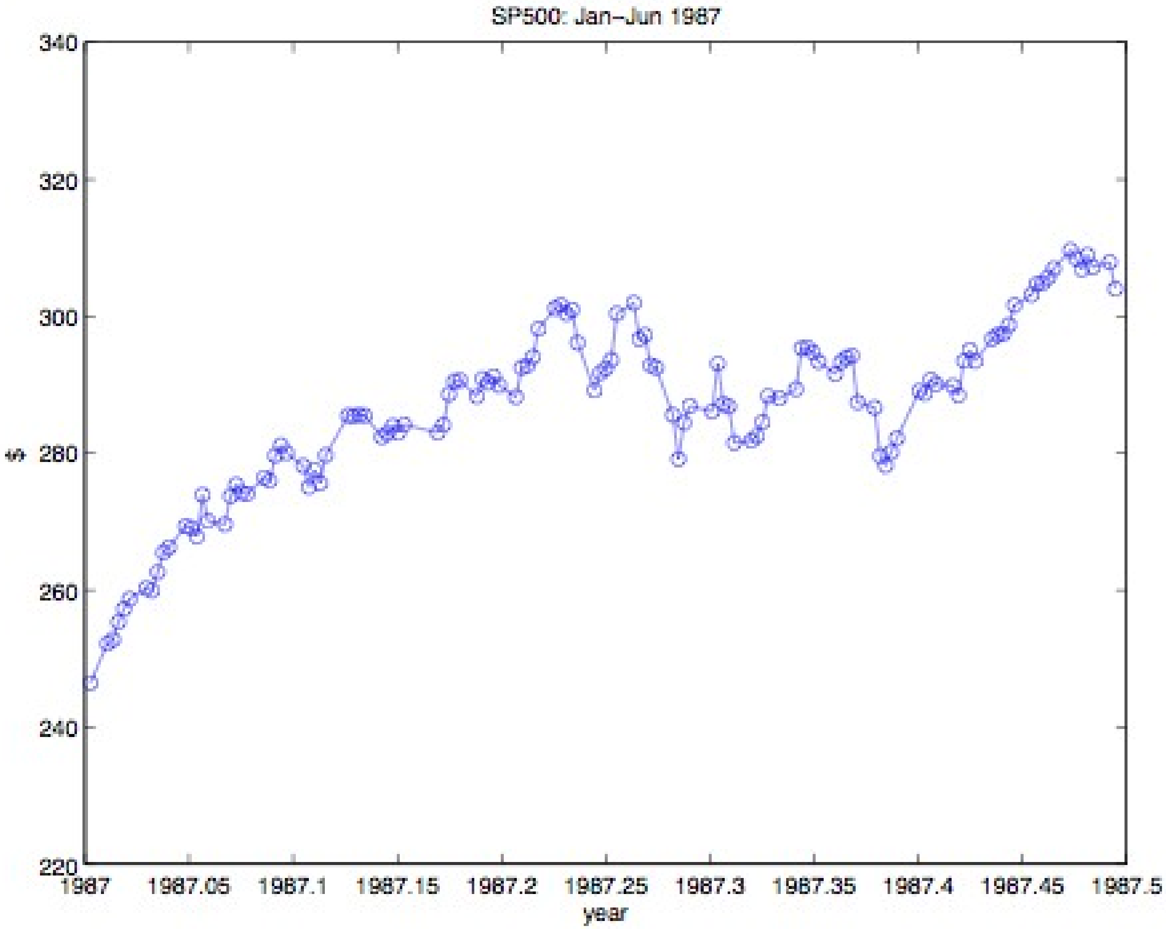
Time series analysis

Example


SP500: Jan–Jun 1987
$
year
340
320
300
280
260
240
220
1987
1987.05
1987.1
1987.15
1987.2
1987.25
1987.3
1987.35
1987.4
1987.45
1987.5

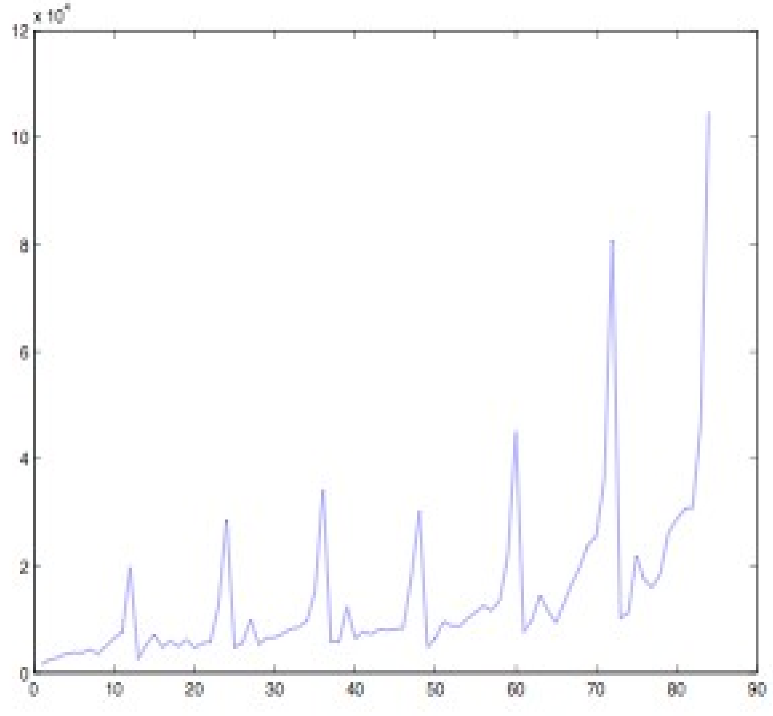
Objectives of time series analysis

1. Compact description of data.
2. Interpretation.
3. Forecasting.
4. Control.
5. Hypothesis testing.
6. Simulation.

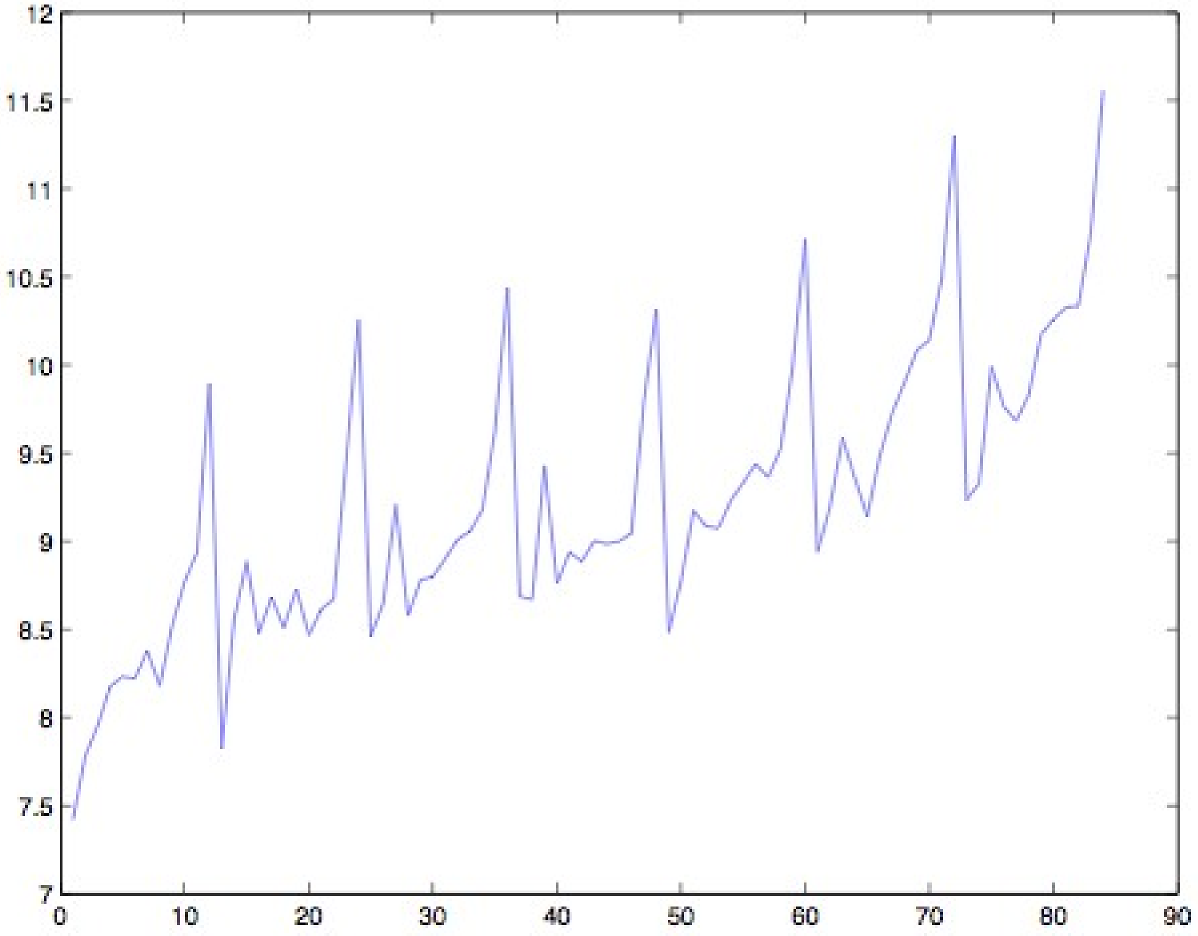
Classical decomposition: An example

Monthly sales for a souvenir shop at a beach resort town in Queensland.

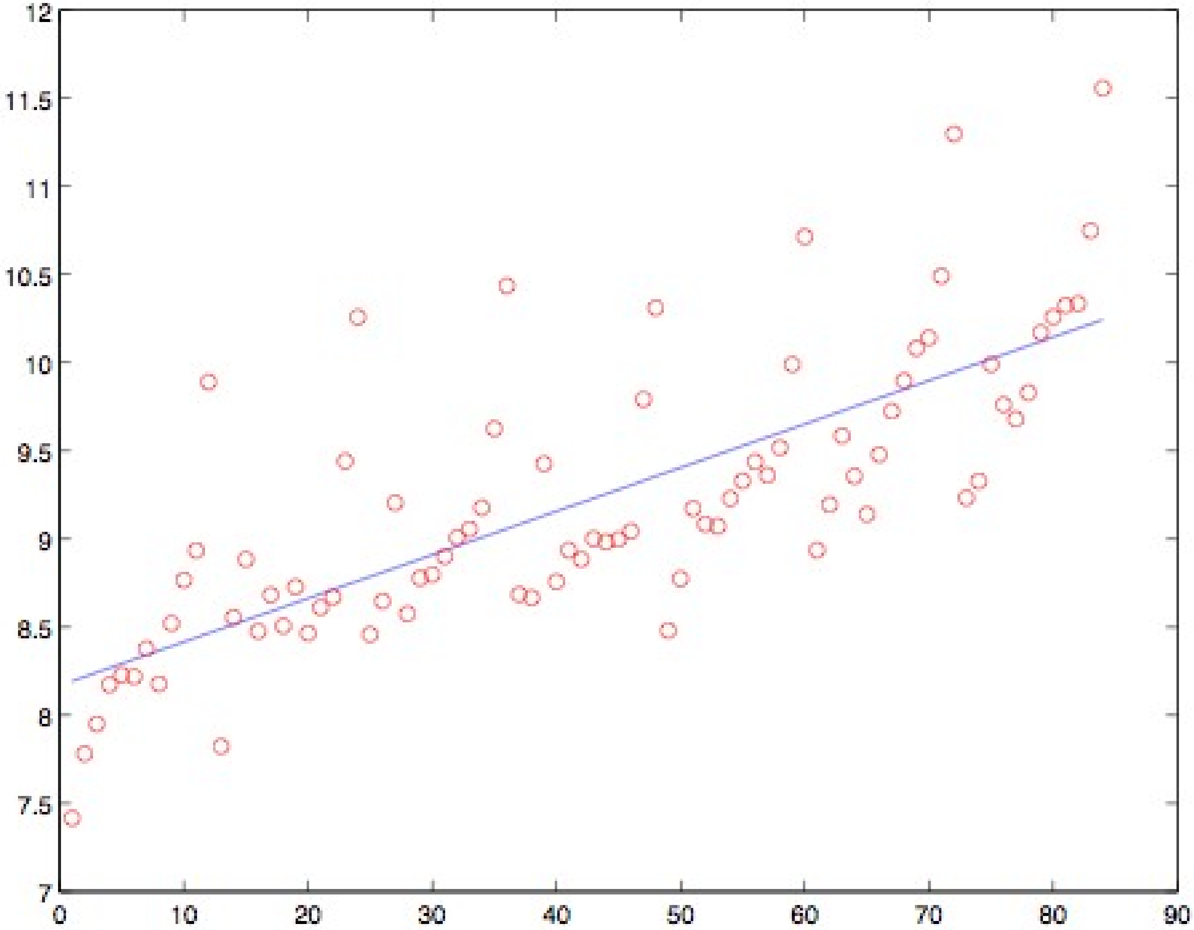
(Makridakis, Wheelwright and Hyndman, 1998)



