1 Forewords

If we imagine that any observed time observations as the part of an infinite sequence of random variables we will be capable of constructing a stochastic process $\{Y_t\}$, i.e., a valid probability model of time series. This infinite set of time observations is called "Ensemble" of time series. Every member of this ensemble is a possible realization of that stochastic process $\{Y_t\}$. This realization may be sometimes discrete, i.e.,$t = 1, 2, \cdots, N$ or continuous, i.e. $X(t), 0 < t < T$. So one can imagine stationary stochastic process in the context of time series too. Likewise a stationary time series can be split in two categories–strict stationary time series and cov./mean stationary time series. Stationarity remains an important character in the study of time series. Briefly speaking, for a stationary time series, probability structure is same over a specified time interval. The term "univariate time series" refers to a time series that consists of single (scalar) observations recorded sequentially over equal time increments, for example, LakeHuron data or atmospheric concentration of CO2 (Look at R).

The most important part of time series analysis is to forecast or predict on future observations collecting information from the available data. To address this issue of forecasting we need to build a mathematical relationship among the observations on different time points. This mathematical function is called time series model.

Definition 1. ***Autocovariance:*** *Think about a time series ensemble $\{X_t\}$. Autocoviance is a covariance between two observations of same ensemble. Why "Auto"? as if we are calculating the correlations of two values of same variables. Autocovariance of order k,*

$$\gamma_k = cov(X(t), X(t+k)) = E(X(t) - \mu)(X(t+k) - \mu)$$

, where μ being the mean of the time ensemble $\{X(t)\}$. In sample structure $\hat{\gamma_k} = \frac{1}{n}\sum_{t=1}^{n-k} X_t X_{t+k} - \overline{X}^2$. Naturally, $\gamma_0 = var(X(t))$.

Definition 2. ***Autocorrelation:*** *It is a correlation between two observations of same ensemble.*

$$\rho_k = corr(\rho(t), \rho(t+k)) = \gamma_k/\gamma_0$$

.It is also called serial correlation. A function of order of autocorrelations k, $\gamma(k)$ is autocorrelation function.k is said order/lag of autocorrelation.

1.1 Properties of ρ_k

- The ACF is a way to measure the linear relationship between an observation at time t and the observations at previous times.
- For a stationary time series, $\rho_k = \rho_{-k}$, means autocorrelation/autocovarince of $X(t), X(t+k)$ and $X(t), X(t-k)$ are same. Moving k time periods apart forward or backward do not change the inter association between the individuals.
- $|\rho_k| < 1 \, \forall k$ as it is a correlations, so lying between ± 1.
- A time series model always has a specific autocorrelation function but the converse is not true.
- The autocorrelation function can be used for the following two purposes
  1. To detect non-randomness in data.
  2. To identify an appropriate time series model if the data are not random.

1.2 Checking of nonrandomness/randomness with Autocorrelation Function

Randomness is one of the key assumptions in determining if a univariate statistical process is in control. Most standard statistical tests depend on randomness. The validity of the test conclusions is directly linked to the validity of the randomness assumption. If the assumptions of constant location and scale, randomness, and fixed distribution are reasonable, then the univariate time series process can be modeled as:

$$Y(t) = A_0 + \epsilon_i$$

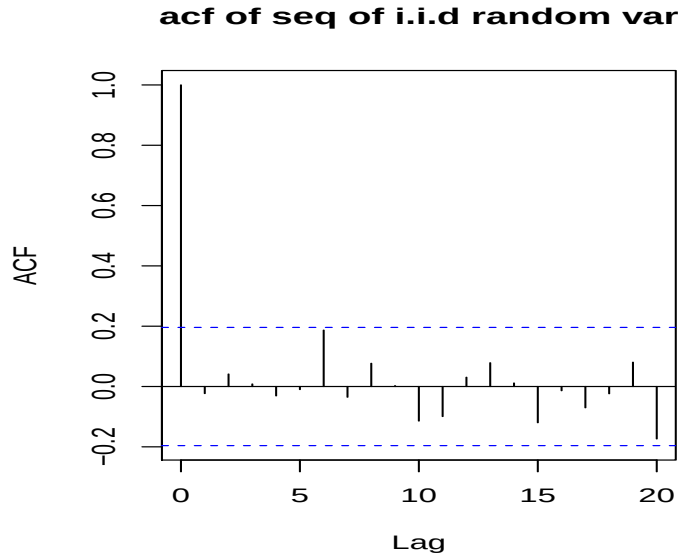
where ϵ_i is an error term. If the randomness assumption is not valid, then a different model needs to be used.

When the autocorrelation is used to detect non-randomness, it is usually only the first (lag 1) autocorrelation that is of interest.This randomness may be ascertained by computing autocorrelations for data values at varying time lags but mostly by acf of lag 1. If random, such autocorrelations

should be near **zero** for any and all time-lag separations. If non-random, then one or more of the autocorrelations will be significantly non-zero.

