Time Series Concepts

Herbert (Yuming) Liu and Eric Zivot

3/28/2021

Stochastic (Random) Process

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

sequence of random variables indexed by time

Observed time series of length T

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

Strictly Stationary Processes

- Intuition: $\{Y_t\}$ is stationary if all aspects of its behavior are unchanged by shifts in time
- *Strict stationarity*: A stochastic process $\{Y_t\}_{t=1}^{\infty}$ 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.

- Strict stationarity implies that the joint distribution of $(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_r})$ is the same as the joint distribution of $(Y_{t_1-t}, Y_{t_2-t}, \ldots, Y_{t_r-t})$ for any value of t.

Remarks

- For example, the distribution of (Y_1, Y_5) is the same as the distribution of (Y_{12}, Y_{16}).
- For a strictly stationary process, Y_t has the same mean, variance (moments) for all t.
- Any function/transformation $g(\cdot)$ of a strictly stationary process, $\{g(Y_t)\}$ is also strictly stationary.
  - if $\{Y_t\}$ is strictly stationary then $\{Y_t^2\}$ is strictly stationary.

Covariance (Weakly) Stationary Processes

$\{Y_t\}$ is a *covariance stationary* process if

1. $E[Y_t] = \mu$ for all t
2. $\mathrm{var}(Y_t) = \sigma^2$ for all t
3. $\mathrm{cov}(Y_t, Y_{t-j}) = \gamma_j$ depends on j and not on t

- $\mathrm{cov}(Y_t, Y_{t-j}) = \gamma_j$ is called the j-lag *autocovariance* and measures the direction of linear time dependence between Y_t and Y_{t-j}
- A strictly stationary process is covariance stationary if $\mathrm{var}(Y_t) < \infty$ and $\mathrm{cov}(Y_t, Y_{t-j}) < \infty$

Autocorrelations

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

- $\mathrm{corr}(Y_t, Y_{t-j}) = \rho_j$ is called the j-lag *autocorrelation* and measures the direction and strength of linear time dependence between Y_t and Y_{t-j}
- By stationarity $\mathrm{var}(Y_t) = \mathrm{var}(Y_{t-j}) = \sigma^2$.
- *Autocorrelation Function* (ACF): Plot of ρ_j against j

ACF Plot

Ergodicity

- In a strictly stationary or covariance stationary stochastic process no assumption is made about the strenth of dependence between random variables in the sequence.
- 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.
- This diminishing dependence assumption is captured by the concept of *ergodicity*.
- Intuitively, a stochastic process $\{Y_t\}$ is *ergodic* if any two collections of random variables partitioned far apart in the sequence are essentially independent.

Ergodicity and Autocorrelations

- If a stochastic process $\{Y_t\}$ is *covariance stationary* and *ergodic* then

$$\rho_j = \mathrm{cor}(Y_t, Y_{t-j}) \to 0 \text{ as } j \text{ gets large}$$

- For example, suppose $\rho_j = \rho^j$. Then, if $|\rho| < 1$

$$\rho_j = \rho^j \to 0 \text{ as } j \text{ gets large}$$

Example: Gaussian White Noise (GWN) Process

$$
\begin{gathered}
Y_t \sim \text{iid } N(0, \sigma^2) \text{ or } Y_t \sim GWN(0, \sigma^2) \\
E[Y_t] = 0, \ \text{var}(Y_t) = \sigma^2 \\
Y_t \text{ independent of } Y_s \text{ for } t \neq s \\
\Rightarrow \text{cov}(Y_t, Y_{t-s}) = 0 \text{ for } t \neq s
\end{gathered}
$$

- "iid" = **i**ndependent and **i**dentically **d**istributed.
- $\{Y_t\}$ represents random draws from the same $N(0, \sigma^2)$ distribution.
- Clearly, $\{Y_t\}$ is covariance stationary and ergodic.

