AR and MA processes
Autocorrelation func. (ACF)
Partial autocorrelation func. (PACF)

Hamid Raza
Associate Professor in Economics
raza@business.aau.dk

Aalborg University

Time Series Econometrics - Økomometri II

Outline

Univariate time series

- Imagine we have the following, general univariate time series model

$$Y_t = f\left(Y_{t-1}, Y_{t-2}, ..., u_t\right).$$

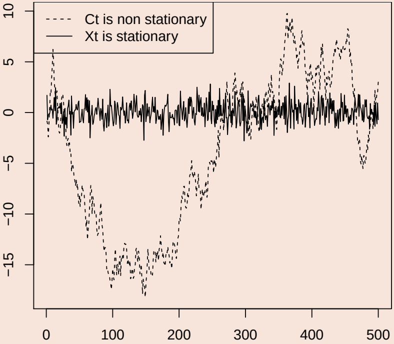
- To make this model operational we must specify three things;
  1. the functional form of $f\left(\right)$,
  2. the number of lags,
  3. and the structure of the disturbance term, u_t.

Stationarity:

- A stationary stochastic process is one whose ensemble statistics are the same for any value of time.
  - Strict stationarity (joint probability distribution of a stochastic variable is invariant over time)
  - Weak stationarity (first two moments of the joint probability distribution of a stochastic variable (i.e. the mean and variance-covariance matrix) are invariant over time)

- Nonstationary Process:
  - A nonstationary process will have a time-varying mean and/or a time-varying variance. This prevents us from making inferences outside of the sample period.

```r
TT <- 500; mean=0; sd=1
Xt <- ts(rnorm(TT, mean, sd), start=1, freq=1)
Ct <- ts(cumsum(rnorm(TT)), start=1, freq=1)
ts.plot(Ct, Xt, lty=2:1)
legend("topleft", c("Ct is non stationary", "Xt is stationary"), lty=2:1)
```



- Purely Random Process:
  - A stochastic process which has a zero mean, constant variance, and is serially uncorrelated is called a purely random process, or, "white noise".

$$\varepsilon_t \sim IID(0, \sigma^2).$$

Moving Average Process:

An $MA(1)$ Process:

$$Y_t = \mu + \varepsilon_t + \alpha\varepsilon_{t-1}$$
$$\varepsilon_t \sim IID(0, \sigma^2).$$

- Mean of Y_t

$$\begin{aligned} E\{Y_t\} &= E\{\mu + \varepsilon_t + \alpha\varepsilon_{t-1}\} \\ &= E\{\mu\} + E\{\varepsilon_t\} + E\{\alpha\varepsilon_{t-1}\} \\ &= \mu \end{aligned}$$

- Variance of Y_t

$$\begin{aligned} V\{Y_t\} &= E\{(Y_t - \mu)^2\} \\ &= E\{(\varepsilon_t + \alpha\varepsilon_{t-1})^2\} \\ &= (1 + \alpha^2)\sigma^2. \end{aligned}$$

Moving Average Process:

An $MA(1)$ Process:

$$Y_t = \mu + \varepsilon_t + \alpha\varepsilon_{t-1}$$
$$\varepsilon_t \sim IID(0, \sigma^2).$$

- Autocovariance between Y_t and Y_{t-1}

$$\begin{aligned} cov\{Y_t, Y_{t-1}\} &= E\{(Y_t - \mu)(Y_{t-1} - \mu)\} \\ &= \alpha\sigma^2. \end{aligned}$$

- Autocovariance between Y_t and Y_{t-2}

$$\begin{aligned} cov\{Y_t, Y_{t-2}\} &= E\{(Y_t - \mu)(Y_{t-2} - \mu)\} \\ &= 0. \end{aligned}$$

In general, $cov\{Y_t, Y_{t-k}\} = 0$ for $k = 2, 3, 4, \ldots$

Autoregressive Processes:

- An example of an $AR(1)$ process

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

$$\varepsilon_t \sim IID(0, \sigma^2).$$

- Mean of Y_t

$$\begin{aligned} E\{Y_t\} &= E\{\mu + \theta Y_{t-1} + \varepsilon_t\} \\ &= \frac{\mu}{1-\theta} \text{ if } |\theta| < 1. \end{aligned}$$