Note that uncorrelated does not necessarily mean random. Data that has significant autocorrelation is not random. However, data that does not show significant autocorrelation can still exhibit non-randomness in other ways. Therefore, some time series data require a more rigorous checking of randomness. In these cases, a battery of tests, which might include checking for autocorrelation, are applied since data can be non-random in many different and often subtle ways.

Let us have two ACF plots below (in R command is acf, acf(y) where y is your series). Remember for a stationary time series model ACF will be

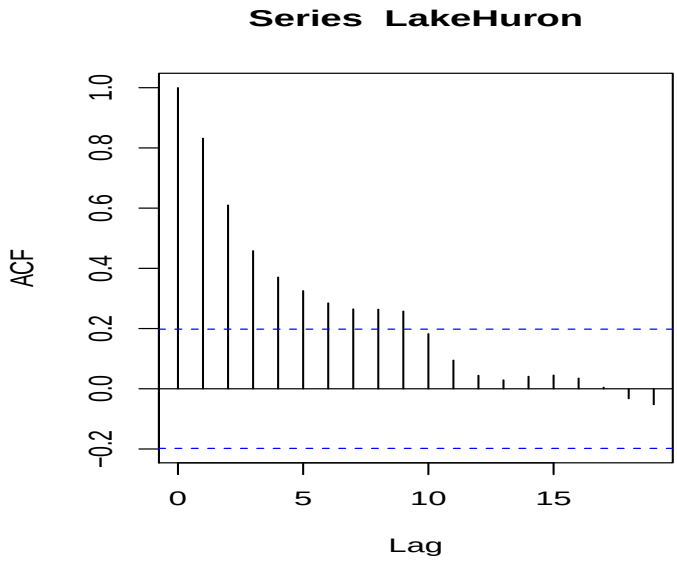


tailing off (getting smaller **not** zero) as the value of lag will be increased.

2 Basic Models

2.1 Pure Random Process(White Noise Process)

A white noise process is one with a mean zero and no correlation between its values at different times. Any sequence of i.i.d. random variables $\{Z_t\}$ can be taken as an example of white noise process (not necessarily normal, the observations should be i.i.d.–that's it!!). In white noise process mean is taken



as zero. For a white noise process $\gamma_k = cov(Z_t, Z_{t+k}) = \begin{cases} \sigma_z^2 \text{ if } k = 0 \\ 0 \text{ if } k \neq 0 \end{cases}$.
Naturally, autocorrelations of every order are zero.The process is mean and strict stationary. (why?)

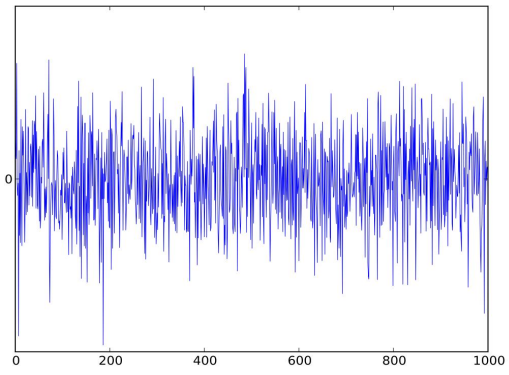


Figure 1: White noise process

2.2 Random Walk

Random Walk model is the time series model where t th time observation depends on $(t-1)$ th time observation and on a random observation. Mathematically, $X(t) = X(t-1) + Z(t)$ where Z_t coming from a pure random process, $E(Z(t)) = 0$ and $V(Z(t)) = \sigma^2$ (see it could be normal/nonnormal). For the sake of computational simplicity, we consider $X(0) = 0$. So, $X(1) = Z(1)$, $X(2) = z(1) + z(2)$, in general, $X(t) = \sum_k = 1^t Z(t)$...Thus $E(X(t)) = 0$, $V(X(t)) = t\sigma^2$.
$cov(X(t), X(t+h)) = t\sigma^2$ (why?).Covariance depends on t. So random walk is not a stationary process.

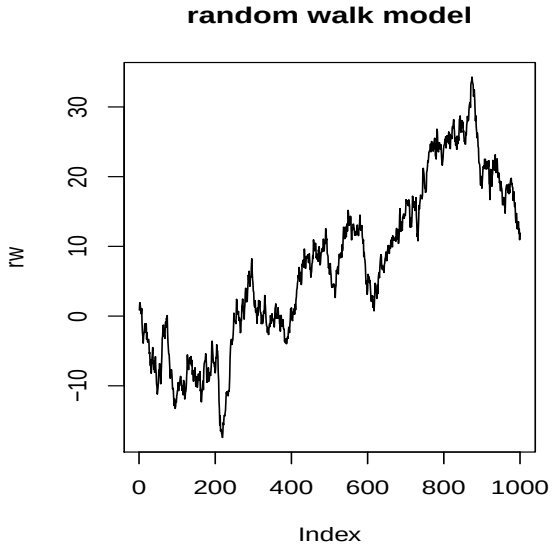


Figure 2: Random Walk time series

2.3 Questions to be solved by R

1. . Generate a set of 200 observations from N(0,4)
2. Let the observations are z_t. Plot the series joining by line.
3. Plot the ACF function of z_t. Comment
4. Next consider $X_t = X_{t-1} + Z_t$. Generate 200 such X_t.Plot X_t.