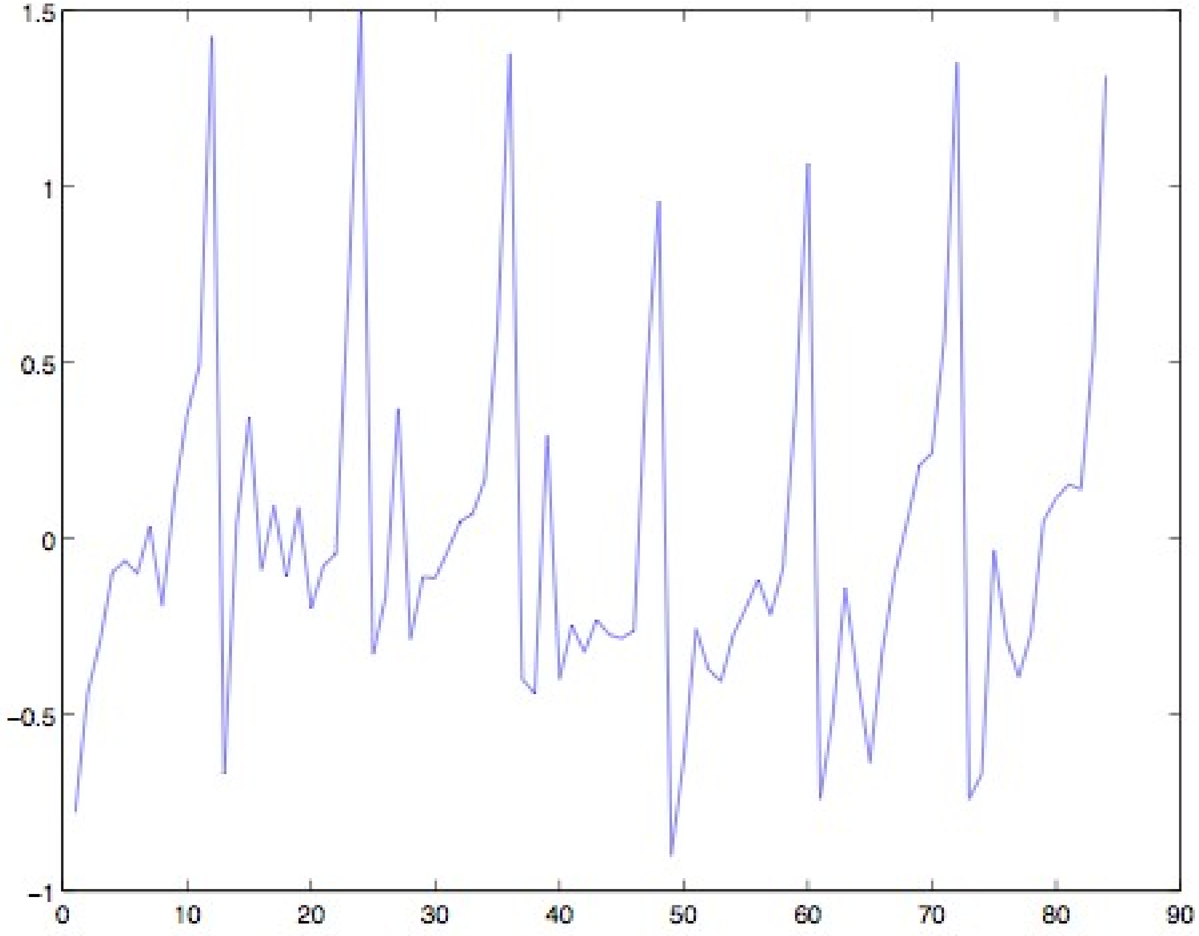
Transformed data



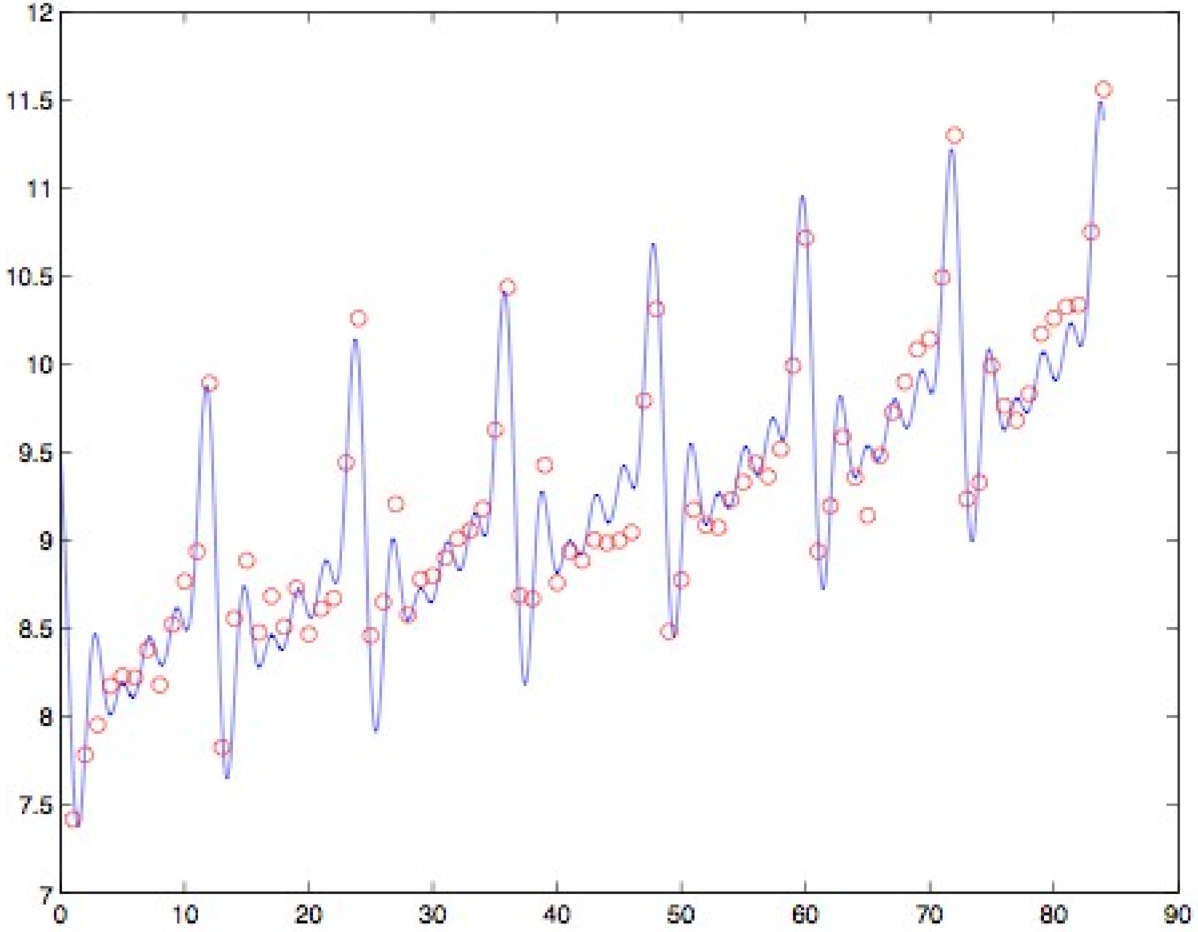
Trend



Residuals



Trend and seasonal variation

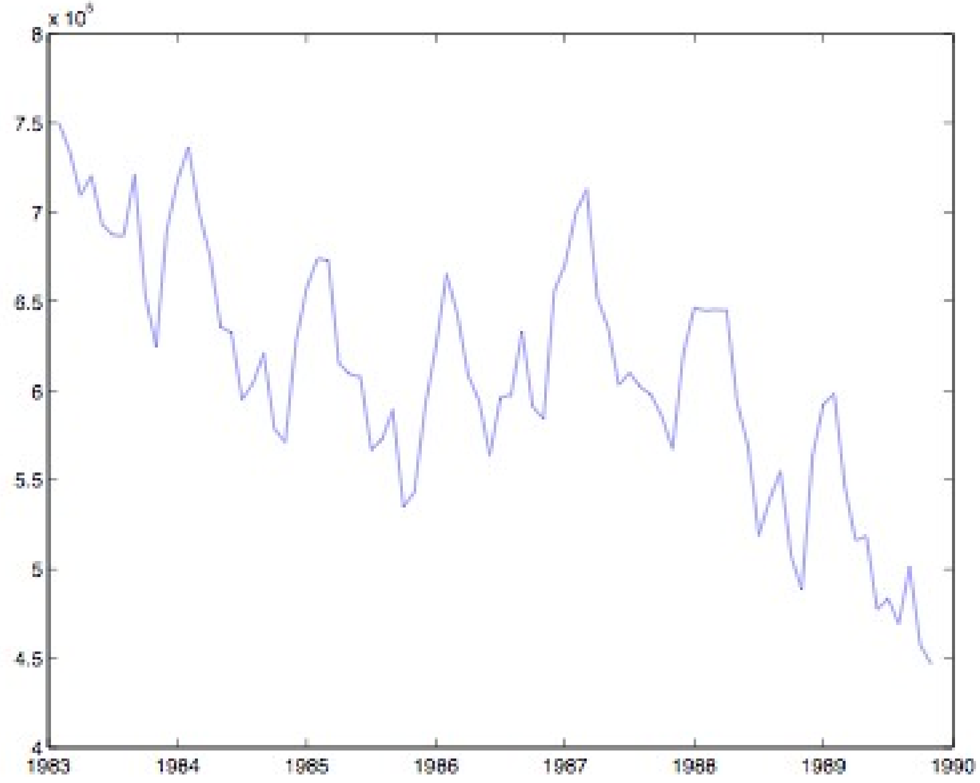


Objectives of time series analysis

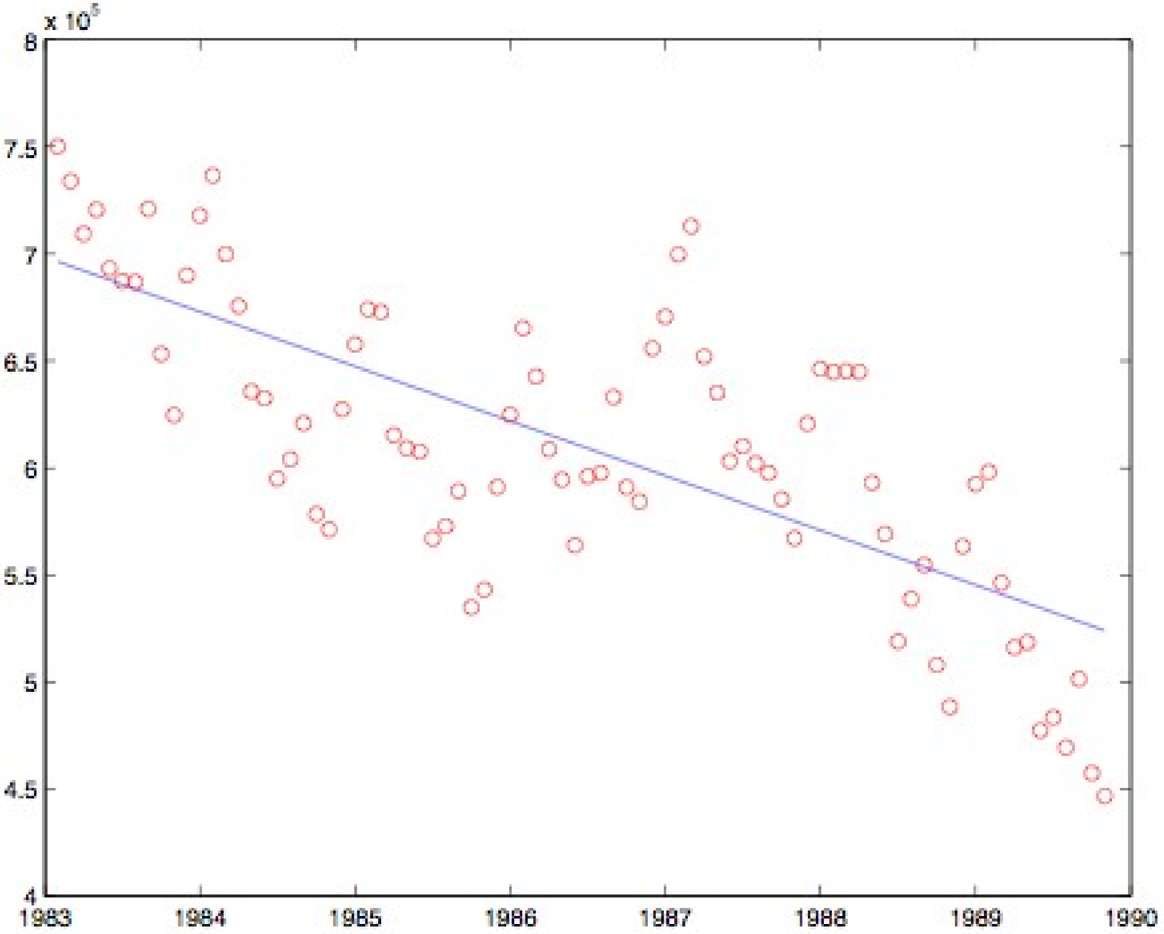
1. Compact description of data.
   Example: Classical decomposition: $X_t = T_t + S_t + Y_t$.
2. Interpretation. Example: Seasonal adjustment.
3. Forecasting. Example: Predict sales.
4. Control.
5. Hypothesis testing.
6. Simulation.