- Variance of Y_t (ignoring μ in the eq. for simplicity):

$$\begin{aligned} V\{Y_t\} &= E\{(Y_t - E\{Y_t\})^2\} \\ &= E\{(\theta Y_{t-1} + \varepsilon_t)^2\} \\ &= \frac{\sigma^2}{1-\theta^2} \text{ if } |\theta| < 1. \end{aligned}$$

Autoregressive Processes:

- Calculate the autocovariance between Y_t and Y_{t-1}

$$\begin{aligned} cov\{Y_t, Y_{t-1}\} &= E\{(Y_t - \mu)(Y_{t-1} - \mu)\} \\ &= \theta \frac{\sigma^2}{1-\theta^2} \text{ if } |\theta| < 1. \end{aligned}$$

- In general, the autocovariance between Y_s and Y_t can be written as

$$cov\{Y_s, Y_t\} = \theta^{|s-t|} \frac{\sigma^2}{1-\theta^2} \text{ if } |\theta| < 1.$$

Outline

Random walk model (without a drift):

- let us now take the $AR(1)$ model from above and set $\theta = 1$ and $\mu = 0$. This results in the well known random walk model

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

$$\varepsilon_t \sim IID(0, \sigma^2).$$

- Is this process stationary? No. Why not?

$$\begin{aligned} Y_T &= Y_{T-1} + \varepsilon_T \\ Y_T &= Y_{T-2} + \varepsilon_{T-1} + \varepsilon_T \\ Y_T &= Y_0 + \Sigma_1^T \varepsilon_t \end{aligned} \tag{1}$$

Let us apply expectations Equation 1:

$$E\{Y_T\} = = E\left\{Y_0 + \Sigma_1^T \varepsilon_t\right\} = Y_0$$

The mean is a finite constant. The variance, however, is not (recall the definition of stationarity)

Random walk model (without a drift):

- Let us derive the variance of the random walk model (taking the variance of eq. 1):

$$
\begin{aligned}
Var\left\{Y_T\right\} &= Var\left\{Y_0 + \Sigma_1^T \varepsilon_t\right\} \\
Var\left\{Y_T\right\} &= 0 + \Sigma_1^T Var\left\{\varepsilon_t\right\} \\
Var\left\{Y_T\right\} &= T\sigma^2
\end{aligned}
$$

- Note that variance in this model is a function of time (T)

Unit Root Setting $\theta = 1$ gives rise to the unit root problem. A unit root is synonymous with nonstationarity. The name unit root is due to the fact that the solution to the characteristic equation of an $AR(1)$ process has one root equal to unity when $\theta = 1$.

Random walk model (without a drift):

- Difference Stationary Process:
  - The random walk model can be made stationary by differencing the time series Y_t

$$\begin{aligned} Y_t - Y_{t-1} &= Y_{t-1} - Y_{t-1} + \varepsilon_t \rightarrow \\ \Delta Y_t &= \varepsilon_t. \end{aligned}$$

  We will make use of this fact quite often in this course. Processes which can be made stationary by differencing are called difference stationary processes.

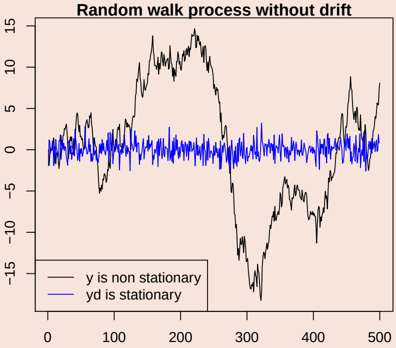
- Integrated Stochastic Process:
  - A stochastic process which can be made stationary through differencing is also known as an integrated stochastic process. If a stochastic process can be made stationary by taking first differences, it is said to be integrated of order one, I(1). If a stochastic process can be made stationary by taking d differences, it is said to be integrated of order d, I(d). Processes which are already stationary are said to be integrated of order zero, I(0).