5. Plot ACF of X_t and hence comment.

Remark 1. *It also follows for a stationary process that the nature of the joint probability distribution $p(Z_t, Z_{t+k})$ of values separated by k intervals of time can be inferred by plotting a scatter diagram using pairs of values (z_t, z_{t+k}) of the time series, separated by a constant interval or lag k.*

2.4 Question to be solved

The following are temperature measurements z_t made every minute on a chemical reactor:

200	202	208	204	204	207	204	202	199	201	198	200	207
202	203	205	207	211	204	206	203	203	201	198	200	206

1. Plot z_{t+1} versus z_t
2. Plot z_{t+2} versus z_t.
3. After inspecting the graphs, do you think that series is autocorrelated?

3 General linear process

Any stochastic process can always be written in terms of white noises $\{z_t\}$ and sometimes a part of the series itself. Time series being a derivative of stochastic process can also be expressed in terms of (white noise) z_t. Since we need to have series in infinite nature we can think a time series sequence $\{X_t\}$ in terms of linear combination of infinite white noises.

$$X_t = Z_t + \psi_1 Z_{t-1} + \psi_2 Z_{t-2} + \cdots = \sum_{j=1}^{\infty} \psi_i z_i \tag{1}$$

where ψ_i's are the coefficients and we assume X_t is taken as the deviated form the mean $E(X_t) = \mu$. When X_t is claimed to be stationary it is necessary for the coefficients ψ_j to be absolutely summable,i.e., $\sum_{j=0}^{\infty} |\psi_j| < \infty$.

In contrast to (1) another representation of $\{X_t\}$can be thought in terms of past observations of the same series,i,e, on $X_{t-1}, X_{t-2}, \cdots$ plus an added white noise a_t.,i,e,

$$X_t = \pi_1 X_{t-1} + \pi_2 X_{t-2} + \cdots + z_t = \sum_{j=1}^{\infty} \pi_j X_{t-j} + z_t \tag{2}$$

Equation (2) may be thought of as one where the current deviation X_t from the level μ is regressed on past deviations $X_{t-1}, X_{t-2}, \cdots$ of the process.

3.1 Relationship between the ψ weights and π weights

First let us define **backward shift** operator B such that $Bz_t = z_{t-1}$ and hence $B^j z_t = z_{t-j}$. So in terms of backward shift operator equation (1) is $X_t = (1 + \psi_1 B + \psi_2 B^2 + \cdots)Z_t = \psi(B)Z_t$ where $\psi_(B)$ is an infinite series on B,i.e., $\psi_B = 1 + \psi_1 B + \psi_2 B^2 + \cdots$, ($\psi_0 = 1$, conventionally).

Similarly equation (2) can be written as

$$X_t - \pi_1 X_{t-1} - \pi_2 X_{t-2} - \cdots = Z_t$$

or

$$\pi(B)X_t = Z_t$$

where $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \cdots$. Replacing X_t we get,

$$\begin{aligned} \psi(B)\pi(B)X_t = \psi(B)Z_t = X_t \\ \psi(B)\pi(B) = 1 \\ \psi(B) = [\pi(B)]^{-1} \end{aligned}$$

- This relationship is helpful in determining π weights knowing ψ weights and vice-versa.
- These two representations of time series are identical to each other.

Stationarity and **invertibility** are two important concepts on which the convergence(stability) of time series can be built. $\pi(B)$ and $\psi(B)$ play a great role to conceptualize these two concepts.

Question: Write the following models in B notation, write first three π and ψ weights

$$X_t - .5X_{t-1} = z_t$$

$$X_t = z_t - 1.3z_{t-1} + .4z_{t-2}$$

$$X_t - X_{t-1} = z_t - .5z_{t-1}$$

$$X_t - X_{t-1} = z_t - 1.3z_{t-1} + .3z_{t-2}$$

.

3.2 Conditions for stationarity and invertibility

3.2.1 Stationarity

The linear process is stationary if the polynomial of coefficients attached to white noise has roots on or within unit circle,i.e.,$\psi(B) = 0$, the roots of this equation lie on or within unit circle which means $\sum \psi_j < \infty$.