Unemployment data

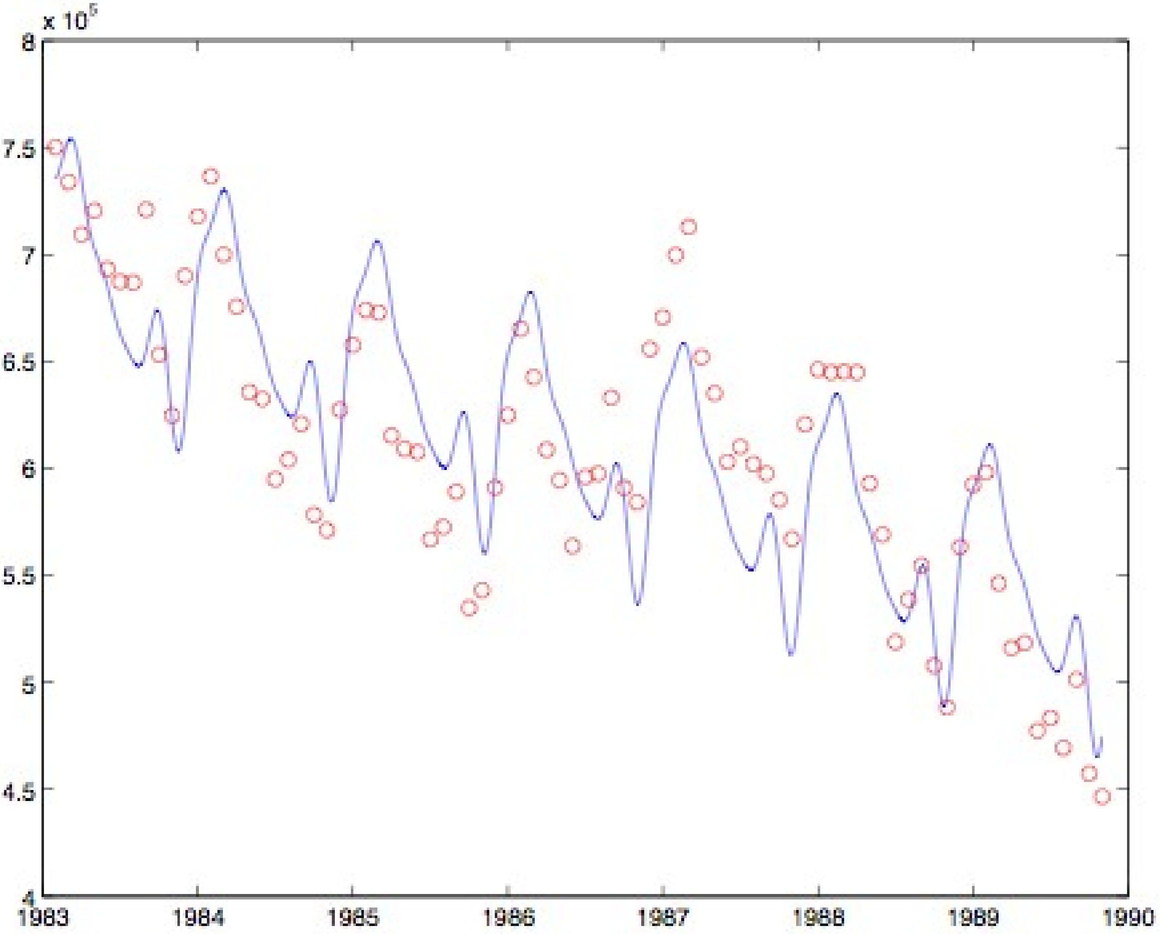
Monthly number of unemployed people in Australia. (Hipel and McLeod, 1994)



Trend



Trend plus seasonal variation



Objectives of time series analysis

1. Compact description of data:

$$X_t = T_t + S_t + f(Y_t) + W_t.$$

2. Interpretation. Example: Seasonal adjustment.
3. Forecasting. Example: Predict unemployment.
4. Control. Example: Impact of monetary policy on unemployment.
5. Hypothesis testing. Example: Global warming.
6. Simulation. Example: Estimate probability of catastrophic events.

Time series models

A **time series model** specifies the joint distribution of the sequence $\{X_t\}$ of random variables.
For example:

$$P[X_1 \le x_1, \ldots, X_t \le x_t] \text{ for all } t \text{ and } x_1, \ldots, x_t.$$

Notation:
$X_1, X_2, \ldots$ is a stochastic process.
$x_1, x_2, \ldots$ is a single realization.

We'll mostly restrict our attention to **second-order properties** only:
$\mathrm{E}X_t, \mathrm{E}(X_{t_1}, X_{t_2})$.

Time series models

Example: White noise: $X_t \sim WN(0, \sigma^2)$.
i.e., $\{X_t\}$ uncorrelated, $\mathrm{E}X_t = 0$, $\mathrm{Var}X_t = \sigma^2$.

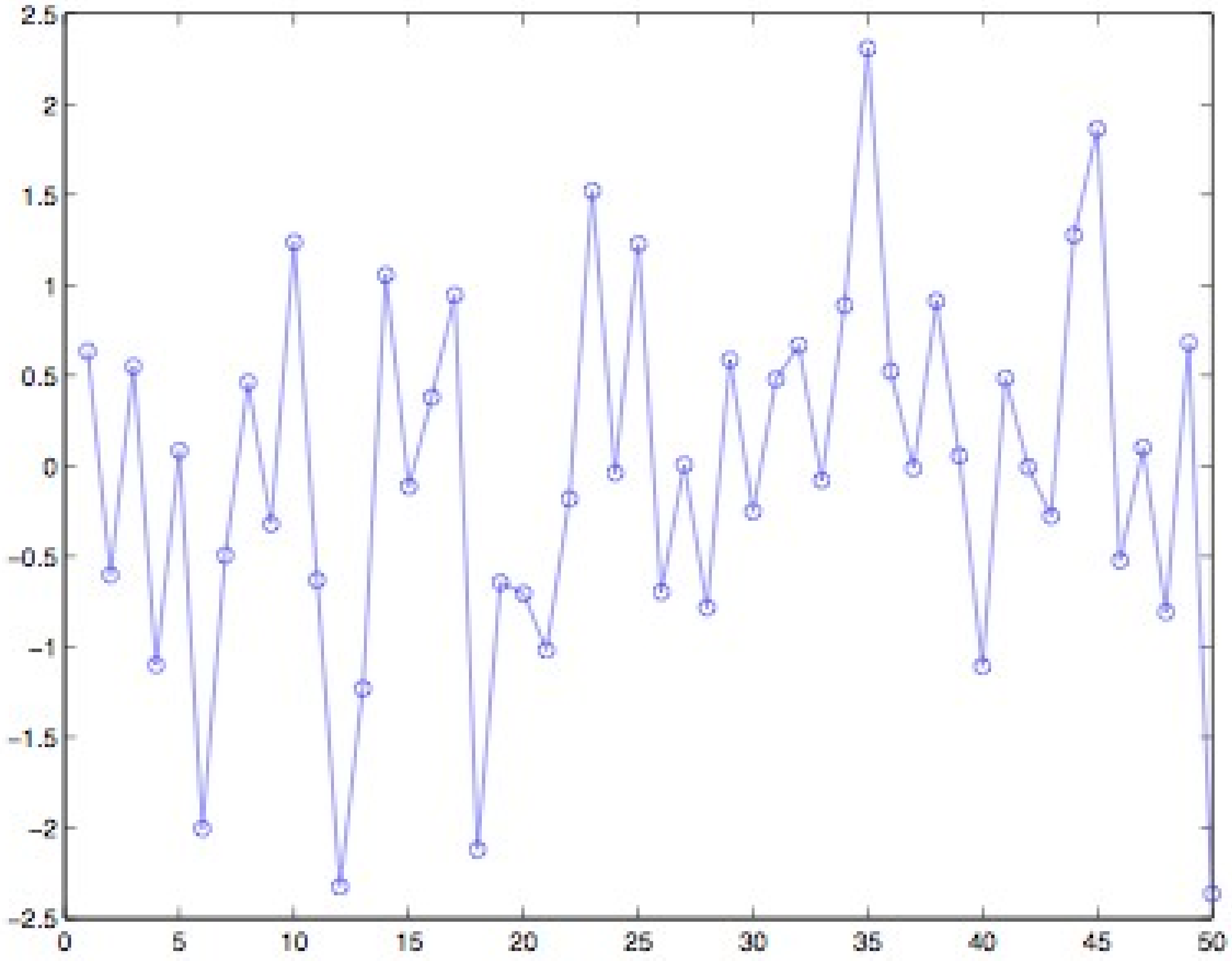
Example: i.i.d. noise: $\{X_t\}$ independent and identically distributed.

$$P[X_1 \le x_1, \ldots, X_t \le x_t] = P[X_1 \le x_1] \cdots P[X_t \le x_t].$$

Not interesting for forecasting:

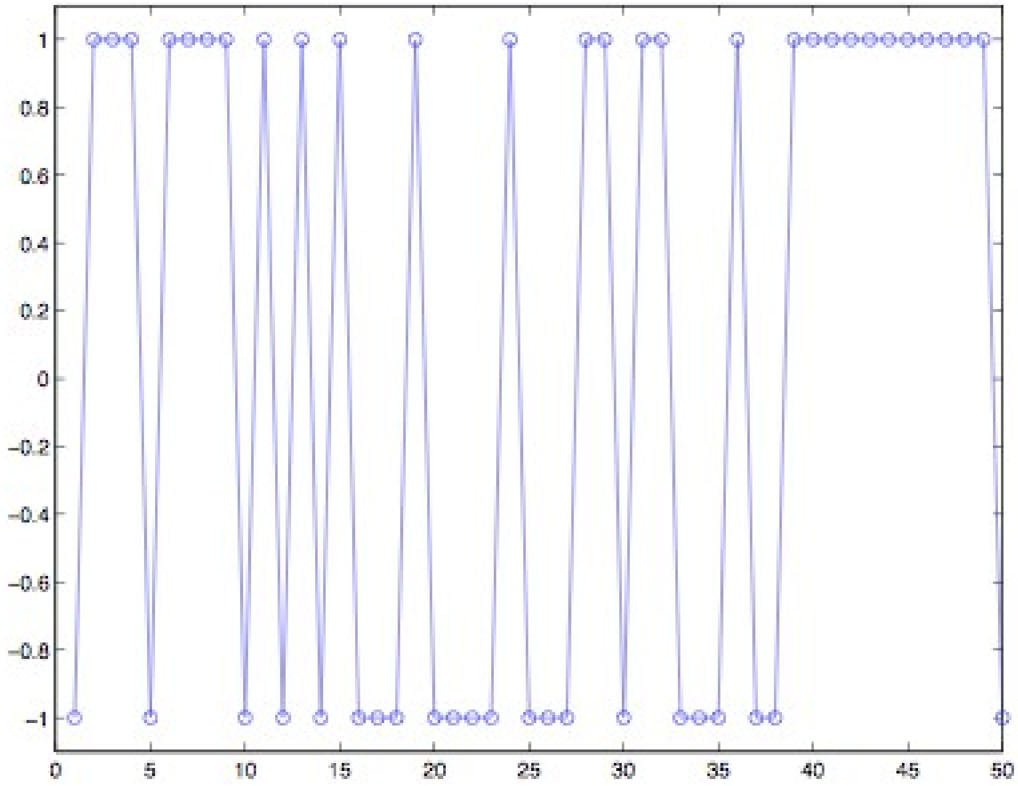
$$P[X_t \le x_t | X_1, \ldots, X_{t-1}] = P[X_t \le x_t].$$