The Random Walk Model without a Drift in R:

- Let us create a randowm walk (non-stationary) and make it stationary by differncing:

```
set.seed(154)
e  = rnorm(500,0,1)  # e is white noise
y <- 0
for (t in 2:500) {
    y[t] <- y[t-1] + e[t]
}

yd = diff(y) # I make it stationay by differncing
ts.plot(y, main = "Random walk process without drift", xlab="")
```



The Random Walk Model with Drift:

- let us now take the $AR(1)$ model from above and, once again, set $\theta = 1$. This time, however, we will allow $\mu \neq 0$. This results in a random walk model with drift:

$$
\begin{aligned}
Y_t &= \mu + Y_{t-1} + \varepsilon_t \\
\varepsilon_t &\sim IID(0, \sigma^2).
\end{aligned}
$$

- Stochastic Trend:
  - The time series Y_t will tend to drift upwards or downwards, depending on the sign of the drift parameter, μ. This drift is called a stochastic trend.
  - It will trend upwards if $\mu > 0$ and vice versa.
- Stationarity of the process:
  - Neither the mean nor the variance are constant
  - This process can also be made stationary by differencing

The Random Walk Model with Drift:

- Mean of the process:

$$
\begin{aligned}
E\{Y_T\} &= \mu + Y_{T-1} + \varepsilon_T \\
&= \mu + \mu + Y_{T-2} + \varepsilon_{T-1} + \varepsilon_T \\
&= Y_0 + \Sigma_1^T \mu + \Sigma_1^T \varepsilon_T \qquad (2)\\
&\quad \text{Apply expectations:} \\
&= E\left\{Y_0 + \Sigma_1^T \mu + \Sigma_1^T \varepsilon_T\right\} = Y_0 + T\mu
\end{aligned}
$$

- Variance of the process (take variance of eq. 2):

$$
\begin{aligned}
Var\{Y_T\} &= Var\left\{Y_0 + \Sigma_1^T \mu + \Sigma_1^T \varepsilon_t\right\} \\
Var\{Y_T\} &= 0 + 0 + \Sigma_1^T Var\{\varepsilon_t\} \\
Var\{Y_T\} &= T\sigma^2
\end{aligned}
$$

The Random Walk Model with Drift:

- The random walk model with drift can also be made stationary by differencing the time series Y_t

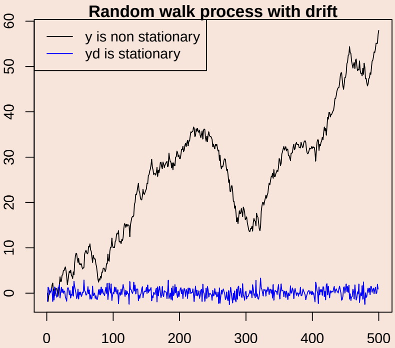
$$\begin{aligned} Y_t - Y_{t-1} &= \mu + Y_{t-1} - Y_{t-1} + \varepsilon_t \rightarrow \\ \Delta Y_t &= \mu + \varepsilon_t \\ \text{where } \varepsilon_t &\sim IID(0, \sigma^2). \end{aligned}$$

The Random Walk Model with Drift in R:

- Let us create a randowm walk with drift (non-stationary) and make it stationary by differncing:

```
set.seed(154)
e  = rnorm(500,0,1)  # e is white noise
y <- 0
mu <- 0.1 # stochastic trend is + so the series will trend upwards
for (t in 2:500) {
    y[t] <- mu + y[t-1] + e[t]
}

yd = diff(y) # I make it stationay by differncing
ts.plot(y, main = "Random walk process with drift", xlab="")
```



Trend Stationary Processes:

- Now let us add a time trend, t, to an otherwise stationary $AR(1)$ process, i.e. assume $|\theta| < 1$.

$$Y_t = \mu + t + \theta Y_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim IID(0, \sigma^2).$$

- This is a nonstationary process with mean

$$E\{Y_t\} = \frac{\mu + t}{1 - \theta}$$

and variance

$$V\{Y_t\} = \frac{\sigma^2}{1 - \theta^2}.$$

- Trend Stationary Processs:
  - This model can be made stationary by removing the deterministic time trend from the original time series. Processes which can be made stationary by removing a deterministic trend are called trend stationary processes (TSP).[a]

[a] H-P trends and quadratic, cubic, etc. trends are also deterministic even though they are not linear.