3.2.2 Invertibility

On the contrary, the process is invertible if $\pi(B) = 0$ has roots on or within unit circle,means $\sum_{j=0}^{\infty} |\pi_j| < \infty$.
Example Let us consider two model:
Model A: $X_t = Z_t + \theta Z_{t-1}$
Model B: $X_t = Z_t + Z_{t-1}/\theta$.
But in Model A,

$$\begin{aligned} Z_t &= X_t - \theta Z_{t-1} = X_t - \theta(X_{t-1} - \theta Z_{t-2}) \\ &= X_t - \theta X_{t-1} + \theta^2 Z_{t-2} \\ &= X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \theta^3 X_{t-3} + \cdots \end{aligned}$$

Similarly for Model B
$Z_t = X_t - \frac{1}{\theta} X_{t-1} + \frac{1}{\theta^2} X_{t-2} - \frac{1}{\theta^3} X_{t-3} + \cdots$. But Which one will be convergent series and for what condition? **Think**. Well take $\theta = 5$. What would be your conclusion?

4 Few Particular Models: Moving Average Process and Autoregressive Process

An autoregressive model is when a value from a time series is regressed on previous values from that same time series. for example, y_t on y_{t-1}.

An autoregressive process of order p, AR(p) is written as

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + Z_t,$$

where $Z_t \sim WN(0, \sigma^2)$ and Z_t is uncorrelated with X_t with X_s for each $s < t$. By backward shift operator AR(p) can be written as

$$\phi(B) X_t = Z_t$$

where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$.

4.1 Autoregressive Process(1)

$$X_t = \phi X_{t-1} + Z_t$$

. In backward shift operator language, $(1-\phi B)X_t = Z_t$ which can be further written as $X_t = \phi_k X_{t-k} + \sum_{j=0}^{k-1} \phi^j Z_{t-j}$.

4.1.1 Invertibility

Remember from earlier discussion invertibility depends on the convergence of polynomial attached to X_t, i.e., $\Pi(B)$. If $\sum \pi_i < \infty$ then the process is invertible. In AR(1), $\Pi(B) = 1 - \phi B$. This is always finite hence AR(1) is always invertible.

4.1.2 Stationarity

Remember from earlier discussion stationarity depends on the convergence of polynomial attached to Z_t (white noise), i.e., $\psi(B)$ (or roots of the equation $\psi(B) = 0$ lies within the unit circle). For AR(1), $\psi(B) = (1 - \phi(B))^{-1}$. The right hand term is an infinite series which is convergent if $|\phi| < 1$. You can address the problem by considering roots of $\psi(B)$ here. Do the answers found in both ways match? (Do it yourself)

Note For a simulated series $X_t = 1.02X_{t-1} + Z_t$ (see the series is not stationary) we can see the values of the time series quickly become large in magnitude, even for ϕ just slightly above 1. Such process is called explosive(Try to draw the time series curve of the above equation, you will see how the curve progresses.

So AR(1) process is stationary when $|\phi| < 1$.

4.1.3 Autocorrelation

Remember $E(X_t) = 0$. If it is $E(X_t) = \mu$, we start with $X_t^{'}$ such that $E(X_t - \mu) = 0$

$$\begin{aligned} Cov(X_t, X_{t-k}) &= Cov(\sum_{j=0}^{\infty} \phi^j \, z_{t-j}, \sum_{j=0}^{\infty} \phi^j \, z_{t-(j+k)}) \\ &= (z_t + \phi z_{t-1} + \phi^2 z_{t-2} + \cdots + \phi^k z_{t-k} + \phi^{k+1} z_{t-(k+1)} + \cdots). \\ &\quad (z_{t-k} + \phi z_{t-(k+1)} + \phi^2 z_{t-(k+2)} + \cdots) \\ &= \phi^k + \phi\phi^{k+1} + \phi^2\phi^{k+2} + \cdots \\ &= \frac{\phi^k}{1 - \phi^2} \end{aligned}$$

Choose $k = 0$, $Var(X_t) = \frac{1}{1-\phi^2}$ and $\rho_k = \phi^k$.

