Chapter 6

Univariate time series modelling and forecasting

Univariate Time Series Models

- Where we attempt to predict returns using only information contained in their past values.

Some Notation and Concepts

- A Strictly Stationary Process

A strictly stationary process is one where

$$P\{y_{t_1} \leq b_1, \ldots, y_{t_n} \leq b_n\} = P\{y_{t_1+m} \leq b_1, \ldots, y_{t_n+m} \leq b_n\}$$

- A Weakly Stationary Process

Univariate Time Series Models (Cont'd)

If a series satisfies the next three equations, it is said to be weakly or covariance stationary

$$
\begin{aligned}
&(1)\;\; E(y_t) = \mu \quad t = 1, 2, \ldots, \infty \\
&(2)\;\; E(y_t - \mu)(y_t - \mu) = \sigma^2 < \infty \\
&(3)\;\; E(y_{t_1} - \mu)(y_{t_2} - \mu) = \gamma_{t_2 - t_1} \quad \forall\, t_1, t_2
\end{aligned}
$$

- So if the process is covariance stationary, all the variances are the same and all the covariances depend on the difference between t_1 and t_2. The moments

$$E(y_t - E(y_t))(y_{t-s} - E(y_{t-s})) = \gamma_s, s = 0, 1, 2, \ldots$$

are known as the covariance function.

Univariate Time Series Models (Cont'd)

- The covariances, γ_s, are known as autocovariances.
- However, the value of the autocovariances depend on the units of measurement of y_t.
- It is thus more convenient to use the autocorrelations which are the autocovariances normalised by dividing by the variance:

$$\tau_s = \frac{\gamma_s}{\gamma_0}, \quad s = 0, 1, 2, \ldots$$

- If we plot τ_s against s=0,1,2,... then we obtain the autocorrelation function or correlogram.

A White Noise Process

- A white noise process is one with (virtually) no discernible structure. A definition of a white noise process is

$$E(y_t) = \mu$$
$$\mathrm{var}(y_t) = \sigma^2$$
$$\gamma_{t-r} = \begin{cases} \sigma^2 & \text{if} \quad t = r \\ 0 & \quad \textit{otherwise} \end{cases}$$

- Thus the autocorrelation function will be zero apart from a single peak of 1 at s=0. $\hat{\tau}_s \sim \textit{approx}.\ N(0, 1/T)$ where T = sample size

- We can use this to do significance tests for the autocorrelation coefficients by constructing a confidence interval.

A White Noise Process (Cont'd)

- For example, a 95 % confidence interval would be given by

$$\pm 1.96 \times \frac{1}{\sqrt{T}}$$

If the sample autocorrelation coefficient, $\hat{\tau}_s$, falls outside this region for any value of s, then we reject the null hypothesis that the true value of the coefficient at lag s is zero.

Joint Hypothesis Tests

- We can also test the joint hypothesis that all m of the τ_k correlation coefficients are simultaneously equal to zero using the Q-statistic developed by Box and Pierce:

$$Q = T \sum_{k=1}^{m} \hat{\tau}_k^2$$

where T=sample size, m=maximum lag length

- The Q-statistic is asymptotically distributed as a χ_m^2.
- However, the Box Pierce test has poor small sample properties, so a variant has been developed, called the Ljung-Box statistic:

$$Q^* = T(T+2) \sum_{k=1}^{m} \frac{\hat{\tau}_k^2}{T-k} \sim \chi_m^2$$

- This statistic is very useful as a portmanteau (general) test of linear dependence in time series.

An ACF Example

- Question:

  Suppose that a researcher had estimated the first 5 autocorrelation coefficients using a series of length 100 observations, and found them to be (from 1 to 5): 0.207, -0.013, 0.086, 0.005, -0.022.

  Test each of the individual coefficient for significance, and use both the Box-Pierce and Ljung-Box tests to establish whether they are jointly significant.

- Solution:

  A coefficient would be significant if it lies outside (-0.196,+0.196) at the 5% level, so only the first autocorrelation coefficient is significant.