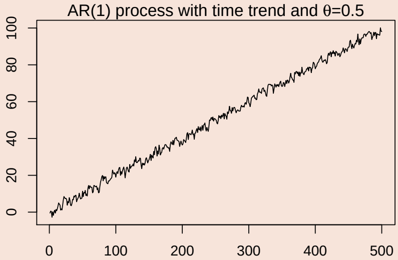
Autoregressive Process:

- Autoregressive Process AR(1) Process in R:

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

- We simulate an AR(1) process with deterministic trend in R:

```
set.seed(213)
n=500
e <- rnorm(n, mean = 0, sd = 1.5)
Y <- rnorm(1)
mu <- 0.1
t=1:n
for (t in 2:n) {
    Y[t] <-  mu +  0.5*Y[t-1] + e[t] + 0.1*t
}
```



Outline

The Lag Operator

- See pages 39-42 in Enders.
- Some basic rules for using lag operators:
  1. $L\alpha = \alpha$
  2. $L(x_t) = x_{t-1}$
  3. $L^2(x_t) = L\left[L(x_t)\right] = L(x_{t-1}) = x_{t-2}$
  4. $L^n(x_t) = x_{t-n}$
  5. $(1 - L)\, x_t = x_t - L(x_t) = x_t - x_{t-1} = \Delta x_t$
  6. $L\,(1 - L)\, x_t = (1 - L)x_{t-1} = x_{t-1} - Lx_{t-1} = x_{t-1} - x_{t-2} = \Delta x_{t-1}$

Converting an AR(1) process into MA(∞)

- Now, let's use these rules to show that an AR(1) process can be converted into an MA(∞) process (and vice versa). First, write down an AR(1) process:

$$Y_t = \alpha Y_{t-1} + \varepsilon_t \implies \alpha L Y_t + \varepsilon_t$$

$$Y_t - \alpha L Y_t = \varepsilon_t$$

$$Y_t(1 - \alpha L) = \varepsilon_t$$

$$Y_t = \frac{\varepsilon_t}{(1 - \alpha L)}$$

Recall the geometric series theorem: $\sum_{n=0}^{\infty} = ar^0 + ar^1 + ar^2 + ... + ar^n + ...$ converges to $a/(1-r)$, if $|r| < 1$

Stationarity condition: It is here, since we are dealing with a stationary AR(1) process with $|\alpha| < 1$

Assume $r = \alpha L$ in this case:

$$Y_t = (\alpha L)^0 \varepsilon_t + (\alpha L)^1 \varepsilon_t + (\alpha L)^2 \varepsilon_t + ... + (\alpha L)^n \varepsilon_t$$

Applying rule no. 2 and 4 in the lag operators, we can get our MA(∞) process:

$$Y_t = \varepsilon_t + \alpha^1 \varepsilon_{t-1} + \alpha^2 \varepsilon_{t-2} + ... + \alpha^n \varepsilon_{t-n} = \sum_{n=0}^{\infty} \alpha^n \varepsilon_{t-n}$$

Inverting an MA(1) process into AR(∞)

Exercise: Can you invert an MA(1) process into AR(∞)?

Outline

- Assuming weak stationarity, we can define the k-th order autocovariance, γ_k, as

$$\gamma_k = cov\{Y_t, Y_{t-k}\} = cov\{Y_t, Y_{t+k}\}.$$

- The autocovariance of a stochastic process can be standardized and presented as an autocorrelation function (ACF), ρ_k

$$\rho_k = \frac{cov\{Y_t, Y_{t-k}\}}{V\{Y_t\}} = \frac{\gamma_k}{\gamma_0}.$$

- What do we use the ACF for?
  - Characterize the development of Y_t over time.
  - Show us how strongly current observations are correlated with past observations
  - Show how shocks today affect future values of the stochastic variable.
  - Help us find unit roots
  - Choose models
  - Run regression diagnostics.

- The ACF of an $AR(1)$ process is

$$\rho_k = \frac{cov\{Y_t, Y_{t-k}\}}{V\{Y_t\}} = \frac{\theta^k \frac{\sigma^2}{1-\theta^2}}{\frac{\sigma^2}{1-\theta^2}} = \theta^k.$$