4.1.4 Autocorrelation function

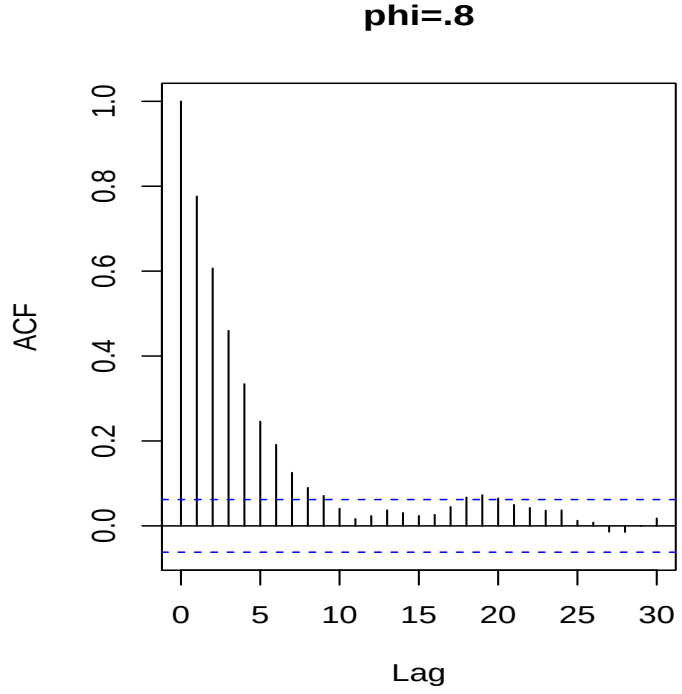
Autocorrelation function of AR(1) satisfies the difference equation $\rho_k = \phi_1 \rho_{k-1}$, $k > 0$ (why?). The autocorrelation function decays exponentially when ϕ is positive but oscillates in sign when ϕ is negative. Below are two graphs of ACF of simulated AR(1) model where a) $X_t = .8X_t + z_t$ and b) $X_t = -.8X_{t-1} + z_t$. The couple of blue dotted lines indicate the significance of autocorrelations which is $(exp(2{*}1.96/\sqrt{N-3}{-}1), exp(2{*}1.96/\sqrt{N-3}{+}1))$. The lines give the values beyond which the autocorrelations are (statistically) significantly different from zero.

4.2 Moving average of order 1:MA(1)

MA(1) model: $X_t = z_t + \theta \, z_{t-1}$. Many textbooks and software programs define the model with negative signs before the θ terms. This doesn't change the general theoretical properties of the model, although it does flip the algebraic signs of estimated coefficient values and (unsquared) θ terms in formulas for ACFs and variances.

4.2.1 Stationarity and Invertibility

Remember what are the conditions of the stationarity and invertibility? MA(1) is always stationary but invertible when $|\theta| < 1$. Why?

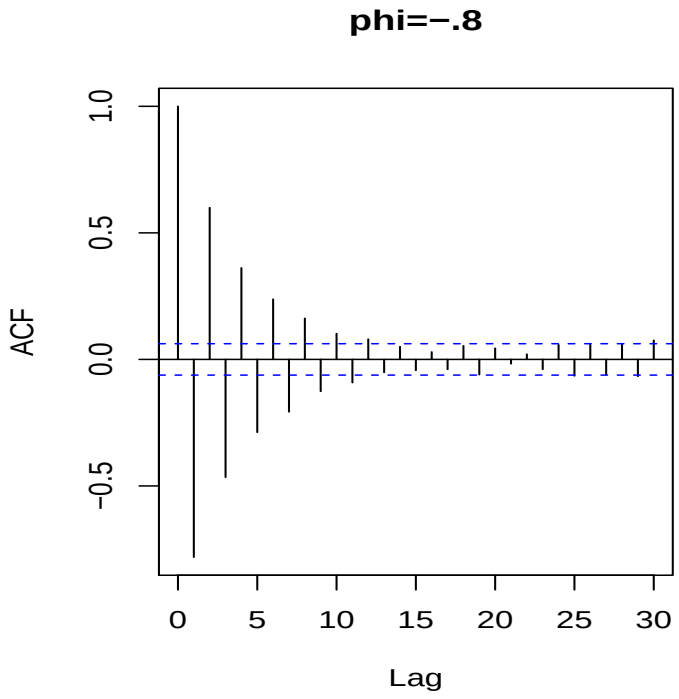


4.2.2 Autocorrelation

1. Variance= $V(X_t) = \sigma^2(1+\theta^2)$ (how?)
2. $\rho_1 = \frac{\theta}{1+\theta^2}$(how?)
3. $\rho_k = 0$ when $k > 1$.

Remark 2. *That the only nonzero value in the theoretical ACF is for lag 1. All other autocorrelations are 0. Thus a sample ACF with a significant autocorrelation only at lag 1 is an indicator of a possible MA(1) model.*

Example Suppose $X_t = z_t + .7z_{t-1}$ where $z_t \sim N(0,1)$.Theoretical ACF is $\rho_1 = \frac{.7}{1+.7^2} = .4698$. All other $\rho_k > 0, k > 1$. The same model one can generate by simulation in R. A plot of ACF is as follows.
Again look at the Moving average model $X_t = Z_t + \frac{Z_{t-1}}{.7}$. This model has exactly same autocorrelation function as $X_t = z_t + .7z_{t-1}$. But this model is not invertible as $1/.7 > 1$ whereas $X_t = z_t + .7z_{t-1}$ is invertible. **The imposition of the invertibility condition ensures that there is a**