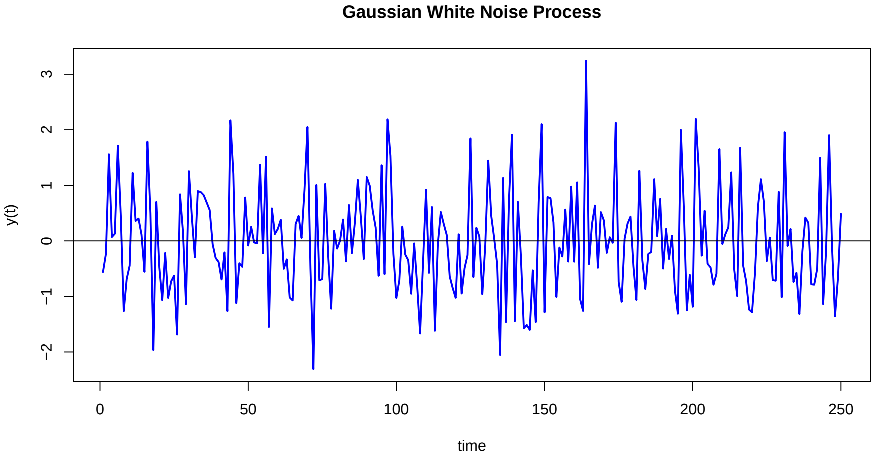
Example: GWN Process

Use `rnorm()` to simulate 250 observations from GWN(0,1)

```
# simulate Gaussian White Noise process
set.seed(123)
y = rnorm(250)
```

- `set.seed(123)` initializes the random number generator in R
- Setting the seed allows the random numbers to be reproduced by anyone using the same seed

Example: GWN Process



Example: Independent White Noise (IWN) Process

$$
\begin{gathered}
Y_t \sim \text{iid }(0, \sigma^2) \text{ or } Y_t \sim IWN(0, \sigma^2) \\
E[Y_t] = 0, \ \text{var}(Y_t) = \sigma^2 \\
Y_t \text{ independent of } Y_s \text{ for } t \neq s
\end{gathered}
$$

- Here, $\{Y_t\}$ represents random draws from the same distribution. However, we don't specify exactly what the distribution is - only that it has mean zero and variance σ^2.
- For example, Y_t could be iid Student's t with variance equal to σ^2. This is like GWN but with fatter tails (i.e., more extreme observations).

Example: IWN with Standardized Student-t errors

- Recall, $Y \sim t_v$ has $E[Y] = 0$ and $\mathrm{var}(Y) = v/(v-2)$
- $X = \sqrt{(v-2)/v} \times Y$ has $E[X] = 0$ and $\mathrm{var} = 1$. X is called a *standardized Student's t* rv

```r
## simulate IWN using scaled student's t with 3 df
## distribution to have variance 1
set.seed(123)
y = (1/sqrt(3))*rt(250, df=3)
```

Example: IWN with Standardized Student-t errors

Example: Weak White Noise (WN) Process

$$
\begin{gathered}
Y_t \sim WN(0, \sigma^2) \\
E[Y_t] = 0,\ \mathrm{var}(Y_t) = \sigma^2 \\
\mathrm{cov}(Y_t, Y_s) = 0 \text{ for } t \neq s
\end{gathered}
$$

- Here, $\{Y_t\}$ represents an uncorrelated stochastic process with mean zero and variance σ^2.
- Recall, the uncorrelated assumption does not imply independence unless $\{Y_t\}$ is normally distributed .
- Hence, Y_t and Y_s can exhibit non-linear dependence (e.g. Y_t^2 can be correlated with Y_s^2)

Nonstationary Processes

- A *nonstationary* stochastic process is a stochastic process that is not covariance stationary.
- A non-stationary process violates one or more of the properties of covariance stationarity. For example, (1) the mean could depend on t; (2) the variance could depend on t; (3) the covariances between Y_t and Y_{t-j} could depend on t.