Gaussian white noise



Time series models

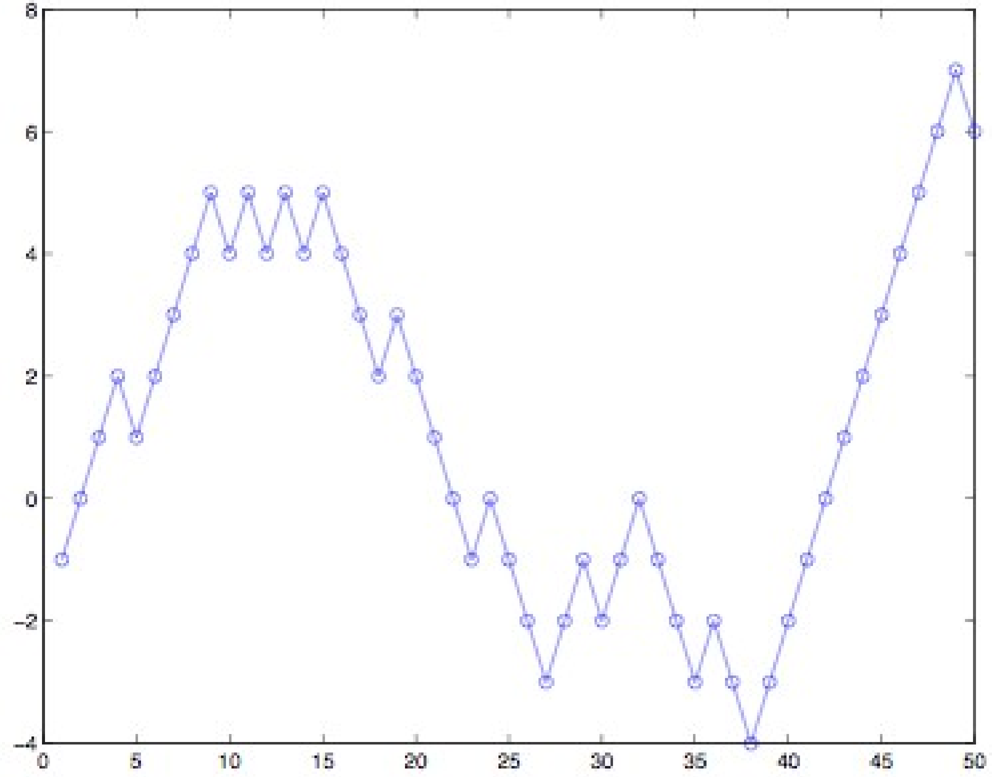
Example: Binary i.i.d. $P[X_t = 1] = P[X_t = -1] = 1/2.$



Random walk

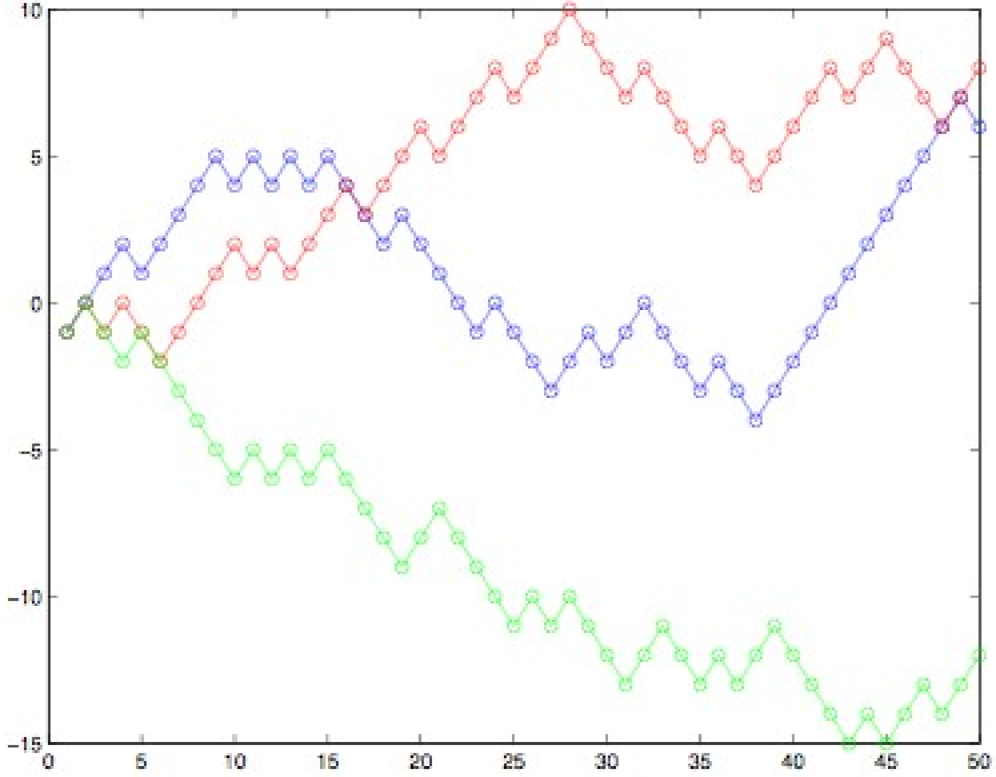
$S_t = \sum_{i=1}^{t} X_i.$

Differences: $\nabla S_t = S_t - S_{t-1} = X_t.$



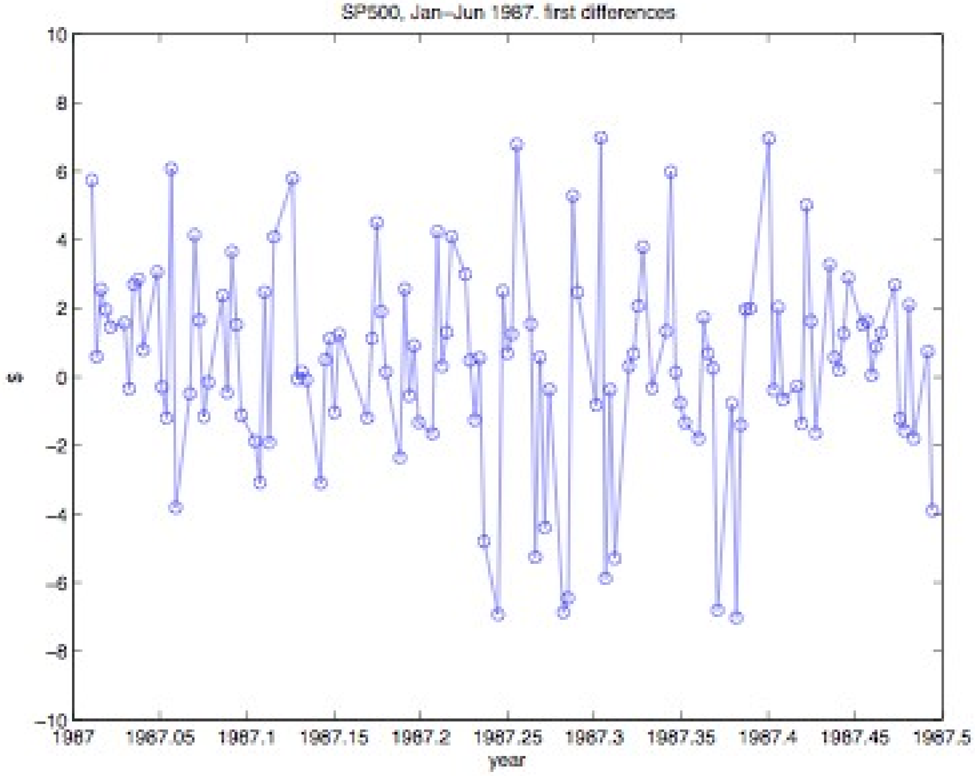
Random walk

$\mathrm{E}S_t$? $\mathrm{Var}S_t$?



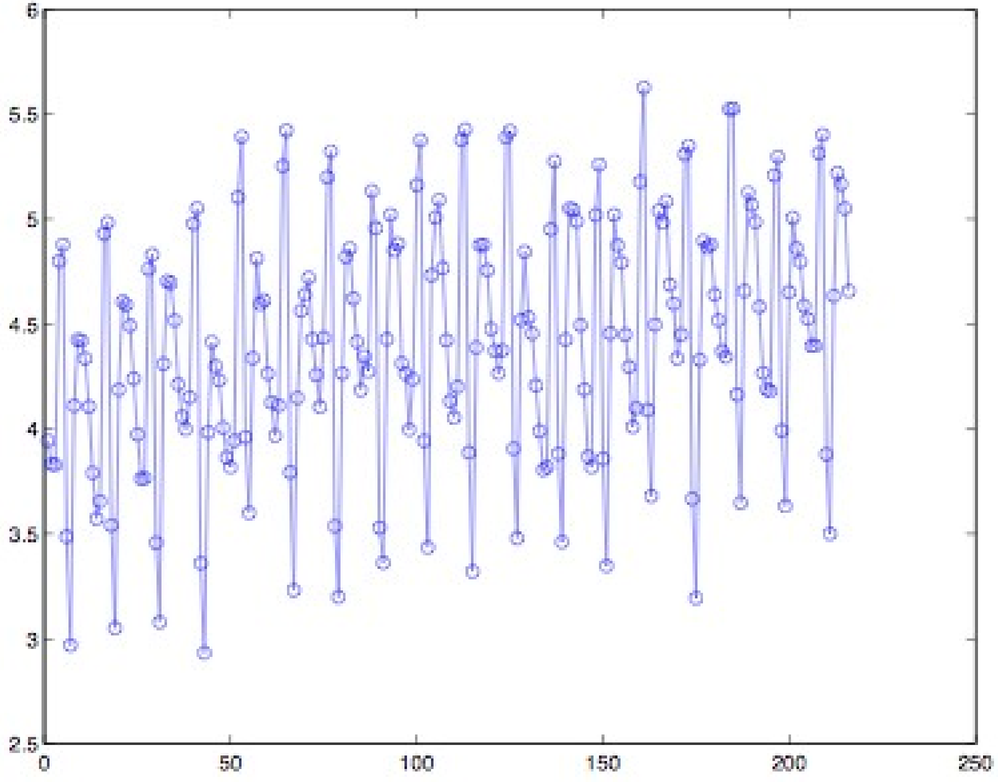
Random walk

Differences: $$\nabla S_t = S_t - S_{t-1} = X_t.$$



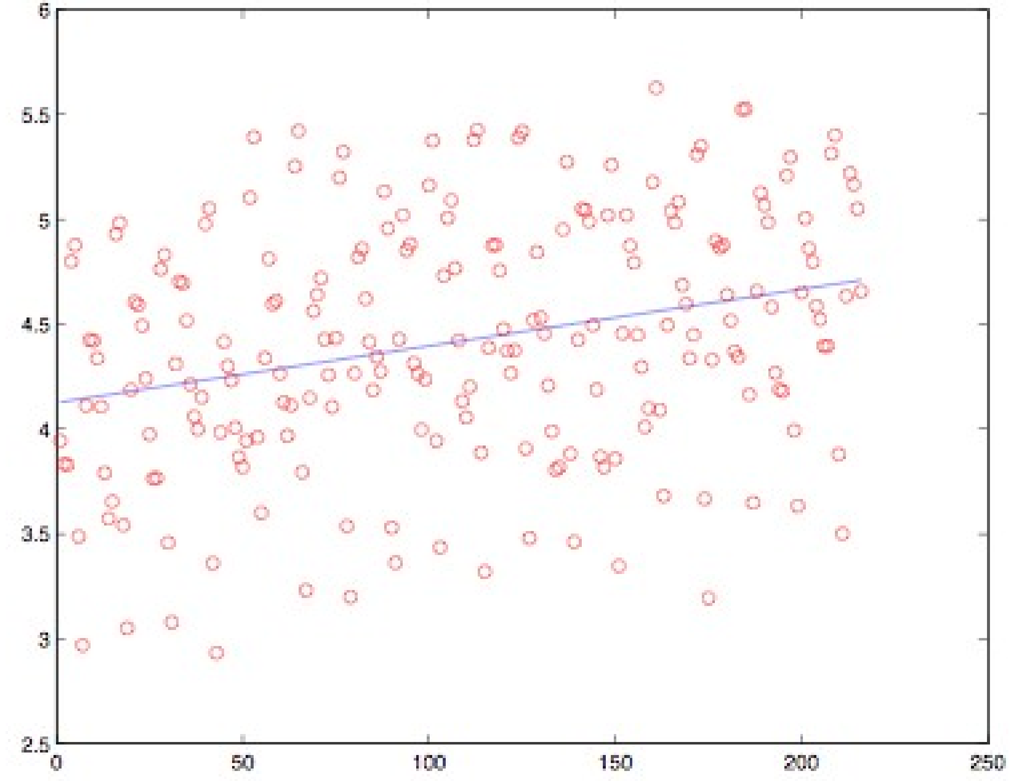
Trend and seasonal models

$$X_t = T_t + S_t + E_t = \beta_0 + \beta_1 t + \sum_i \left(\beta_i \cos(\lambda_i t) + \gamma_i \sin(\lambda_i t)\right) + E_t$$