  Q=5.09 and Q*=5.26

  Compared with a tabulated $\chi^2(5)$=11.1 at the 5% level, so the 5 coefficients are jointly insignificant.

Moving Average Processes

- Let u_t $(t = 1, 2, 3, \dots)$ be a sequence of independently and identically distributed (iid) random variables with $E(u_t) = 0$ and $\text{var}(u_t) = \sigma^2$, then

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q}$$

  is a qth order moving average model MA(q).

- Its properties are

$$\begin{aligned} \mathrm{E}(y_t) &= \mu \\ \mathrm{var}(y_t) &= \gamma_0 = \left(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2\right)\sigma^2 \end{aligned}$$

  Covariances

$$\gamma_s = \begin{cases} (\theta_s + \theta_{s+1}\theta_1 + \theta_{s+2}\theta_2 + \cdots + \theta_q\theta_{q-s}) \quad \sigma^2 & \textit{for} \;\; s = 1, \dots, q \\ 0 & \textit{for} \quad s > q \end{cases}$$

Example of an MA Problem

1. Consider the following MA(2) process:

$$y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$$

where u_t is a zero mean white noise process with variance σ^2.

   i. Calculate the mean and variance of X_t

   ii. Derive the autocorrelation function for this process (i.e. express the autocorrelations, τ_1, τ_2, ...as functions of the parameters θ_1 and θ_2).

   iii. If $\theta_1 = -0.5$ and $\theta_2 = 0.25$, sketch the acf of X_t.

Solution

1. If $E(u_t) = 0$, then $\mathrm{E}(u_{t-i}) = 0 \; \forall \; i$ So

$$\begin{aligned} \mathrm{E}(y_t) &= \mathrm{E}(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}) \\ &= \mathrm{E}(u_t) + \theta_1 \mathrm{E}(u_{t-1}) + \theta_2 \mathrm{E}(u_{t-2}) = 0 \\ \mathrm{var}(y_t) &= \mathrm{E}[y_t - \mathrm{E}(y_t)][y_t - \mathrm{E}(y_t)] \end{aligned}$$

But $\mathrm{E}(y_t) = 0$, so

$$\begin{aligned} \mathrm{var}(y_t) &= \mathrm{E}[(y_t)(y_t)] \\ \mathrm{var}(y_t) &= \mathrm{E}[(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})] \\ \mathrm{var}(y_t) &= \mathrm{E}\left[u_t^2 + \theta_1^2 u_{t-1}^2 + \theta_2^2 u_{t-2}^2 + \textit{cross-products}\right] \end{aligned}$$

But $\mathrm{E}[\textit{cross-products}] = 0$ since $\mathrm{cov}(u_t, u_{t-s}) = 0$ for $s \neq 0$.

Solution (Cont'd)

So
$$\begin{aligned} \operatorname{var}(y_t) &= \gamma_0 = E\left[u_t^2 + \theta_1^2 u_{t-1}^2 + \theta_2^2 u_{t-2}^2\right] \\ &= \sigma^2 + \theta_1^2\sigma^2 + \theta_2^2\sigma^2 \\ &= \left(1 + \theta_1^2 + \theta_2^2\right)\sigma^2 \end{aligned}$$

Solution (Cont'd)

ii. The acf of y_t

$$\begin{aligned}
\gamma_1 &= \mathrm{E}[y_t - \mathsf{E}(y_t)][y_{t-1} - \mathsf{E}(y_{t-1})] \\
\gamma_1 &= \mathrm{E}[y_t][y_{t-1}] \\
\gamma_1 &= \mathrm{E}[(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})(u_{t-1} + \theta_1 u_{t-2} + \theta_2 u_{t-3})] \\
\gamma_1 &= \mathrm{E}\left[\left(\theta_1 u_{t-1}^2 + \theta_1\theta_2 u_{t-2}^2\right)\right] \\
\gamma_1 &= \theta_1\sigma^2 + \theta_1\theta_2\sigma^2 \\
\gamma_1 &= (\theta_1 + \theta_1\theta_2)\sigma^2
\end{aligned}$$