Example: Deterministically Trending Process

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$
$$E[Y_t] = \beta_0 + \beta_1 t \text{ depends on } t$$

Note: A simple detrending transformation yields a stationary process:

$$X_t = Y_t - \beta_0 - \beta_1 t = \varepsilon_t$$

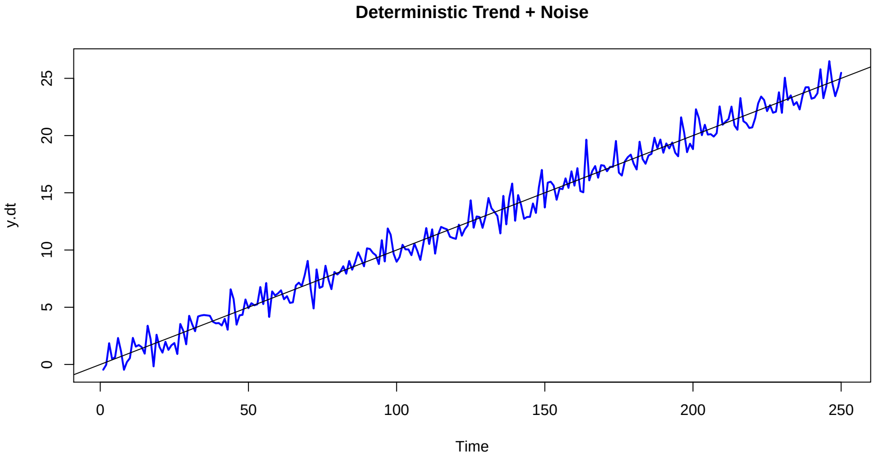
Example: Deterministic Trend + GWN Errors

Simulate $T = 250$ observations from

$$Y_t = 0.1 \times t + \varepsilon_t,\ \varepsilon_t \sim \mathrm{GWN}(0,1), t = 1, \cdots, T$$

```
set.seed(123)
e = rnorm(250)
y.dt = 0.1*seq(1,250) + e
```

Example: Deterministic Trend + GWN Errors



Example: Random Walk

$$
\begin{aligned}
Y_t &= Y_{t-1} + \varepsilon_t,\ \ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2),\ \ Y_0 \text{ is fixed}, t = 1, \cdots, T \\
&= Y_0 + \sum_{j=1}^{t} \varepsilon_j \Rightarrow \mathrm{var}(Y_t) = \sigma_\varepsilon^2 \times t \ \ \text{depends on } t
\end{aligned}
$$

Note: A simple detrending transformation yields a stationary process:

$$\Delta Y_t = Y_t - Y_{t-1} = \varepsilon_t$$

Example: Random Walk - More Detail

Example: Randow Walk

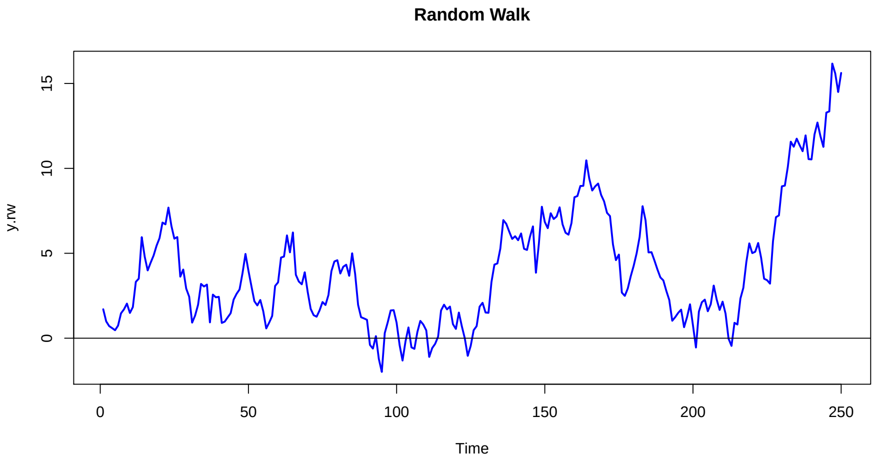
Simulate $T = 250$ observations from RW

$$Y_t = Y_{t-1} + \varepsilon_t,\ \ \varepsilon_t \sim \mathrm{GWN}(0,1), Y_0 = 0, t = 1, \cdots, T$$

```
set.seed(321)
e = rnorm(250)
y.rw = cumsum(e)
```