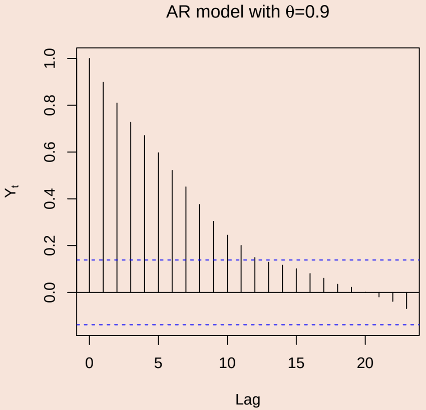
- The ACF of an $MA(1)$ process is

$$\rho_1 = \frac{cov\{Y_t, Y_{t-1}\}}{V\{Y_t\}} = \frac{\alpha\sigma^2}{(1+\alpha^2)\sigma^2} = \frac{\alpha}{1+\alpha^2}; \rho_k = 0, k > 1.$$

- In $\mathbb{R}$ you can graph the ACF of a variable, x, by writing: `acf(x)`
- It is often times difficult to distinguish between different AR(q) processes based solely on an examination of a correlogram (see Enders pp. 61).

ACF of AR(1) process

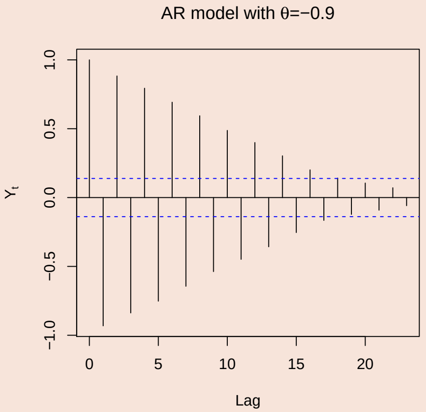
```
TT <- 200
acf(arima.sim(model=list(ar=0.9), n=TT), ylab=expression(Y[t]),
     main=expression(paste("AR model with ", theta, "=0.9")))
```



- Note: In an AR(1) process, y_t and y_{t-2} are correlated even though y_{t-2} does not directly appear in the model.
- The acf function detects these indirect correlations and therefore acf of an AR process has a slow decay

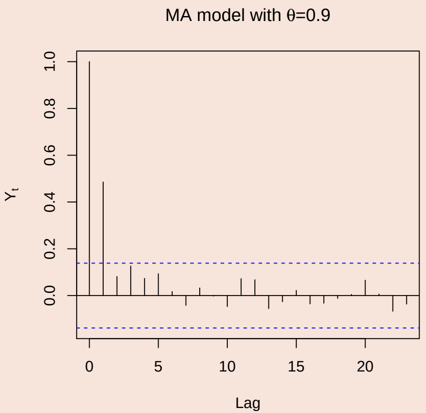
ACF of AR(1) process

```
TT <- 200
acf(arima.sim(model=list(ar=-0.9), n=TT), ylab=expression(Y[t]),
     main=expression(paste("AR model with ", theta, "=-0.9")))
```



ACF of MA(1) process

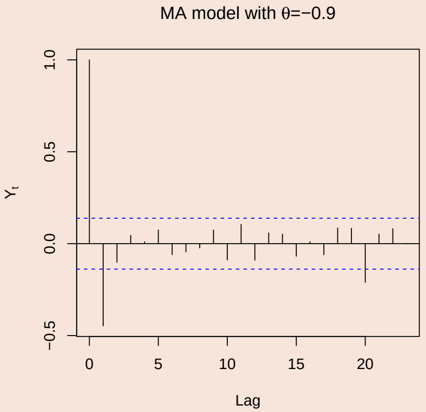
```
TT <- 200
acf(arima.sim(model=list(ma=0.9), n=TT), ylab=expression(Y[t]),
     main=expression(paste("MA model with ", theta, "=0.9")))
```



- Note: In an MA(1) process, the correlation between y_t and y_{t-2} is zero.
- Therefore acf of an MA(1) process has a sudden decay.