Solution (Cont'd)

$$\gamma_2 = \mathsf{E}[y_t - \mathsf{E}(y_t)][y_{t-2} - \mathsf{E}(y_{t-2})]$$

$$\gamma_2 = \mathrm{E}[y_t][y_{t-2}]$$

$$\gamma_2 = \mathrm{E}[(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})(u_{t-2} + \theta_1 u_{t-3} + \theta_2 u_{t-4})]$$

$$\gamma_2 = \mathrm{E}\big[\big(\theta_2 u_{t-2}^2\big)\big]$$

$$\gamma_2 = \theta_2 \sigma^2$$

Solution (Cont'd)

$$\gamma_3 = \mathrm{E}[y_t - \mathsf{E}(y_t)][y_{t-3} - \mathsf{E}(y_{t-3})]$$

$$\gamma_3 = \mathrm{E}[y_t][y_{t-3}]$$

$$\gamma_3 = \mathrm{E}[(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})(u_{t-3} + \theta_1 u_{t-4} + \theta_2 u_{t-5})]$$

$$\gamma_3 = 0$$

So $\gamma_s = 0$ for $s > 2$.

Solution (Cont'd)

We have the autocovariances, now calculate the autocorrelations:

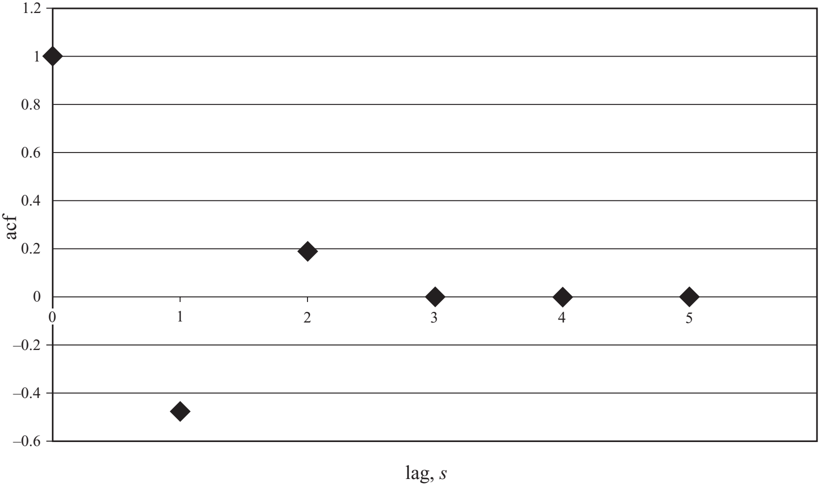
$$\begin{aligned}
\tau_0 &= \frac{\gamma_0}{\gamma_0} = 1 \\
\tau_1 &= \frac{\gamma_1}{\gamma_0} = \frac{(\theta_1 + \theta_1\theta_2)\sigma^2}{\left(1 + \theta_1^2 + \theta_2^2\right)\sigma^2} = \frac{(\theta_1 + \theta_1\theta_2)}{\left(1 + \theta_1^2 + \theta_2^2\right)} \\
\tau_2 &= \frac{\gamma_2}{\gamma_0} = \frac{(\theta_2)\sigma^2}{\left(1 + \theta_1^2 + \theta_2^2\right)\sigma^2} = \frac{\theta_2}{\left(1 + \theta_1^2 + \theta_2^2\right)} \\
\tau_3 &= \frac{\gamma_3}{\gamma_0} = 0 \\
\tau_s &= \frac{\gamma_s}{\gamma_0} = 0 \ \forall \ s > 2
\end{aligned}$$

Solution (Cont'd)

iii. For $\theta_1 = -0.5$ and $\theta_2 = 0.25$, substituting these into the formulae above gives $\tau_1 = -0.476$, $\tau_2 = 0.190$.