- `cusum()` computes the cumulative sum of the elements of a vector

Example: Random Walk



Example: Random Walk with Drift

$$
\begin{aligned}
Y_t &= \mu + Y_{t-1} + \varepsilon_t,\ \ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2),\ \ Y_0 \text{ is fixed} \\
&= Y_0 + \mu \times t + \sum_{j=1}^{t} \varepsilon_j \\
&\Rightarrow E[Y_t] = Y_0 + \mu \times t \text{ and } \mathrm{var}(Y_t) = \sigma_\varepsilon^2 \times t \ \text{ depend on } t
\end{aligned}
$$

Note: A simple detrending transformation yields a stationary process:

$$
\Delta Y_t = Y_t - Y_{t-1} = \mu + \varepsilon_t
$$

Example: Random Walk with Drift - More Detail

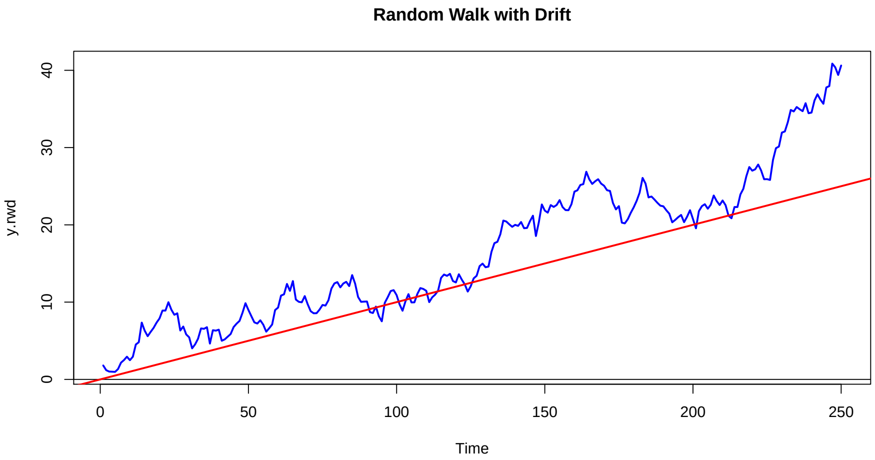
Example: Randow Walk with Drift

Simulate $T = 250$ observations from RW with drift

$$Y_t = 0.1 + Y_{t-1} + \varepsilon_t,\ \ \varepsilon_t \sim \mathrm{GWN}(0,1), Y_0 = 0, t = 1, \cdots, T$$

```
set.seed(321)
e = rnorm(250)
y.rwd = 0.1*seq(1:250) + cumsum(e)
```

Example: Random Walk with Drift



Time Series Models

- A *time series model* is a probability model to describe the behavior of a stochastic process $\{Y_t\}$.
- Typically, a time series model is a simple probability model that describes the time dependence in the stochastic process $\{Y_t\}$.
- Two common time series models are *autoregressive* models and *moving average* models.

Moving Average (MA) Process

Consider a stochastic process that only exhibits one period linear time dependence.

