, Y_{t-1}) &= \gamma_1 = \sigma^2\phi, & (1.6)\\
\mathrm{cor}(Y_t, Y_{t-1}) &= \rho_1 = \gamma_1/\sigma^2 = \phi, & (1.7)\\
\mathrm{cov}(Y_t, Y_{t-j}) &= \gamma_j = \sigma^2\phi^j, & (1.8)\\
\mathrm{cor}(Y_t, Y_{t-j}) &= \rho_j = \gamma_j/\sigma^2 = \phi^j. & (1.9)
\end{aligned}$$

Notice that the restriction $|\phi| < 1$ implies that

$$\lim_{j\to\infty} \rho_j = \phi^j = 0,$$

so that Y_t is essentially independent of Y_{t-j} for large j and so $\{Y_t\}_{t=-\infty}^{\infty}$ is ergodic. For example, if $\phi = 0.5$ then $\rho_{10} = (0.5)^{10} = 0.001$; if $\phi = 0.9$ then $\rho_{10} = (0.9)^{10} = 0.349$. Hence, the closer ϕ is to unity the stronger is the time dependence in the process.

Verifying covariance stationarity for the AR(1) model is more involved than for the MA(1) model, and establishing the properties (1.4) - (1.9) involves some tricks. First, consider the derivation for (1.4). We have

$$\begin{aligned}
E[Y_t] &= \mu + \phi(E[Y_{t-1}] - \mu) + E[\varepsilon_t]\\
&= \mu + \phi E[Y_{t-1}] - \phi\mu.
\end{aligned}$$

If we assume that $\{Y_t\}_{t=-\infty}^{\infty}$ is covariance stationary then $E[Y_t] = [Y_{t-1}]$. Substituting into the above and solving for $E[Y_t]$ gives (1.4). A similar trick can be used to derive (1.5):

$$\mathrm{var}(Y_t) = \phi^2(\mathrm{var}(Y_{t-1})) + \mathrm{var}(\varepsilon_t) = \phi^2(\mathrm{var}(Y_t)) + \sigma_\varepsilon^2,$$

which uses the fact that Y_{t-1} is independent of ε_t, and $\mathrm{var}(Y_t) = \mathrm{var}(Y_{t-1})$ provided $\{Y_t\}$ is covariance stationary. Solving for $\sigma^2 = \mathrm{var}(Y_t)$ gives (1.5). To determine (1.6), multiply both sides of (1.3) by $Y_{t-1} - \mu$ and take expectations to give

$$\gamma_1 = E\left[(Y_t - \mu)(Y_{t-1} - \mu)\right] = \phi E\left[(Y_{t-1} - \mu)^2\right] + E\left[\varepsilon_t (Y_{t-1} - \mu)\right] = \phi\sigma^2,$$

which uses the fact that Y_{t-1} is independent of ε_t, and $\mathrm{var}(Y_t) = \mathrm{var}(Y_{t-1}) = \sigma^2$. Finally, to determine (1.8), multiply both sides of (1.3) by $Y_{t-j} - \mu$ and take expectations to give

$$\begin{aligned}\gamma_j &= E\left[(Y_t-\mu)\left(Y_{t-j}-\mu\right)\right] = \phi E\left[(Y_{t-1}-\mu)(Y_{t-j}-\mu)\right] + E\left[\varepsilon_t\left(Y_{t-j}-\mu\right)\right] \\ &= \phi\gamma_{j-1},\end{aligned}$$

which uses the fact that Y_{t-j} is independent of ε_t, and $E\left[(Y_{t-1}-\mu)(Y_{t-j}-\mu)\right] = \gamma_{j-1}$ provided $\{Y_t\}$ is covariance stationary. Using recursive substitution and $\gamma_0 = \sigma^2$ gives (1.8).

The AR(1) Model can be re-expressed in the form of a linear regression model as follows:

$$\begin{aligned}Y_t - \mu &= \phi(Y_{t-1}-\mu) + \varepsilon_t \Rightarrow \\ Y_t &= \mu - \phi\mu + \phi Y_{t-1} + \varepsilon_t \\ &= c + \phi Y_{t-1} + \varepsilon_t,\end{aligned}$$

where $c = (1-\phi)\mu \Rightarrow \mu = c/(1-\phi)$. This regression model form is convenient for estimation by ordinary least squares.

Example 18 *Simulated values from AR(1) process*

The R function `arima.sim()` can be used to simulate observations from an AR(1) process, and the function `ARMAacf()` can be used to compute the theoretical ACF. For example, to simulate 250 observations from (1.3) with $\mu = 1$, $\phi = 0.9$ and $\sigma_\varepsilon = 1$, and compute the theoretical ACF use

```
> ar1.model = list(ar=0.9)
> mu = 1
> set.seed(123)
> ar1.sim = mu + arima.sim(model=ar1.model,n=250)
> ar1.acf = ARMAacf(ar=0.9, ma=0, lag.max=10)
```

These values are shown in Figure 1.6. Compared to the MA(1) process in 1.5, the realizations from the AR(1) process are much smoother. That is, when Y_t wanders high above its mean it tends to stay above the mean for a while and when it wanders low below the mean it tends to stay below for a while.

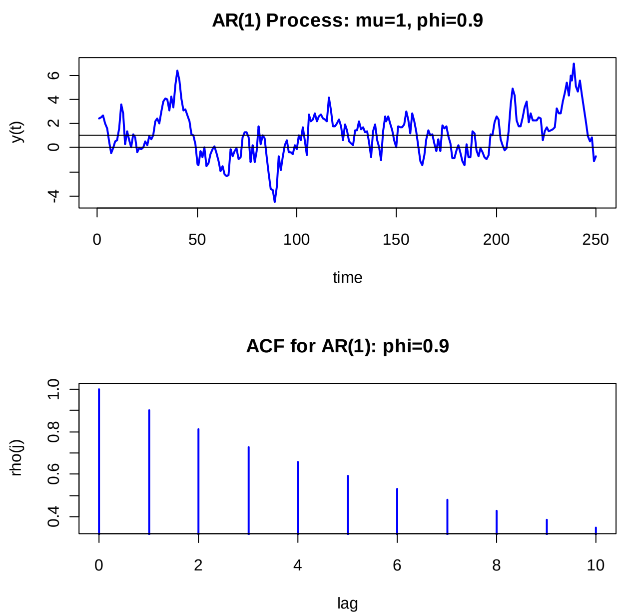

Figure 1.6: Simulated values and ACF from AR(1) model with $\mu = 1, \phi = 0.9$ and $\sigma_\varepsilon^2 = 1$.