ACF Plot

Thus the acf plot will appear as follows:



Autoregressive Processes

- An autoregressive model of order p, an AR(p) can be expressed as

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + u_t$$

- Or using the lag operator notation:

$$Ly_t = y_{t-1} \qquad L^i y_t = y_{t-i}$$

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + u_t$$

- or

$$y_t = \mu + \sum_{i=1}^{p} \phi_i L^i y_t + u_t$$

- or $\phi(L) y_t = \mu + u_t$ where $\phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)$.

The Stationarity Condition for an AR Model

- The condition for stationarity of a general AR(p) model is that the roots of $1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0$ all lie outside the unit circle.

- A stationary AR(p) model is required for it to have an MA(∞) representation.

- Example 1: Is $y_t = y_{t-1} + u_t$ stationary?

  The characteristic root is 1, so it is a unit root process (so non-stationary)

- Example 2: Is $y_t = 3y_{t-1} - 2.75y_{t-2} + 0.75y_{t-3} + u_t$ stationary?

  The characteristic roots are 1, 2/3, and 2. Since only one of these lies outside the unit circle, the process is non-stationary.

The Moments of an Autoregressive Process

- The moments of an autoregressive process are as follows. The mean is given by

$$E(y_t) = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$$

- The autocovariances and autocorrelation functions can be obtained by solving what are known as the Yule-Walker equations:

$$\begin{aligned}
\tau_1 &= \phi_1 + \tau_1\phi_2 + \cdots + \tau_{p-1}\phi_p \\
\tau_2 &= \tau_1\phi_1 + \phi_2 + \cdots + \tau_{p-2}\phi_p \\
&\vdots \\
\tau_p &= \tau_{p-1}\phi_1 + \tau_{p-2}\phi_2 + \cdots + \phi_p
\end{aligned}$$

- If the AR model is stationary, the autocorrelation function will decay exponentially to zero.

Sample AR Problem

- Consider the following simple AR(1) model

$$y_t = \mu + \phi_1 y_{t-1} + u_t$$

1. Calculate the (unconditional) mean of y_t.

   For the remainder of the question, set $\mu = 0$ for simplicity.

2. Calculate the (unconditional) variance of y_t.
3. Derive the autocorrelation function for y_t.

Solution

1. Unconditional mean:

$$\begin{aligned} E(y_t) &= E(\mu + \phi_1 y_{t-1}) \\ E(y_t) &= \mu + \phi_1 E(y_{t-1}) \end{aligned}$$

But also

$$\begin{aligned} E(y_t) &= \mu + \phi_1(\mu + \phi_1 E(y_{t-2})) \\ &= \mu + \phi_1 \mu + \phi_1^2 E(y_{t-2}) \\ &= \mu + \phi_1 \mu + \phi_1^2(\mu + \phi_1 E(y_{t-3})) \end{aligned}$$

Solution (Cont'd)

So

$$\begin{aligned} \mathrm{E}(y_t) &= \mu + \phi_1(\mu + \phi_1 \mathrm{E}(y_{t-2})) \\ &= \mu + \phi_1\mu + \phi_1^2 \mathrm{E}(y_{t-2}) \\ &= \mu + \phi_1\mu + \phi_1^2(\mu + \phi_1 \mathrm{E}(y_{t-3})) \\ \mathrm{E}(y_t) &= \mu + \phi_1\mu + \phi_1^2\mu + \phi_1^3 \mathrm{E}(y_{t-3}) \end{aligned}$$

An infinite number of such substitutions would give

$$\mathrm{E}(y_t) = \mu\left(1 + \phi_1 + \phi_1^2 + \cdots\right) + \phi_1^{\infty} y_0$$

So long as the model is stationary, i.e. $|\phi_1| < 1$, then $\phi_1^{\infty} = 0$.