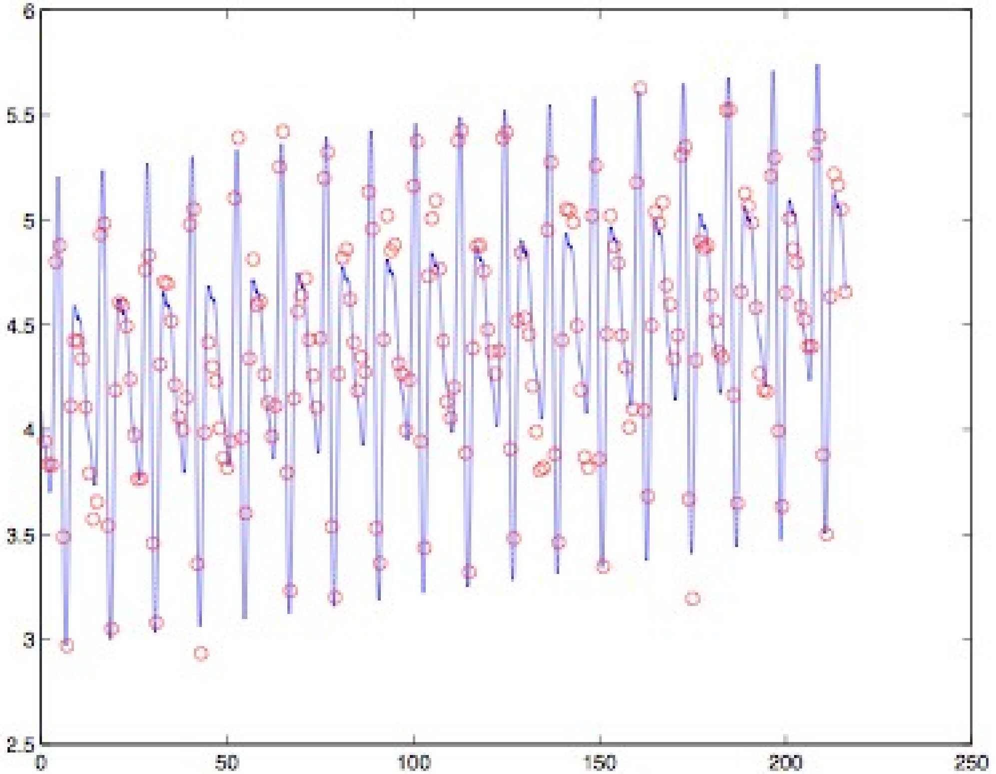
Trend and seasonal models

$$X_t = T_t + E_t = \beta_0 + \beta_1 t + E_t$$



Trend and seasonal models

$$X_t = T_t + S_t + E_t = \beta_0 + \beta_1 t + \sum_i \left(\beta_i \cos(\lambda_i t) + \gamma_i \sin(\lambda_i t)\right) + E_t$$



Time series modeling

1. Plot the time series.

   Look for trends, seasonal components, step changes, outliers.

2. Transform data so that residuals are **stationary**.

   (a) Estimate and subtract T_t, S_t.

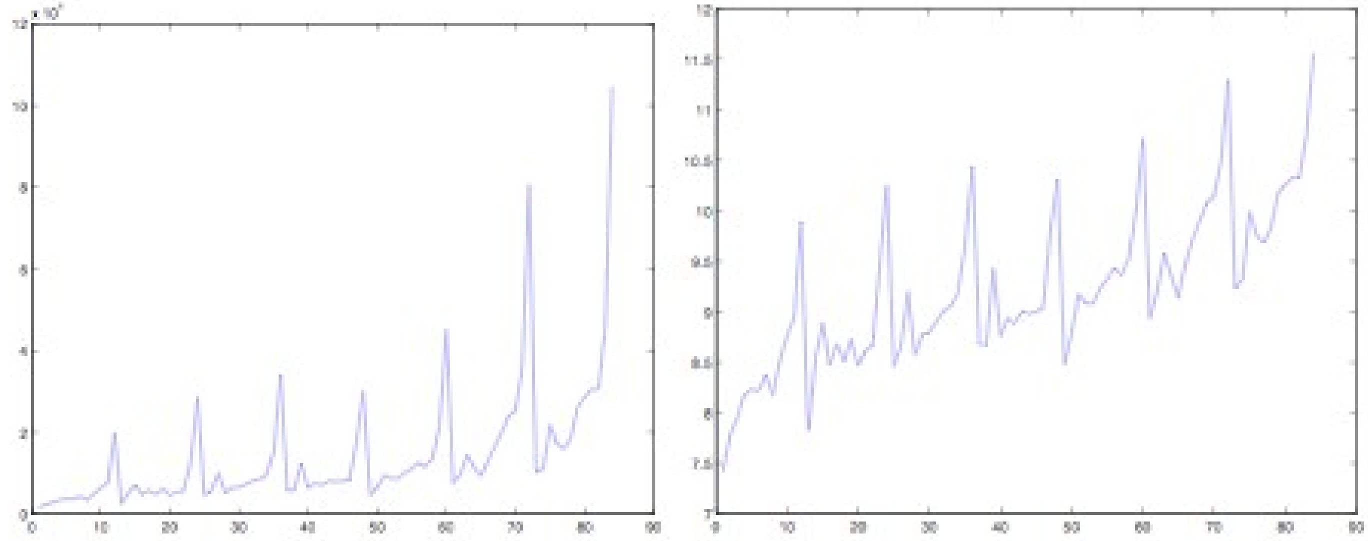
   (b) Differencing.

   (c) Nonlinear transformations (log, $\sqrt{\cdot}$).

3. Fit model to residuals.

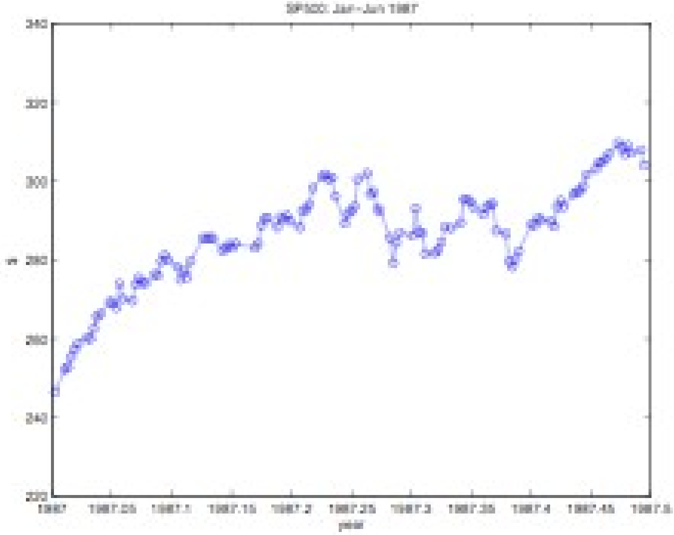
Nonlinear transformation

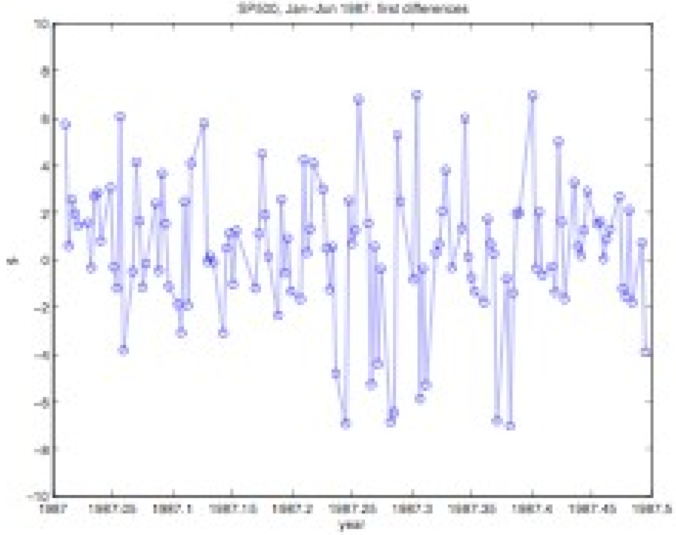
Recall: Monthly sales. (Makridakis, Wheelwright and Hyndman, 1998)



Differencing

Recall: S&P 500 data.





Differencing and trend

Define the lag-1 **difference operator**, (think 'first derivative')

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t,$$

where B is the **backshift** operator, $BX_t = X_{t-1}$.

- If $X_t = \beta_0 + \beta_1 t + Y_t$, then

$$\nabla X_t = \beta_1 + \nabla Y_t.$$

- If $X_t = \sum_{i=0}^{k} \beta_i t^i + Y_t$, then

$$\nabla^k X_t = k!\beta_k + \nabla^k Y_t,$$

where $\nabla^k X_t = \nabla(\nabla^{k-1} X_t)$ and $\nabla^1 X_t = \nabla X_t$.

Differencing and seasonal variation

Define the lag-s **difference operator**,

$$\nabla_s X_t = X_t - X_{t-s} = (1 - B^s) X_t,$$

where B^s is the backshift operator applied s times, $B^s X_t = B(B^{s-1} X_t)$ and $B^1 X_t = B X_t$.

If $X_t = T_t + S_t + Y_t$, and S_t has period s (that is, $S_t = S_{t-s}$ for all t), then

$$\nabla_s X_t = T_t - T_{t-s} + \nabla_s Y_t.$$

Stationarity

$\{X_t\}$ is **strictly stationary** if

for all $k, t_1, \ldots, t_k, x_1, \ldots, x_k$, and h,

$$P(X_{t_1} \leq x_1, \ldots, X_{t_k} \leq x_k) = P(x_{t_1+h} \leq x_1, \ldots, X_{t_k+h} \leq x_k).$$

i.e., shifting the time axis does not affect the distribution.

We shall consider **second-order properties** only.

Mean and Autocovariance

Suppose that $\{X_t\}$ is a time series with $\mathrm{E}[X_t^2] < \infty$.
Its **mean function** is

$$\mu_t = \mathrm{E}[X_t].$$

Its **autocovariance function** is

$$\begin{aligned}\gamma_X(s,t) &= \mathrm{Cov}(X_s, X_t) \\ &= \mathrm{E}[(X_s - \mu_s)(X_t - \mu_t)].\end{aligned}$$

Weak stationarity

We say that $\{X_t\}$ is **(weakly) stationary** if

1. μ_t is independent of t, and
2. For each h, $\gamma_X(t+h,t)$ is independent of t.

In that case, we write

$$\gamma_X(h) = \gamma_X(h,0).$$

Stationarity

The **autocorrelation function (ACF)** of $\{X_t\}$ is defined as

$$\begin{aligned}\rho_X(h) &= \frac{\gamma_X(h)}{\gamma_X(0)}\\ &= \frac{\mathrm{Cov}(X_{t+h}, X_t)}{\mathrm{Cov}(X_t, X_t)}\\ &= \mathrm{Corr}(X_{t+h}, X_t).\end{aligned}$$

Stationarity

Example: i.i.d. noise, $\mathrm{E}[X_t] = 0, \mathrm{E}[X_t^2] = \sigma^2$. We have

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2 & \text{if } h = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Thus,

1. $\mu_t = 0$ is independent of t.
2. $\gamma_X(t+h, t) = \gamma_X(h, 0)$ for all t.

So $\{X_t\}$ is stationary.

Similarly for any white noise (uncorrelated, zero mean), $X_t \sim WN(0, \sigma^2)$.

Stationarity

Example: Random walk, $S_t = \sum_{i=1}^{t} X_i$ for i.i.d., mean zero $\{X_t\}$.
We have $\mathrm{E}[S_t] = 0, \mathrm{E}[S_t^2] = t\sigma^2$, and

$$
\begin{aligned}
\gamma_S(t+h, t) &= \mathrm{Cov}(S_{t+h}, S_t) \\
&= \mathrm{Cov}\left(S_t + \sum_{s=1}^{h} X_{t+s}, S_t\right) \\
&= \mathrm{Cov}(S_t, S_t) = t\sigma^2.
\end{aligned}
$$

1. $\mu_t = 0$ is independent of t, but
2. $\gamma_S(t+h, t)$ is not.

So $\{S_t\}$ is not stationary.

Covariances

$$\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z),$$

$$\mathrm{Cov}(aX,Y) = a\,\mathrm{Cov}(X,Y),$$

Also if X and Y are independent (e.g., $X = c$), then

$$\mathrm{Cov}(X,Y) = 0.$$

Stationarity

Example: MA(1) process (**Moving Average**):

$$X_t = W_t + \theta W_{t-1}, \qquad \{W_t\} \sim WN(0, \sigma^2).$$

We have $\mathrm{E}[X_t] = 0$, and

$$\begin{aligned}\gamma_X(t+h, t) &= \mathrm{E}(X_{t+h} X_t) \\ &= \mathrm{E}[(W_{t+h} + \theta W_{t+h-1})(W_t + \theta W_{t-1})] \\ &= \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2 \theta & \text{if } h = \pm 1, \\ 0 & \text{otherwise.} \end{cases}\end{aligned}$$

Thus, $\{X_t\}$ is stationary.