MA(1) Model:

$$
\begin{aligned}
Y_t &= \mu + \varepsilon_t + \theta\varepsilon_{t-1}, \quad -\infty < \theta < \infty \\
&\varepsilon_t \sim iid\ N(0, \sigma_\varepsilon^2) \text{ (i.e., } \varepsilon_t \sim GWN(0, \sigma_\varepsilon^2)) \\
&\theta \text{ determines the magnitude of time dependence}
\end{aligned}
$$

Properties:

$$
\begin{aligned}
E[Y_t] &= \mu + E[\varepsilon_t] + \theta E[\varepsilon_{t-1}] \\
&= \mu + 0 + 0 = \mu
\end{aligned}
$$

Note: MA(1) is covariance stationary for any value of θ.

Moving Average (MA) Process

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

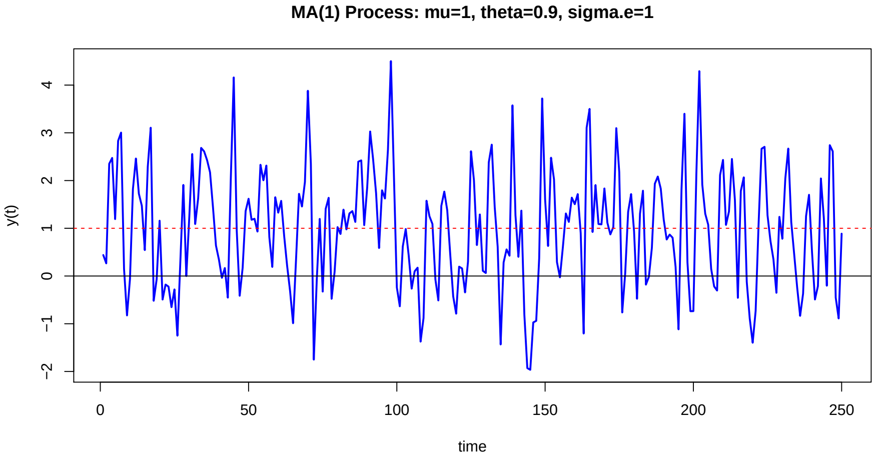
$$\begin{aligned}
\mathrm{cov}(Y_t, Y_{t-1}) &= \gamma_1 = E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})] \\
&= E[\varepsilon_t\varepsilon_{t-1}] + \theta E[\varepsilon_t\varepsilon_{t-2}] \\
&\quad + \theta E[\varepsilon_{t-1}^2] + \theta^2 E[\varepsilon_{t-1}\varepsilon_{t-2}] \\
&= 0 + 0 + \theta\sigma_\varepsilon^2 + 0 = \theta\sigma_\varepsilon^2
\end{aligned}$$

Moving Average (MA) Process

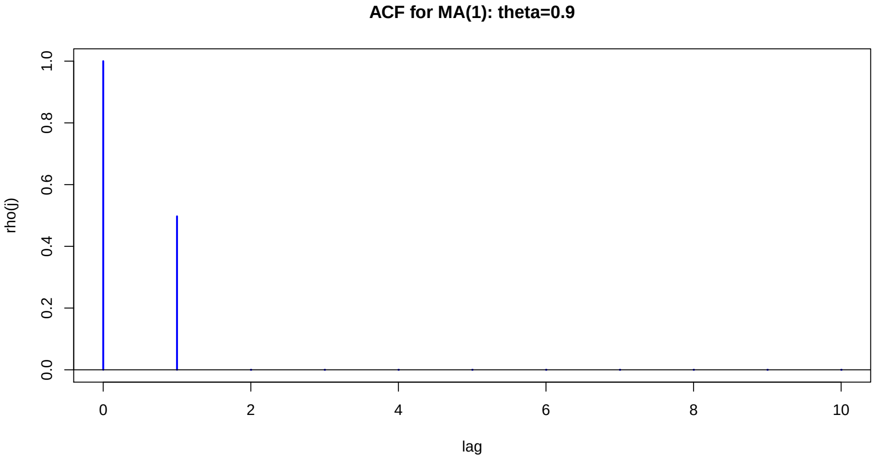
$$Y_t = 1 + \varepsilon_t + 0.9 \times \varepsilon_{t-1}$$

```
# set parameters
mu = 1
theta = 0.9
sigma.e = 1
n.obs = 250
# simulated MA(1) using vectorized calculations
set.seed(123)
e = rnorm(n.obs, sd = sigma.e)
em1 = c(0, e[1:(n.obs-1)])
y = mu + e + theta*em1
```

Moving Average (MA) Process



ACF for MA(1) Process: Theta > 0



Autoregressive (AR) Process

Consider a stochastic process that exhibits multi-period geometrically decaying linear time dependence.