Solution (Cont'd)

So

$$\mathrm{E}(y_t) = \mu\left(1 + \phi_1 + \phi_1^2 + \cdots\right) = \frac{\mu}{1 - \phi_1}$$

ii. Calculating the variance of y_t: $y_t = \phi_1 y_{t-1} + u_t$

From Wold's decomposition theorem:

$$\begin{aligned} y_t(1 - \phi_1 L) &= u_t \\ y_t &= (1 - \phi_1 L)^{-1} u_t \\ y_t &= \left(1 + \phi_1 L + \phi_1^2 L^2 + \cdots\right) u_t \end{aligned}$$

So long as, $|\phi_1| < 1$, this will converge.

$$\mathrm{var}(y_t) = \mathrm{E}[y_t - \mathrm{E}(y_t)][y_t - \mathrm{E}(y_t)]$$

Solution (Cont'd)

but $\mathrm{E}(y_t) = 0$, since μ is set to zero.

$$
\begin{aligned}
\mathrm{var}(y_t) &= \mathrm{E}[(y_t)(y_t)] \\
&= \mathrm{E}\Big[\big(u_t + \phi_1 u_{t-1} + \phi_1^2 u_{t-2} + \cdots\big)\big(u_t + \phi_1 u_{t-1} \\
&\quad + \phi_1^2 u_{t-2} + \cdots\big)\Big] \\
&= \mathrm{E}\left[u_t^2 + \phi_1^2 u_{t-1}^2 + \phi_1^4 u_{t-2}^2 + \cdots + \textit{cross-products}\right] \\
&= \mathrm{E}\left[u_t^2 + \phi_1^2 u_{t-1}^2 + \phi_1^4 u_{t-2}^2 + \cdots\right] \\
&= \sigma_u^2 + \phi_1^2 \sigma_u^2 + \phi_1^4 \sigma_u^2 + \cdots \\
&= \sigma_u^2 \left(1 + \phi_1^2 + \phi_1^4 + \cdots\right) \\
&= \frac{\sigma_u^2}{(1 - \sigma_u^2)}
\end{aligned}
$$

Solution (Cont'd)

iii. Turning now to calculating the acf, first calculate the autocovariances:

$$\gamma_1 = \mathrm{cov}\,(y_t, y_{t-1}) = \mathrm{E}[y_t - E(y_t)][y_{t-1} - E(y_{t-1})]$$

Since a_0 has been set to zero, $\mathrm{E}(y_t) = 0$ and $\mathrm{E}(y_{t-1}) = 0$, so

$$\gamma_1 = \mathrm{E}[y_t y_{t-1}]$$

Solution (Cont'd)

under the result above that $\mathrm{E}(y_t) = \mathrm{E}(y_{t-1}) = 0$. Thus

$$\begin{aligned}
\gamma_1 &= \mathrm{E}\big[\big(u_t + \phi_1 u_{t-1} + \phi_1^2 u_{t-2} + \cdots\big)\big(u_{t-1} + \phi_1 u_{t-2} \\
&\quad + \phi_1^2 u_{t-3} + \cdots\big)\big] \\
\gamma_1 &= \mathrm{E}\big[\phi_1 u_{t-1}^2 + \phi_1^3 u_{t-2}^2 + \cdots + cross - products\big] \\
\gamma_1 &= \phi_1\sigma^2 + \phi_1^3\sigma^2 + \phi_1^5\sigma^2 + \cdots \\
\gamma_1 &= \frac{\phi_1\sigma^2}{(1-\phi_1^2)}
\end{aligned}$$

For the second autocorrelation coefficient,

$$\gamma_2 = \mathrm{cov}(y_t, y_{t-2}) = \mathrm{E}[y_t - \mathrm{E}(y_t)][y_{t-2} - \mathrm{E}(y_{t-2})]$$

Solution (Cont'd)