Stationarity

Example: AR(1) process (**AutoRegressive**):

$$X_t = \phi X_{t-1} + W_t, \qquad \{W_t\} \sim WN(0, \sigma^2).$$

Assume that X_t is stationary and $|\phi| < 1$. Then we have

$$\begin{aligned} \mathrm{E}[X_t] &= \phi \mathrm{E} X_{t-1} \\ &= 0 \qquad \text{(from stationarity)} \\ \mathrm{E}[X_t^2] &= \phi^2 \mathrm{E}[X_{t-1}^2] + \sigma^2 \\ &= \frac{\sigma^2}{1-\phi^2} \qquad \text{(from stationarity)}, \end{aligned}$$

Stationarity

Example: AR(1) process, $X_t = \phi X_{t-1} + W_t, \qquad \{W_t\} \sim WN(0, \sigma^2)$.
Assume that X_t is stationary and $|\phi| < 1$. Then we have

$$\mathrm{E}[X_t] = 0, \qquad \mathrm{E}[X_t^2] = \frac{\sigma^2}{1 - \phi^2}$$

$$\begin{aligned}
\gamma_X(h) &= \mathrm{Cov}(\phi X_{t+h-1} + W_{t+h}, X_t) \\
&= \phi \mathrm{Cov}(X_{t+h-1}, X_t) \\
&= \phi \gamma_X(h-1) \\
&= \phi^{|h|} \gamma_X(0) \qquad \text{(check for } h > 0 \text{ and } h < 0\text{)} \\
&= \frac{\phi^{|h|} \sigma^2}{1 - \phi^2}.
\end{aligned}$$

Linear process

An important class of stationary time series:

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

where $\{W_t\} \sim WN(0, \sigma_w^2)$

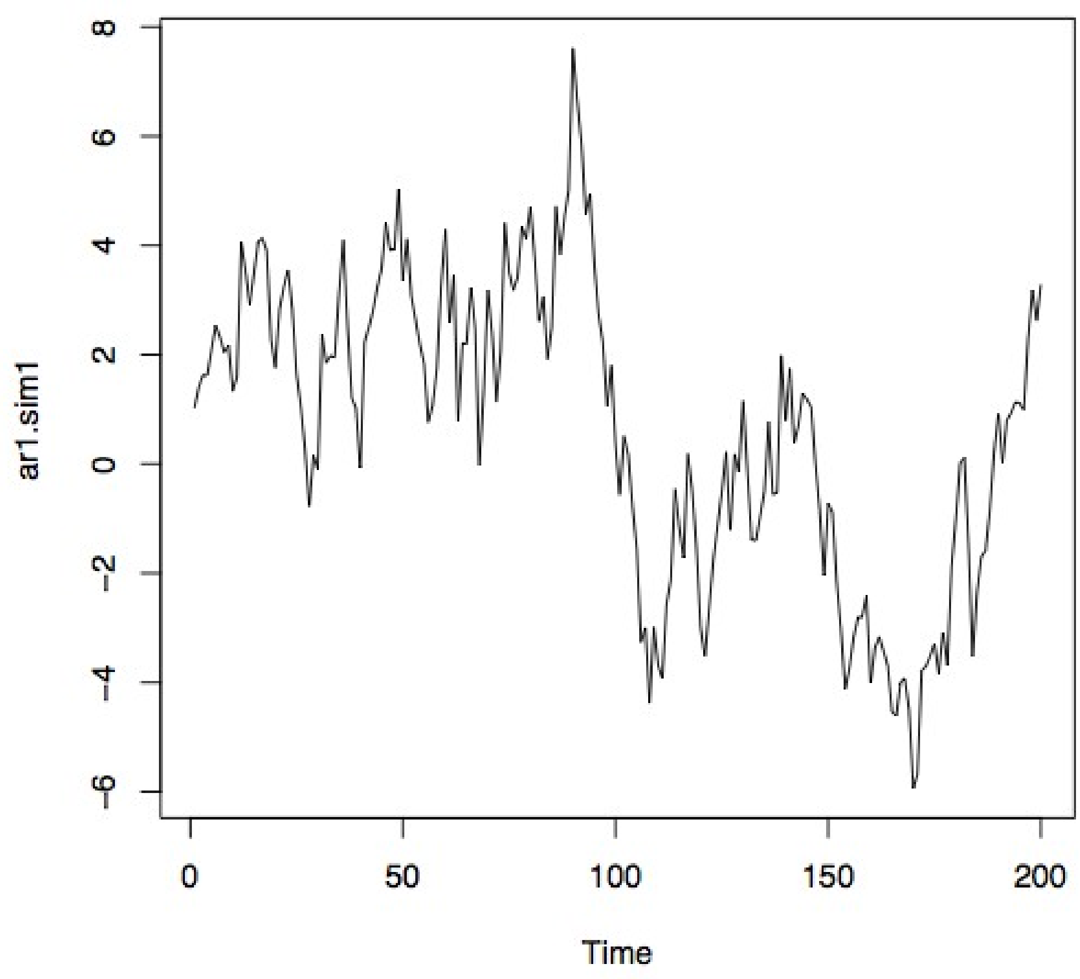
and μ, ψ_j are parameters satisfying

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

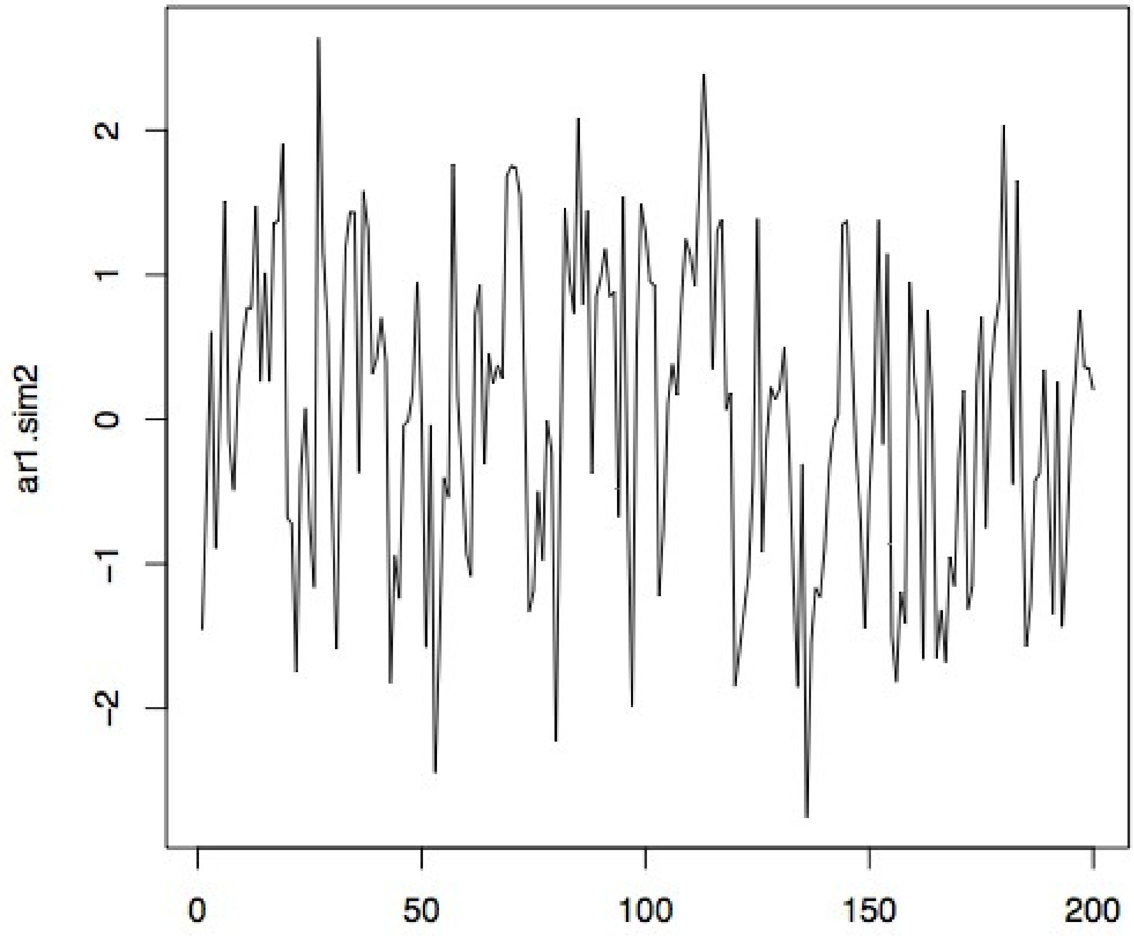
Examples:

- White noise: $\psi_0 = 1$.
- MA(1): $\psi_0 = 1, \psi_1 = \theta$.
- AR(1): $\psi_0 = 1, \psi_1 = \phi, \psi_2 = \phi^2, \ldots$