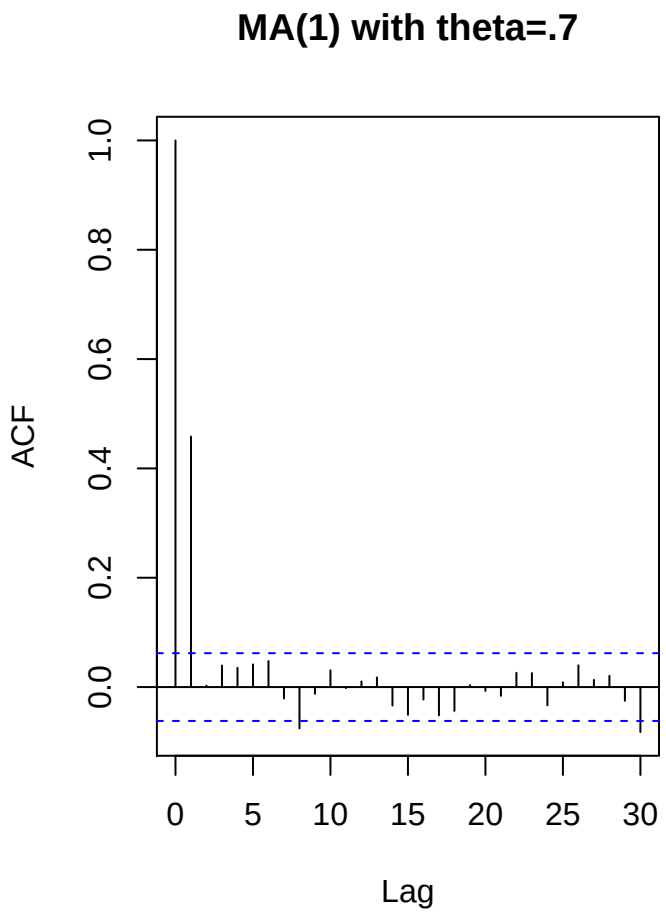
unique MA process for a given autocorrelation function.
Homework

1. Show that AR(1)$\equiv$ $MA(\infty)$.
2. Show that MA(1) $\equiv$ $AR(\infty)$.
3. State whether the following model is stationary/invertible. $X_t - .5X_{t-1} = a_t - .5a_{t-1}$

In R you can generate sample points directly from Moving average/autoregressive process by the following code. `ma1.sim<-arima.sim(list(ma = c(0.3),ar=c(.4)),n = 100, sd=1)`. Remember this is a general code. If you have to generate only MA then erase the word ar and vice versa.

4.3 Moving average of order 2

This process is written as $X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$ where Z_t being white noise of zero mean and sd σ^2. In terms of backward shift operator MA(2)



can be written as

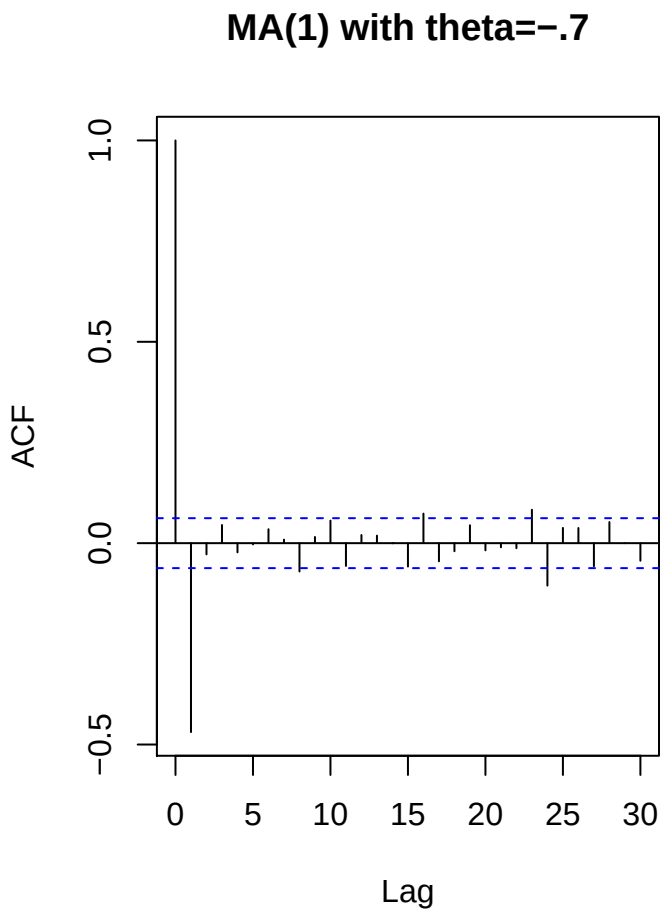
$$X_t = (1 + \theta_1 B + \theta_2 B^2) Z_t$$

.