Using the same rules as applied above for the lag 1 covariance

$$
\begin{aligned}
\gamma_2 &= \mathrm{E}[y_t y_{t-2}] \\
\gamma_2 &= \mathrm{E}\Big[\big(u_t + \phi_1 u_{t-1} + \phi_1^2 u_{t-2} + \cdots\big)\big(u_{t-2} + \phi_1 u_{t-3} \\
&\qquad + \phi_1^2 u_{t-4} + \cdots\big)\Big] \\
\gamma_2 &= \mathrm{E}\big[\phi_1^2 u_{t-2}^2 + \phi_1^4 u_{t-3}^2 + \cdots + \textit{cross-products}\big] \\
\gamma_2 &= \phi_1^2\sigma^2 + \phi_1^4\sigma^2 + \cdots \\
\gamma_2 &= \phi_1^2\sigma^2\big(1 + \phi_1^2 + \phi_1^4 + \cdots\big) \\
\gamma_2 &= \frac{\phi_1^2\sigma^2}{(1-\phi_1^2)}
\end{aligned}
$$

Solution (Cont'd)

If these steps were repeated for γ_3, the following expression would be obtained

$$\gamma_3 = \frac{\phi_1^3 \sigma^2}{\left(1 - \phi_1^2\right)}$$

and for any lag s, the autocovariance would be given by

$$\gamma_s = \frac{\phi_1^s \sigma^2}{\left(1 - \phi_1^2\right)}$$

Solution (Cont'd)

The acf can now be obtained by dividing the covariances by the variance:

$$
\begin{aligned}
\tau_0 &= \frac{\gamma_0}{\gamma_0} = 1 \\
\tau_1 &= \frac{\gamma_1}{\gamma_0} = \frac{\left(\dfrac{\phi_1 \sigma^2}{(1-\phi_1^2)}\right)}{\left(\dfrac{\sigma^2}{(1-\phi_1^2)}\right)} = \phi_1 \\
\tau_2 &= \frac{\gamma_2}{\gamma_0} = \frac{\left(\dfrac{\phi_1^2 \sigma^2}{(1-\phi_1^2)}\right)}{\left(\dfrac{\sigma^2}{(1-\phi_1^2)}\right)} = \phi_1^2 \\
\tau_3 &= \phi_1^3
\end{aligned}
$$

The Partial Autocorrelation Function (denoted τ_{kk})

- Measures the correlation between an observation k periods ago and the current observation, after controlling for observations at intermediate lags (i.e. all lags $<k$).
- So τ_{kk} measures the correlation between y_t and y_{t-k} after removing the effects of $y_{t-k+1}, y_{t-k+2}, \ldots, y_{t-1}$
- At lag 1, the acf $=$ pacf always
- At lag 2,

$$\tau_{22} = \left(\tau_2 - \tau_1^2\right) / \left(1 - \tau_1^2\right)$$

- For lags 3+, the formulae are more complex.

The Partial Autocorrelation Function (denoted τ_{kk}) (Cont'd)

- The pacf is useful for telling the difference between an AR process and an ARMA process.
- In the case of an AR(p), there are direct connections between y_t and y_{t-s} only for $s \leq p$.
- So for an AR(p), the theoretical pacf will be zero after lag p.
- In the case of an MA(q), this can be written as an AR(∞), so there are direct connections between y_t and all its previous values.
- For an MA(q), the theoretical pacf will be geometrically declining.

ARMA Processes

- By combining the AR(p) and MA(q) models, we can obtain an ARMA(p,q) model:

$$\phi(L)y_t = \mu + \theta(L)u_t$$

where

$$\begin{aligned} \phi(L) &= 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p \quad \text{and} \\ \theta(L) &= 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q \end{aligned}$$

or

$$\begin{aligned} y_t &= \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 u_{t-1} \\ &\quad + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} + u_t \end{aligned}$$

with

$$\mathrm{E}(u_t) = 0; E\left(u_t^2\right) = \sigma^2; E(u_t u_s) = 0, t \neq s$$