AR(1) Model:

$$Y_t = (1-\phi)\mu + \phi Y_{t-1} + \varepsilon_t, \ \ -1 < \phi < 1$$
$$\varepsilon_t \sim \text{iid } N(0, \sigma_\varepsilon^2)$$

Remark: AR(1) model is covariance stationary provided $-1 < \phi < 1$.

Autoregressive (AR) Process

Properties:

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

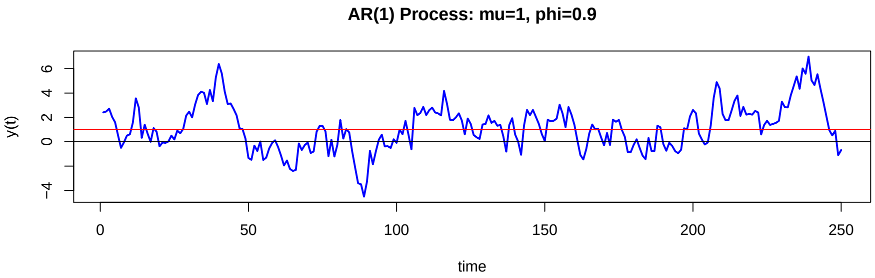
Note: Since $|\phi| < 1$,

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

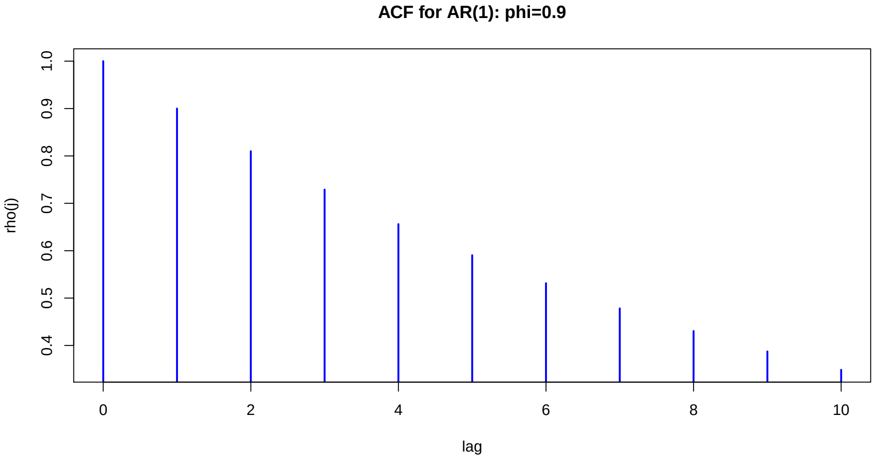
AR(1) Process: Phi > 0

$$Y_t = 1 + 0.9 \times (Y_{t-1} - 1) + \epsilon_t$$

```
ar1.model = list(ar=0.9)
mu = 1
set.seed(123)
ar1.sim = mu + arima.sim(model=ar1.model, n=250)
```



ACF for AR(1) Process: Phi $>$ 0



The AR(1) model and Economic and Financial Time Series}

The AR(1) model is a good description for the following time series:

- Interest rates on U.S. Treasury securities, dividend yields, unemployment
- Growth rate of macroeconomic variables
  - Real GDP, industrial production, productivity
  - Money, velocity, consumer prices
  - Real and nominal wages