4.3.1 Autocorrelations

$$\gamma(k) = cov(X_t, X_{t+k}) = \begin{cases} (1 + \theta_1^2 + \theta_2^2)\sigma^2 \text{ for } k = 0 \\ (\theta_1 + \theta_1\theta_2)\sigma^2 \text{ for } k = 1 \\ \theta_2\sigma^2 \text{ for } k = 2 \\ 0 \text{ for } k > 2 \end{cases}$$

Now find out the the autocorrelation function of MA(2).



Problem

Take two series with z_t white noise with zero mean and variance 1.

A:$X_t = Z_t + .5Z_{t-1} + .5Z_{t-2}$

B:$X_t = Z_t + 5Z_{t-1} + 5Z_{t-2}$. Simulate 100 observations from each of the model and plot them (use `plot.ts` function). Which one you think invertible? (Note stationarity does not hamper but span of series is inflated).

MA(2) is always stationary (why?) but invertible when $\theta_1 + \theta_2 < 1$,$\theta_1 - \theta_2 < 1$,$|\theta_2| < 1$ (how? we will see this later)

Remember For MA(2), the only nonzero values in the theoretical ACF are for lags 1 and 2. Autocorrelations **for higher lags are** 0. So, a sample ACF with significant autocorrelations at lags 1 and 2, but non-significant autocorrelations for higher lags indicates a possible MA(2) model.

4.4 Yule Walker Equations:For AR process

An important recurrence relation for the autocorrelation function of a stationary autoregressive process is found by multiplying throughout in $X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + z_t$ (General Autoregressive process of order p) by X_{t-k} for $k \geq 0$, to obtain

$$X_{t-k}X_t = \phi_1 X_{t-k}X_{t-1} + \phi_2 X_{t-k}X_{t-2} + \cdots + \phi_p X_{t-k}X_{t-p} + X_{t-k}z_t.$$

On taking expected values of the above expression we obtain the difference equation

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + \cdots + \phi_p\gamma_{k-p}, k > 0.$$

Note that $E(X_{t-k}z_t)$ vanishes since X_{t-k} involves the term upto z_{t-k} and z_t is uncorrelated with z_{t-k}. On dividing by γ_0 we have difference equation in autocorrelation.

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \cdots + \phi_p\rho_{k-p} \tag{3}$$

If we substitute $k = 1, 2, \cdots, p$ in the above we obtain a set of linear equations in $\phi_1, \phi_2, \cdots, \phi_p$ in terms of $\rho_1, \rho_2, \cdots, \rho_p$ as follows.

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \cdots + \phi_p\rho_{p-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \cdots + \phi_p\rho_{p-2} \\
&\cdots \\
\rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \cdots + \phi_p
\end{aligned}$$

These equations are called Yule-Walker equations. We obtain Yule Walker estimates of the parameters by replacing the theoretical autocorrelations ρ_k by estimated autocorrelatons r_k.

4.4.1 Variance

When $k = 0$ from the equation 4.4 we obtain

$$\gamma_0 = \phi_1\gamma_{-1} + \phi_2\gamma_{-2} + \cdots + \phi_p\gamma_{-p} + \sigma_z^2.$$

Writing as $\gamma_k = \gamma_o\rho_k$,we have the variance

$$\gamma_0 = \frac{\sigma^2}{1 - \phi_1\rho_1 - \phi_2\rho_2 - \cdots - \phi_p\rho_p}.$$

4.5 AR(2)

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$$

where Z_t is white noise with mean 0 and variance σ^2.

4.5.1 Stationarity of AR(2)

See the file **Stationarity AR(2) ACF**.pdf.
Remember for AR(2) when the roots of the difference equation in backward shift operator are real ACF looks as a mixture of damped exponentials. Moreover if it has a positive dominant root ACF will be positive exponential decay, negative dominant root ACF alternates its sign rapidly as its damps out. If the roots are complex, ACF will be damped sine wave.

4.5.2 Autocorrelations

For AR(2) the autocorrelation function satisfies the second order difference equation

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}.$$

Substituting $p = 2$ in Yule Walker equations, the Yule Walker equations are

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$
$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

Solving equations we obtain

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}$$
$$\rho_2 = \phi_2 + \frac{\phi_1^2}{1 - \phi_2}$$

Using ρ_1 and ρ_2 one can find out ρ_3,ρ_4 as well. Find $V(X_t)$.
Question For a AR(2) process $X_t = .75X_{t-1} - .50X_{t-2} + Z_t$,

1. draw the series.
2. Construct the difference equation on Backward shift operator.
3. Can you tell anything about the nature of the roots.
4. Deduce first four theoretical autocorrelations.
5. Plot by ACF command in R. How does the ACF look?

4.6 Partial Autocorrelation

Initially we may not know the order of autoregressive process to fit to an observed time series, means how to determine that X_t is dependent on X_{t-1} only or X_{t-1}, X_{t-2} or $X_{t-1}, X_{t-2}, \cdots, X_{t-p}$? The ACF of all autoregressive process are almost of similar type (either exponential decaying or damped sinusoidal), so deciding about the order of AR process from autocorrelation function can not be possible also. The problem is analogous to deciding on the number of independent variables to be included in a multiple regression.