Summary of the Behaviour of the acf for AR and MA Processes

An autoregressive process has

- a geometrically decaying acf
- number of spikes of pacf = AR order

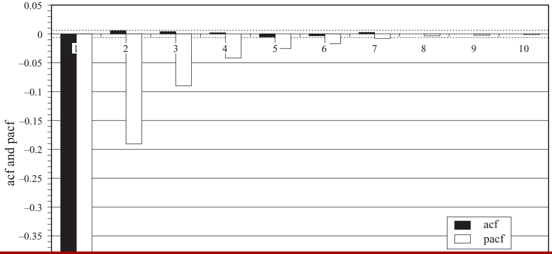
A moving average process has

- Number of spikes of acf = MA order
- a geometrically decaying pacf

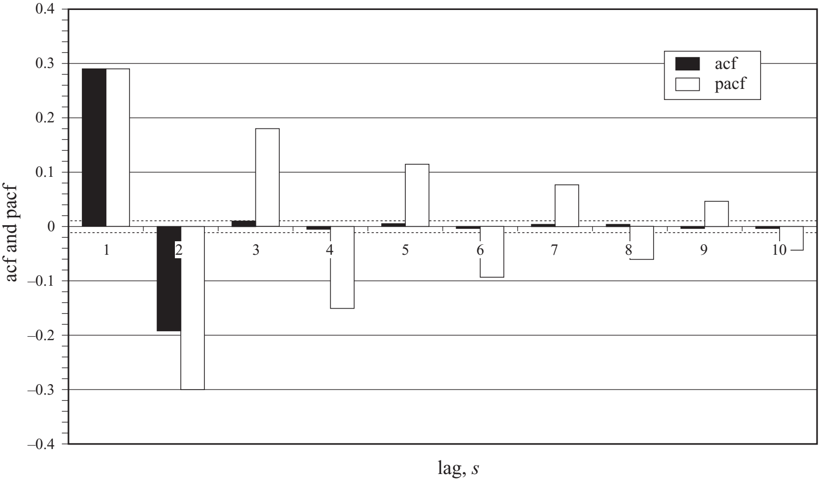
Some sample acf and pacf plots for standard processes

- The acf and pacf are not produced analytically from the relevant formulae for a model of that type, but rather are estimated using 100,000 simulated observations with disturbances drawn from a normal distribution.

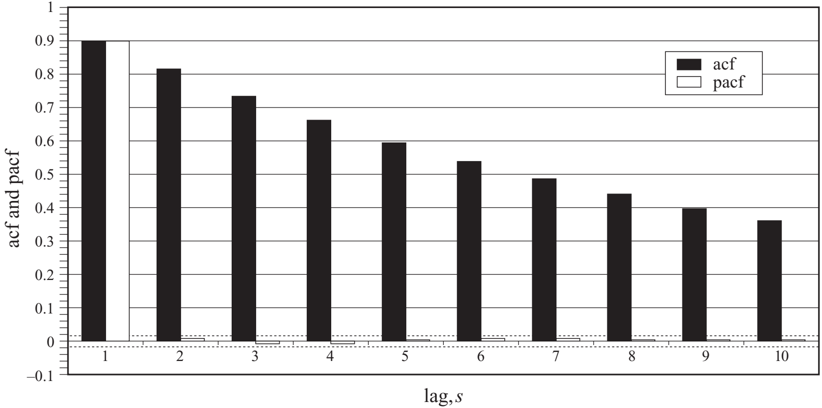
Figure: Sample autocorrelation and partial autocorrelation functions for an MA(1) model: $y_t = -0.5u_{t-1} + u_t$