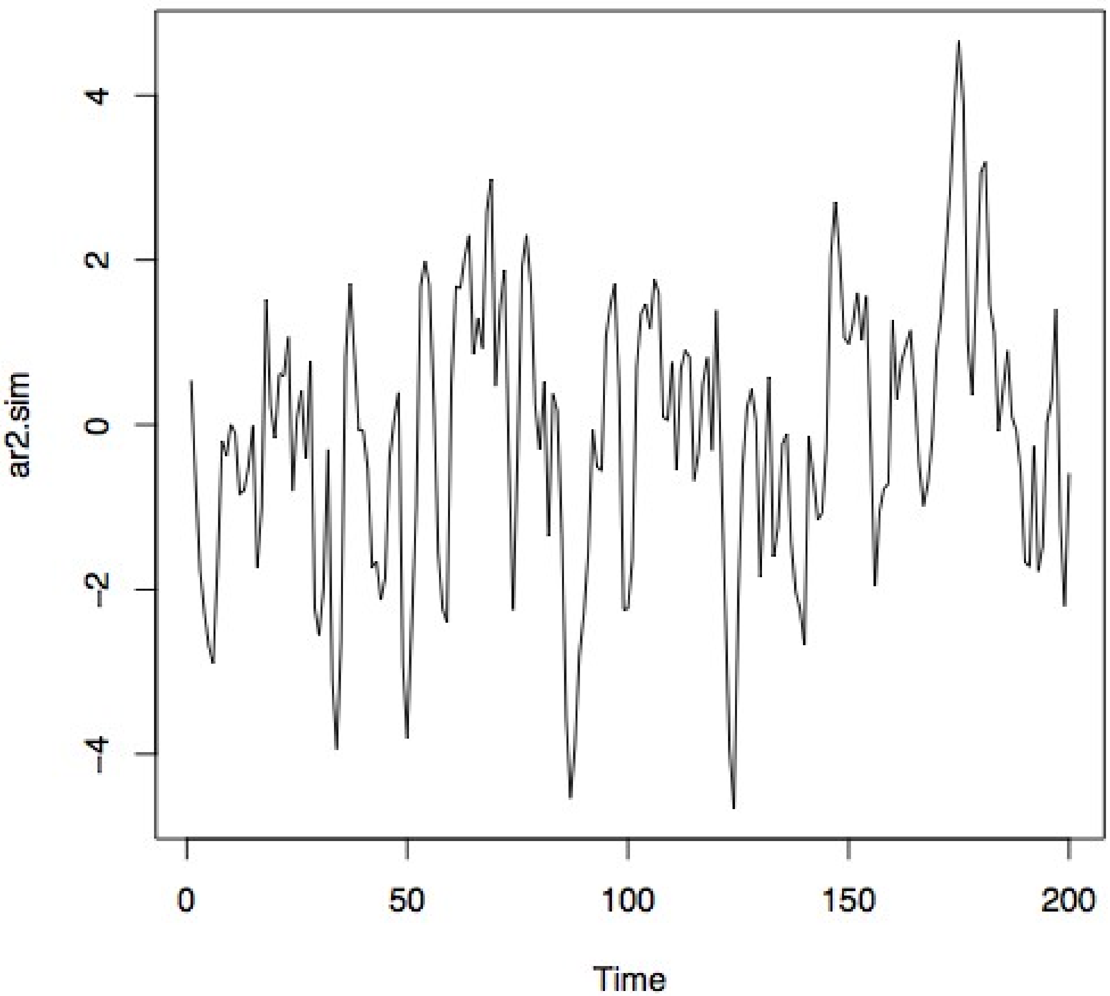
AR(1) ϕ :0.95



AR(1) ϕ:0.5



AR(2): 0.9, 0.2



Sample ACF

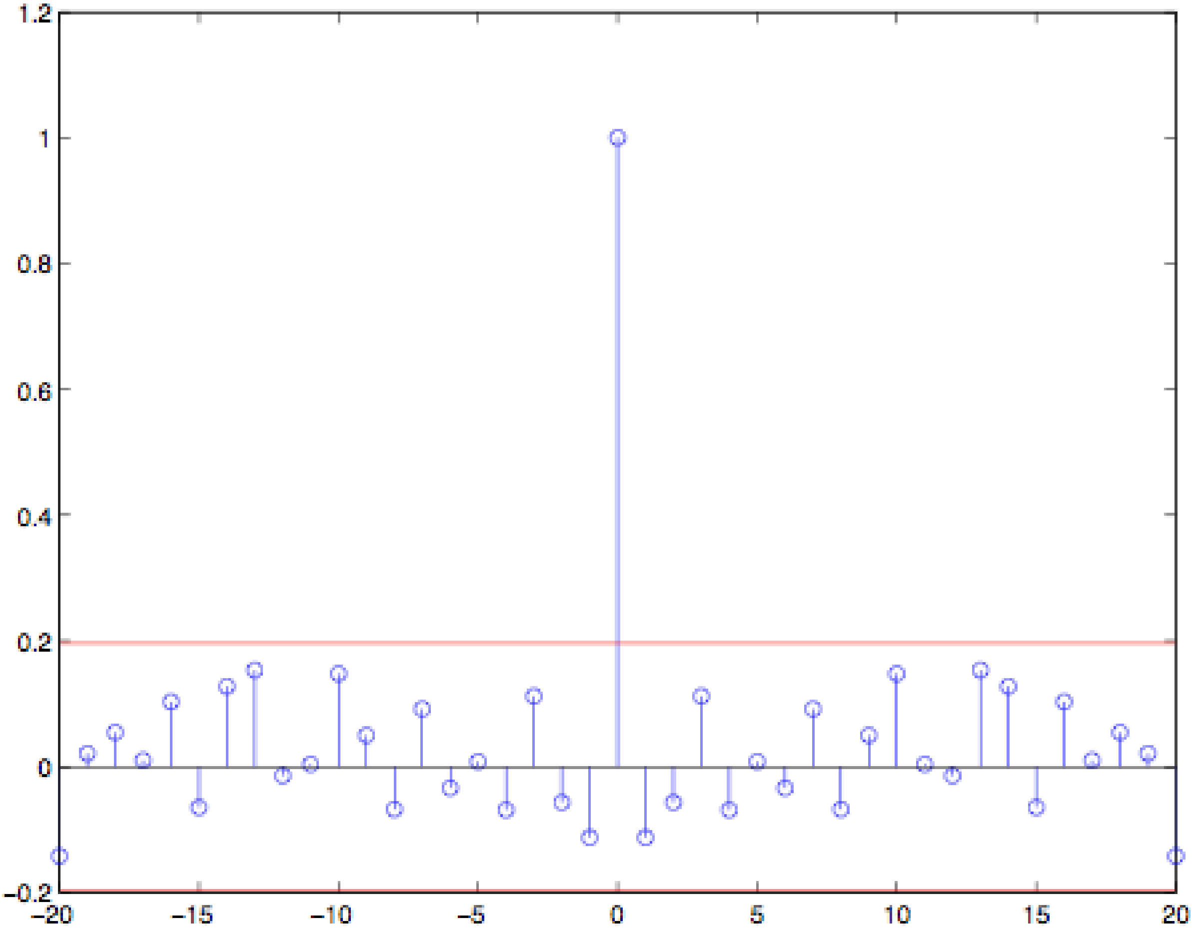
Sample autocovariance function:

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}).$$

$\approx$ the sample covariance of $(x_1, x_{h+1}), \ldots, (x_{n-h}, x_n)$, except that

- we normalize by n instead of $n - h$, and
- we subtract the full sample mean.

Sample ACF for Gaussian noise

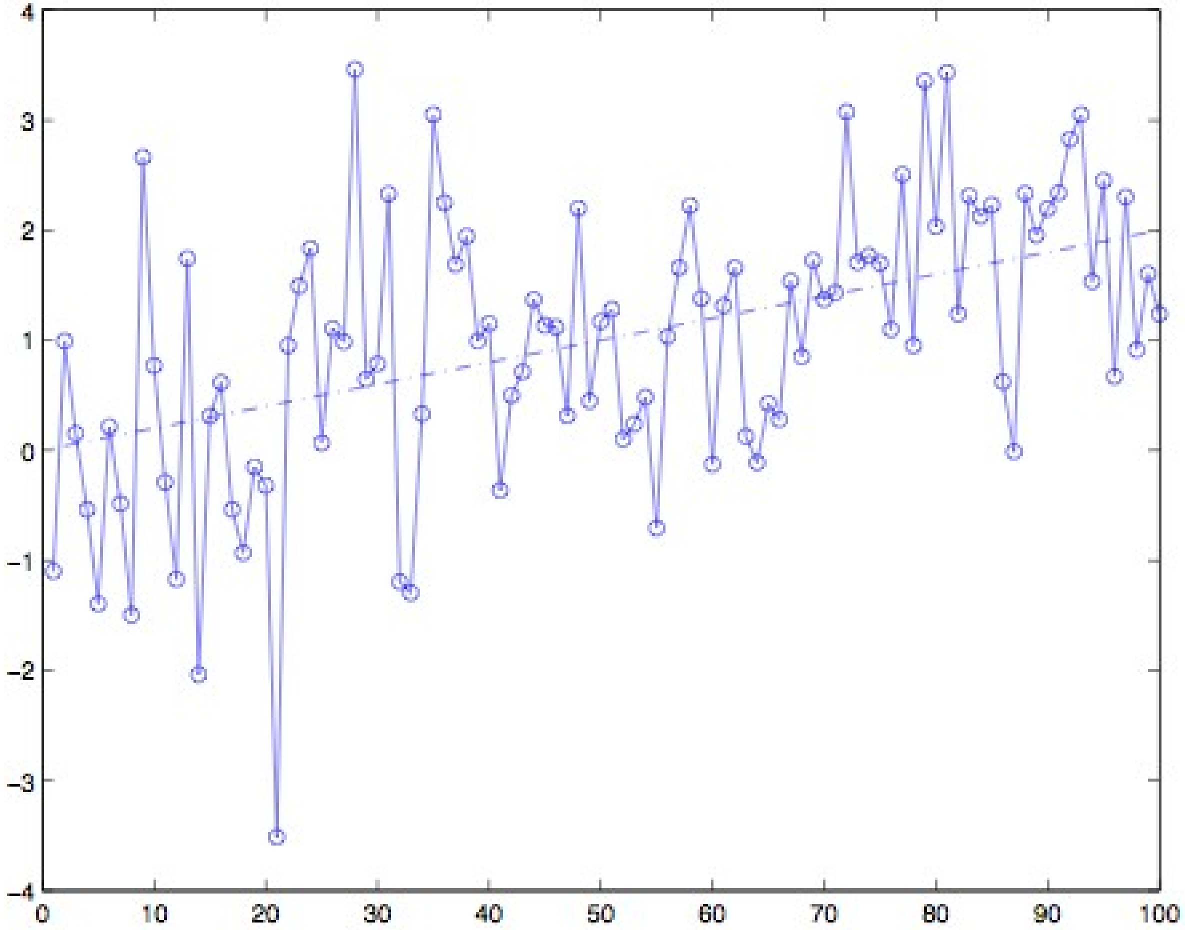


Summary for sample ACF

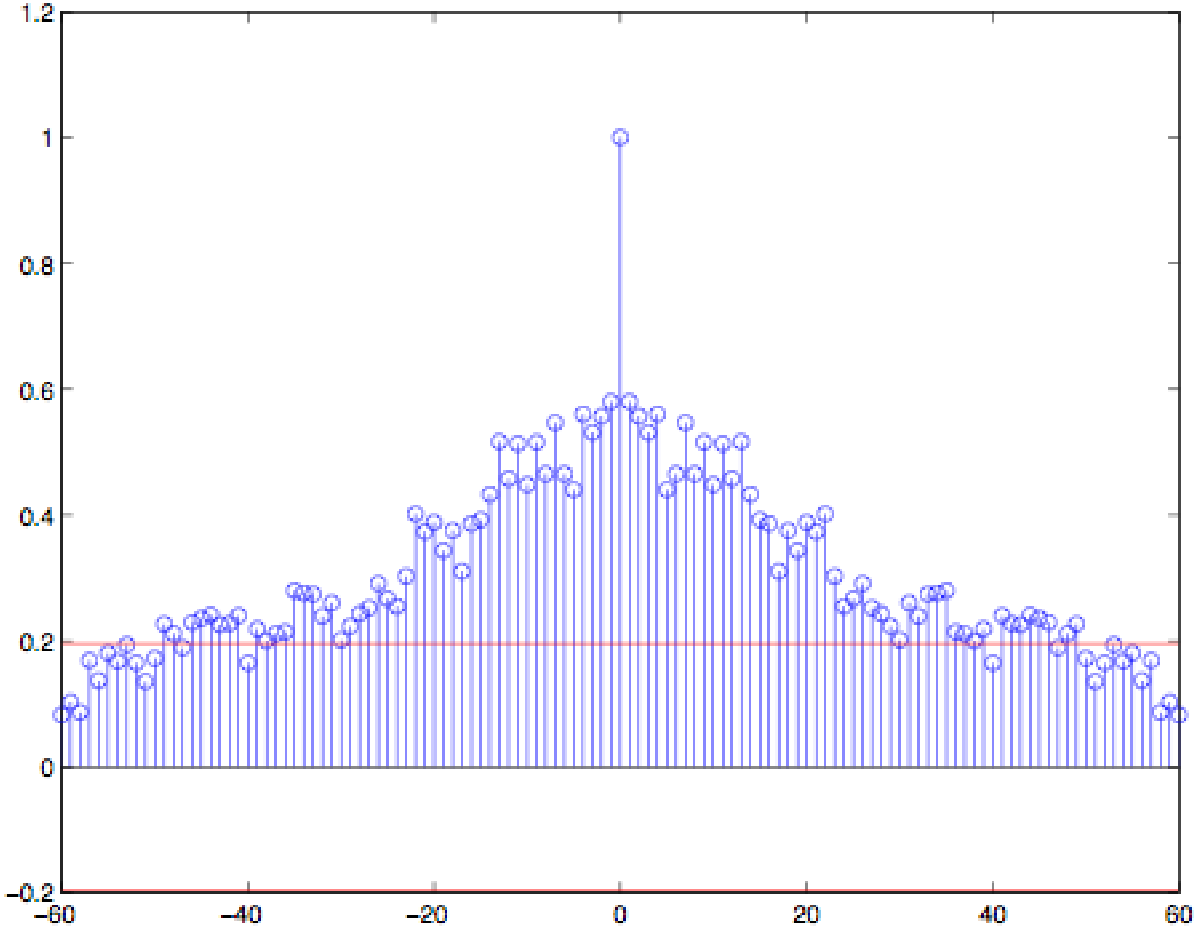
We can recognize the sample autocorrelation functions of many non-white (even non-stationary) time series.

Time series:	**Sample ACF:**
White	zero
Trend	Slow decay
Periodic	Periodic
MA(q)	Zero for $\lvert h \rvert > q$
AR(p)	Decays to zero exponentially