The autocorrelation for an observation and an observation at a prior time step is comprised of both the direct correlation and indirect correlations. These indirect correlations are a linear function of the correlation of the observation, with observations at intervening time steps. It is these indirect correlations that the partial autocorrelation function seeks to remove. Consider the autocorrelation at lag 2. Observation 1 effects observation 2. Observation 1 affects observation 3 through two channels i.e. directly and indirectly through its effect on observation 2 and observations 2's effect on observation 3.The autocorrelation measures both effects. The partial autocorrelation measures only the direct effect.

Definition: A partial autocorrelation is a summary of the relationship between an observation in a time series with observations at prior time steps with the relationships of intervening observations removed.

Specifically, partial autocorrelations are useful in identifying the order of an autoregressive model. The partial autocorrelation of an AR(p) process is zero at lag $p+1$ and greater.It can be described in terms of p nonzero functions of the autocorrelations. If the sample autocorrelation plot indicates that an AR model may be appropriate, then the sample partial autocorrelation plot is examined to help identify the order.

4.7 Derivation

To find partial autocorrelations we may set out this procedure as follows.

$$\begin{aligned} X_t &= \phi_{11}X_{t-1} + \epsilon_t \\ X_t &= \phi_{21}X_{t-1} + \phi_{22}X_{t-2} + \epsilon_t \\ X_t &= \phi_{31}X_{t-1} + \phi_{32}X_{t-2} + \phi_{33}X_{t-3} + \epsilon_t \\ \cdots & \\ X_t &= \phi_{k1}X_{t-1} + \cdots + \phi_{kk}X_{t-k} + \epsilon_t \end{aligned}$$

The sequence $\phi_{11}, \phi_{22}, \cdots, \phi_{kk}$ are the partial autocorrelations. Multiply the final equation above by X_{t-k}, take expectations and diving by the variance. Do the same operation with $X_{t-1}, X_{t-2}, X_{t-3}, \cdots, X_{t-k}$ successively to get the following a set of k Yule Walker equations.

$$\begin{aligned} \rho_1 &= \phi_{k1} + \phi_{k2}\rho_1 + \cdots + \phi_{kk}\rho_{k-1} \\ \rho_2 &= \phi_{k1}\rho_1 + \phi_{k2} + \cdots + \phi_{kk}\rho_{k-2} \\ \cdots & \\ \rho_k &= \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \cdots + \phi_{kk} \end{aligned}$$

So the solution would be

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \cdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdots \\ \rho_k \end{bmatrix} \tag{4}$$

Solving these equations for $k = 1, 2, 3, \cdots$ successively, we obtain

$$\begin{aligned} \phi_{11} &= \rho_1 \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \end{aligned}$$

4.7.1 Partial autocorrelations of AR(1) and AR(2)

PACF of AR(1) is just $\phi_{11} = \rho_1 = \phi$ as we are regressing X_t on X_{t-1} and using equation (4).

PACF of AR(2): There are two PACF ϕ_{11} and ϕ_{22}. Using equation (4), $\phi_{11} = \rho_1 = \frac{\phi_1}{1-\phi_2}$ and $\phi_{22} = \frac{\rho_2-\rho_1^2}{1-\rho_1^2} = \phi_2$. In R you can use the function pacf() to obtain simulated partial autocorrelation.

```
ar1.sim<-arima.sim(list(ar = c(0.6)),n = 100, sd=2)\\
pacf(ar1.sim)
```

You would have only one significance partial autocorrelation function in this case. Lower the value of ϕ partial autocorrelation would be lowered (Why?). Check with AR(2) model you will have spikes of all partial correlation coefficients within the line of significance **except the the first two autocorrelations**. By partial autocorrelation it is easy to identify the model.

4.7.2 Partial autocorrelation of MA(1) and MA(2)

For MA(1) $X_t = Z_t + \theta Z_{t-1}$, $\rho_1 = \frac{\theta}{1+\theta^2}$ and $\rho_k > 0$ when $k > 1$. Now replace all these in equation (4). $\phi_{11} = \rho_1$ and $\phi_{22} = \dfrac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = -\frac{\rho_1^2}{1-\rho_1^2}$.

What about ϕ_{33}?

Get back to equation (4) and we have $\phi_{33} = \dfrac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \dfrac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & 0 \\ 0 & \rho_1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & 0 \\ \rho_1 & 1 & \rho_1 \\ 0 & \rho_1 & 1 \end{vmatrix}}$.

Similarly one can find out ϕ_{kk}.

Thus for MA(1) partial autocorrelations is infinite in nature rather tails off sinusoidally. Check it in R code as well.
ACF and PACF are two identification tools for AR process and MA process.