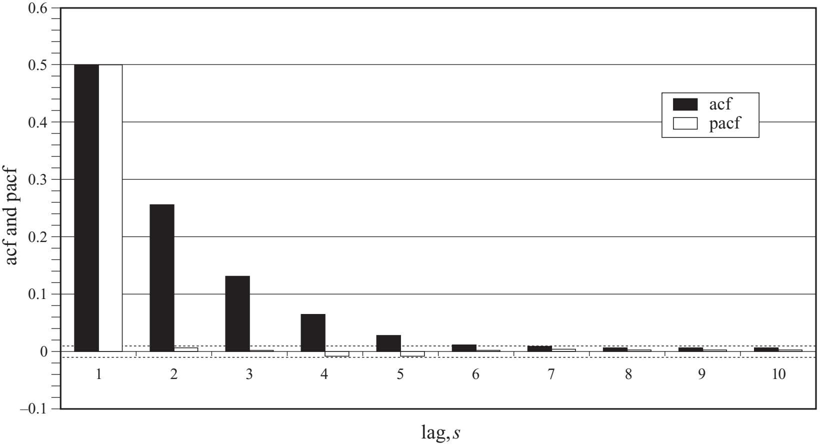
ACF and PACF for an MA(2) Model: $y_t = 0.5u_{t-1} - 0.25u_{t-2} + u_t$



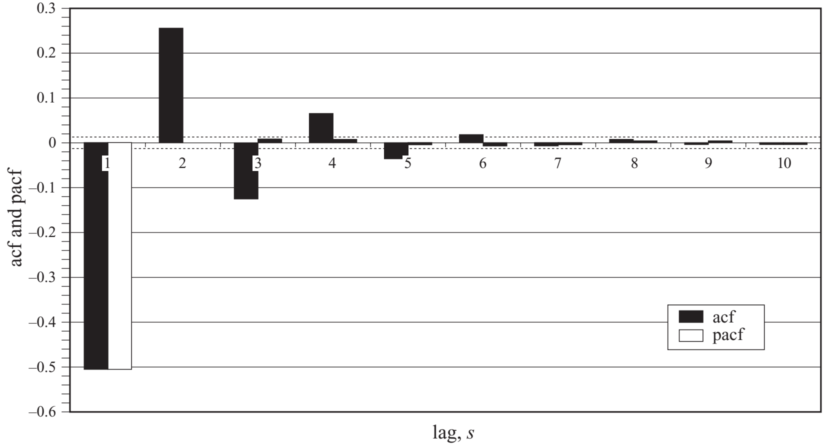
ACF and PACF for a slowly decaying AR(1) Model: $y_t = 0.9y_{t-1} + u_t$



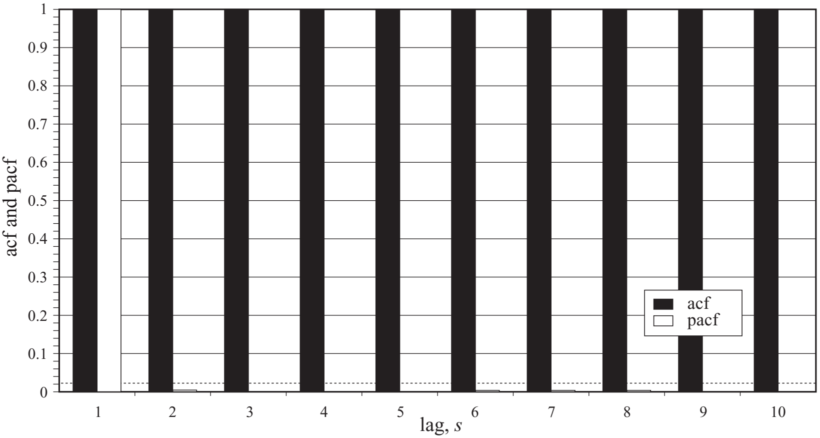
ACF and PACF for a more rapidly decaying AR(1) Model: $y_t = 0.5y_{t-1} + u_t$



ACF and PACF for a more rapidly decaying AR(1) Model with Negative Coefficient: $y_t = -0.5y_{t-1} + u_t$



ACF and PACF for a Non-stationary Model (i.e. a unit coefficient): $y_t = y_{t-1} + u_t$



ACF and PACF for an ARMA(1,1): $y_t = 0.5y_{t-1} + 0.5u_{t-1} + u_t$