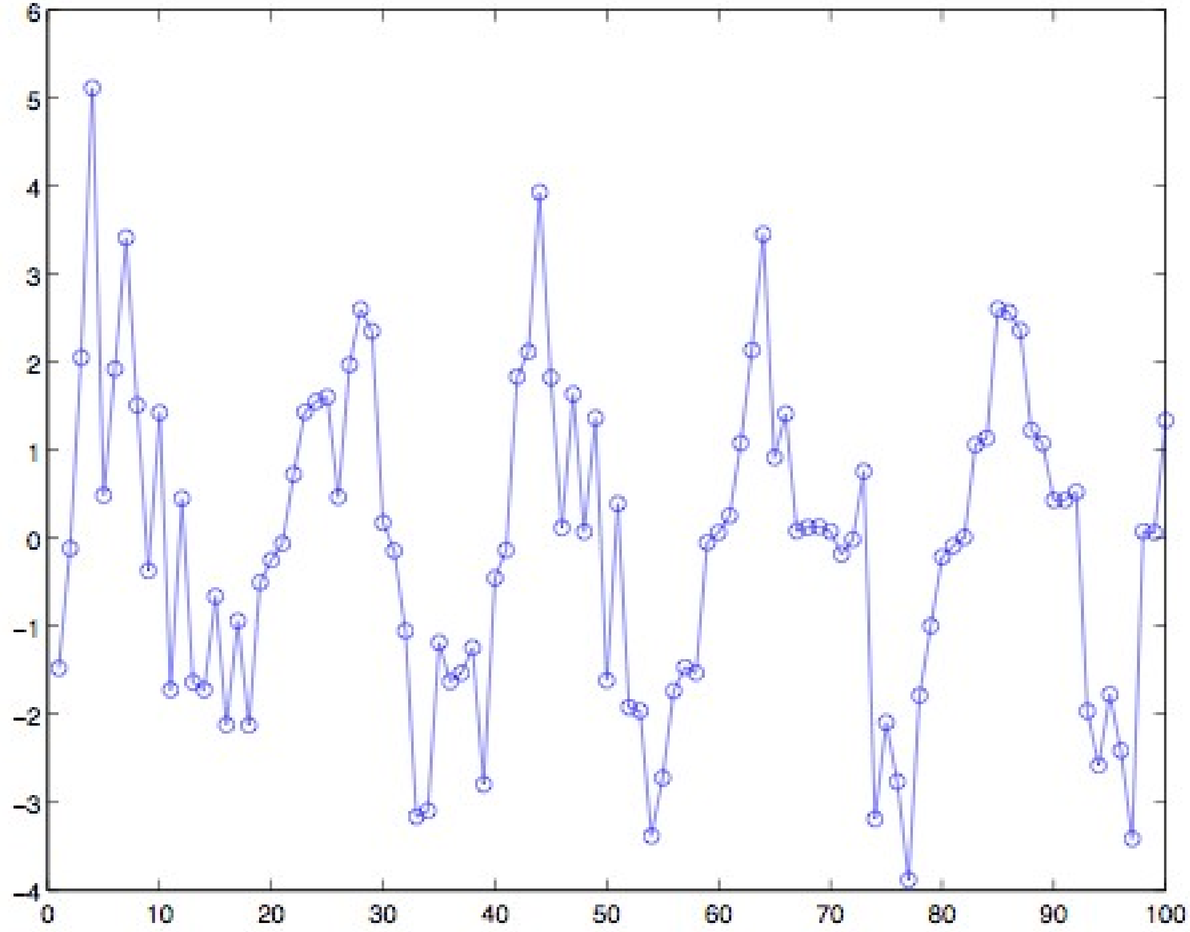
Trend



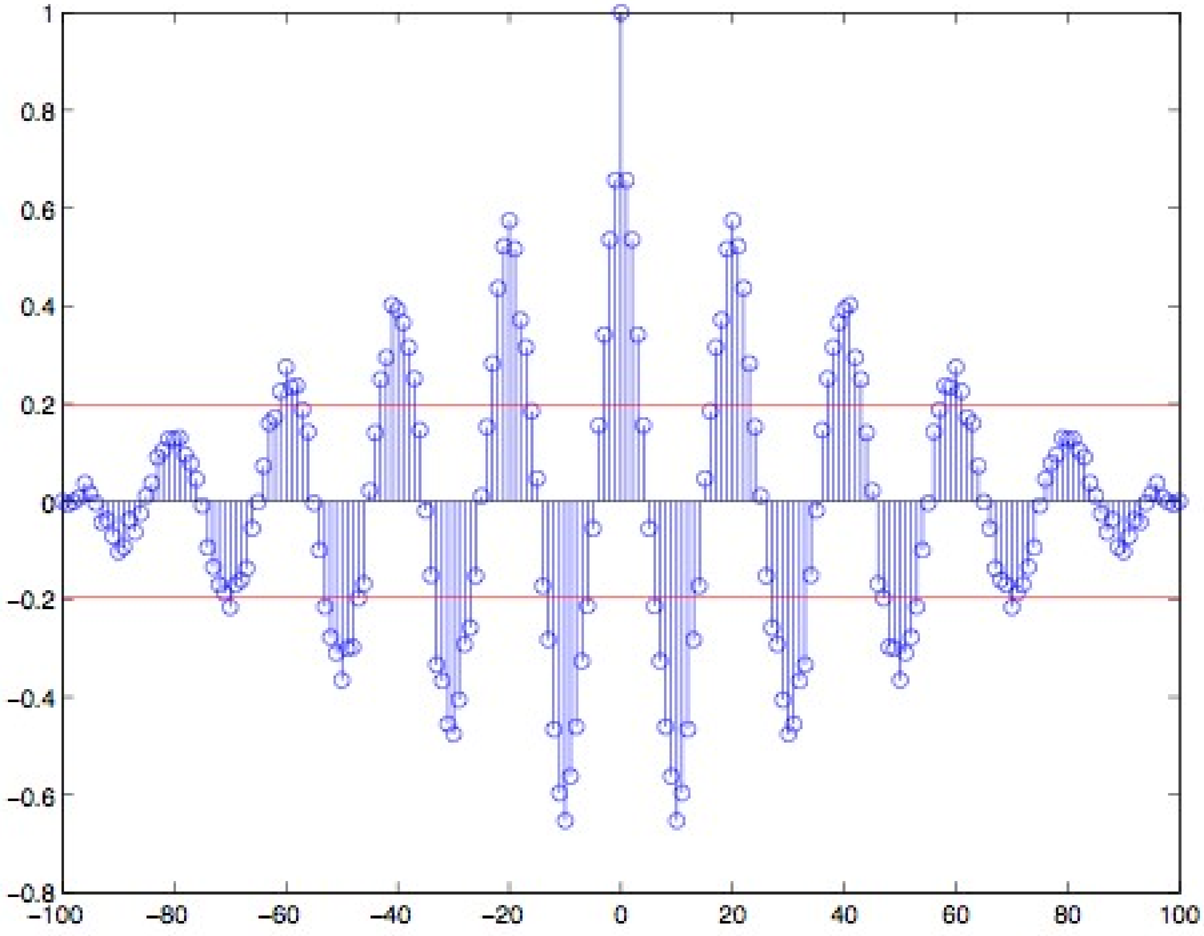
Sample ACF: Trend



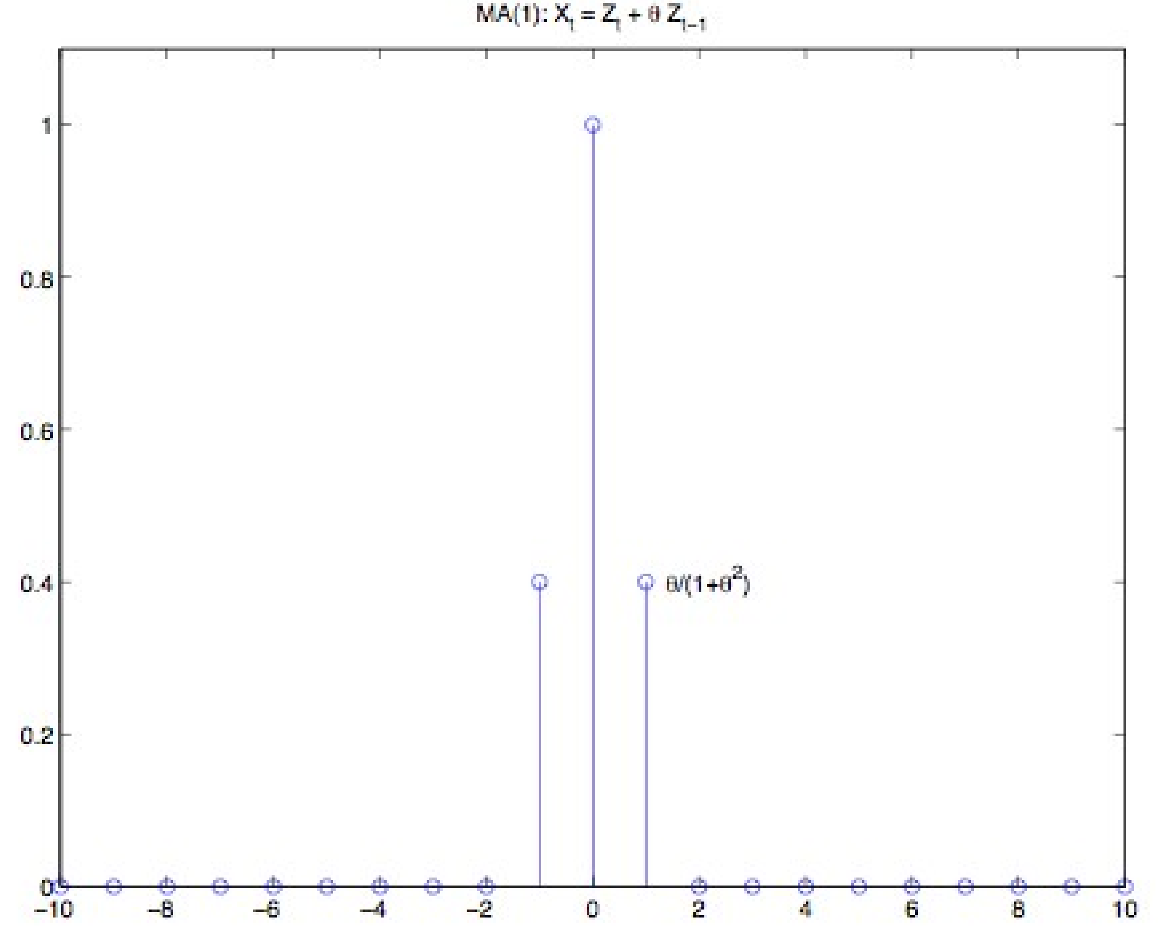
Periodic



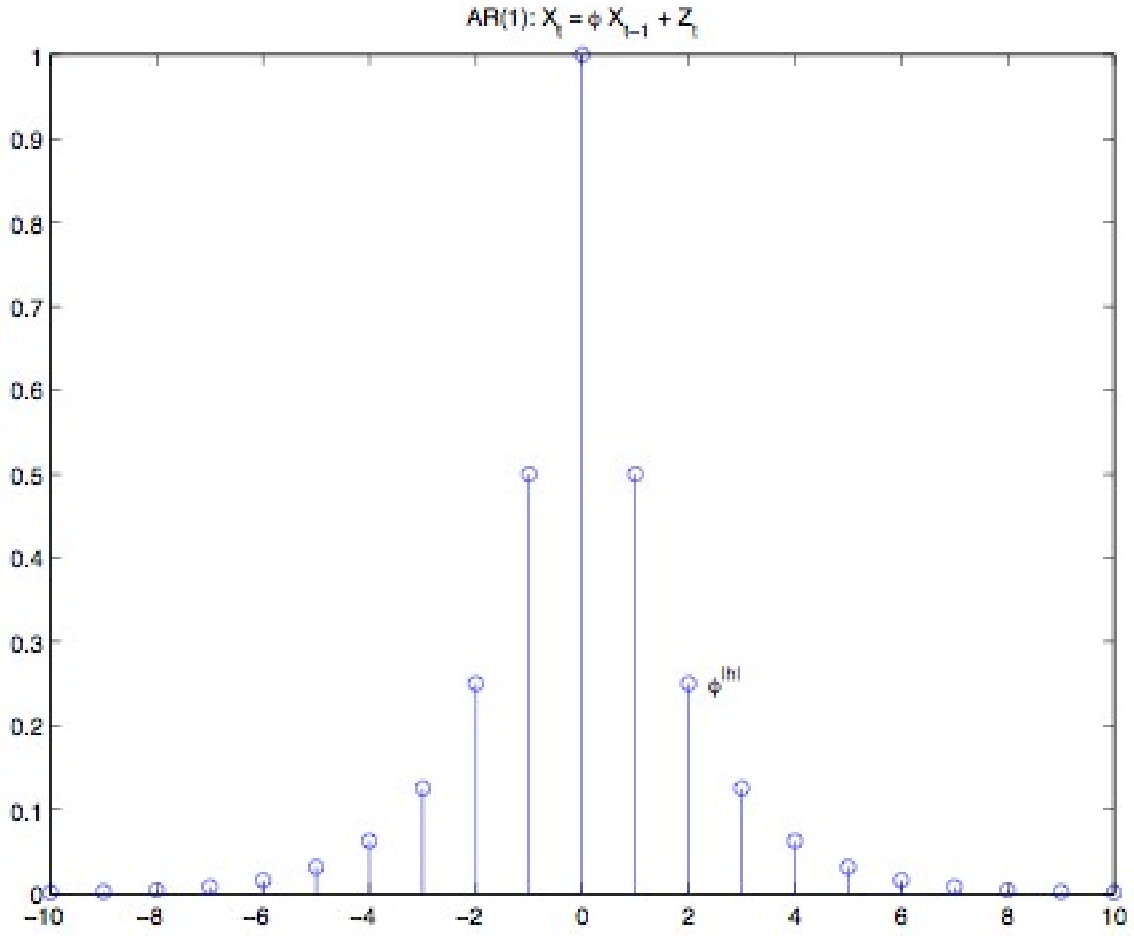
Sample ACF: Periodic



ACF: MA(1)


MA(1): $X_t = Z_t + \theta Z_{t-1}$
$\theta/(1+\theta^2)$
1
0.8
0.6
0.4
0.2
0
−10
−8
−6
−4
−2
0
2
4
6
8
10

ACF: AR



AR(1): $X_t = \phi X_{t-1} + Z_t$
$\phi^{|h|}$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−10
−8
−6
−4
−2
0
2
4
6
8
10

ARMA

An **ARMA(p,q) process** $\{X_t\}$ is a stationary process that satisfies

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q},$$

where $\{W_t\} \sim WN(0, \sigma^2)$.

Also, $\phi_p, \theta_q \neq 0$ and $\phi(z), \theta(z)$ have no common factors.