ACF of MA(1) process

```
TT <- 200
acf(arima.sim(model=list(ma=-0.9), n=TT), ylab=expression(Y[t]),
     main=expression(paste("MA model with ", theta, "=-0.9")))
```



Outline

Partial autocorrelation function(PACF)

- Assume an AR(1) process:

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

- Recall, the acf function detects indirect correlations between Y_t and Y_{t-2} although Y_{t-2} is not part of the original model.
- PACF provides a more clear picture
- Assume an AR(2) process:

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \varepsilon_t$$

- θ_2 is the partial correlation between Y_t and Y_{t-2} while holding the effect of Y_{t-1} fixed.
- One can adjust for the effect of indirect correlations by computing partial autocorrelations from the autocorrelations as:

$$\theta_1 = \rho_1$$

$$\theta_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

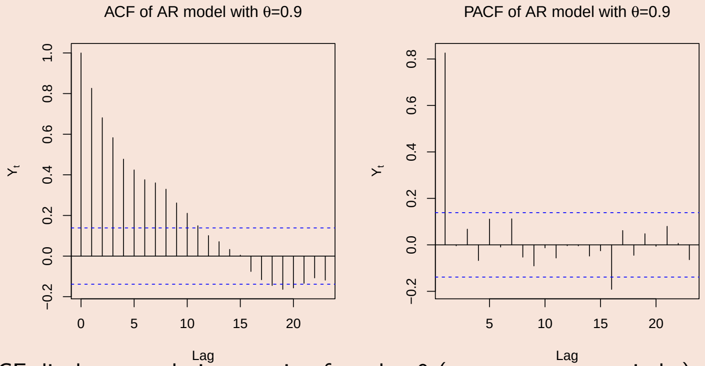
- These can be graphed in $\mathbb{R}$ by writing: `pacf(x)`

Partial autocorrelation function(PACF)

```
TT <- 200
x= arima.sim(model=list(ar=0.9), n=TT)

acf(x, ylab=expression(Y[t]), main= expression(paste("ACF of AR model with ", theta, "=0.9")))

pacf(x, ylab=expression(Y[t]), main= expression(paste("PACF of AR model with ", theta, "=0.9")))
```



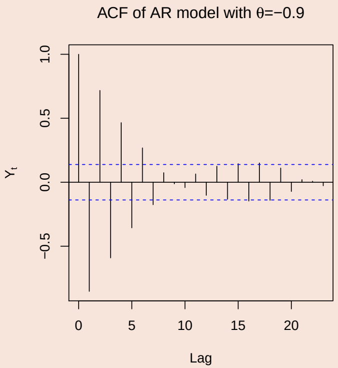
Note: ACF displays correlation starting from lag 0 (means current period t) / PACF shows partial correlation starting from lag 1 (means t-1)

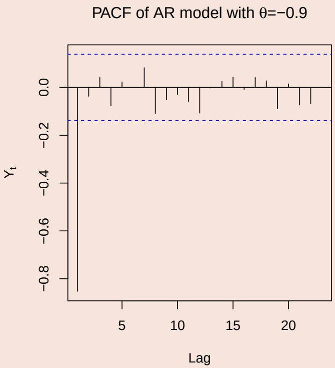
Partial autocorrelation function(PACF)

```
TT <- 200
x= arima.sim(model=list(ar=-0.9), n=TT)

acf(x, ylab=expression(Y[t]), main= expression(paste("ACF of AR model with ", theta, "=-0.9")))

pacf(x, ylab=expression(Y[t]), main= expression(paste("PACF of AR model with ", theta, "=-0.9")))
```



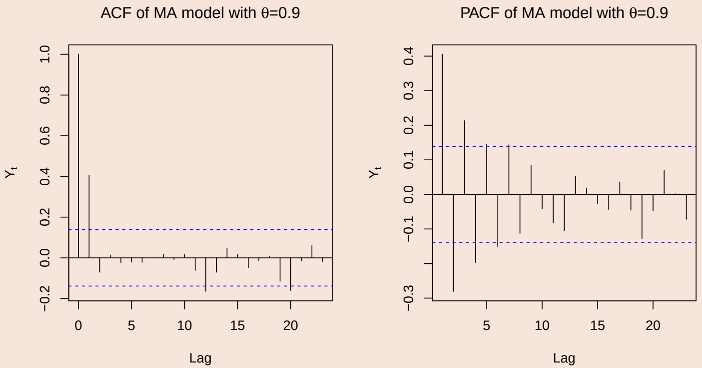


Partial autocorrelation function(PACF)

```
TT <- 200
x= arima.sim(model=list(ma=0.9), n=TT)

acf(x, ylab=expression(Y[t]), main= expression(paste("ACF of MA model with ", theta, "=0.9")))

pacf(x, ylab=expression(Y[t]), main= expression(paste("PACF of MA model with ", theta, "=0.9")))
```

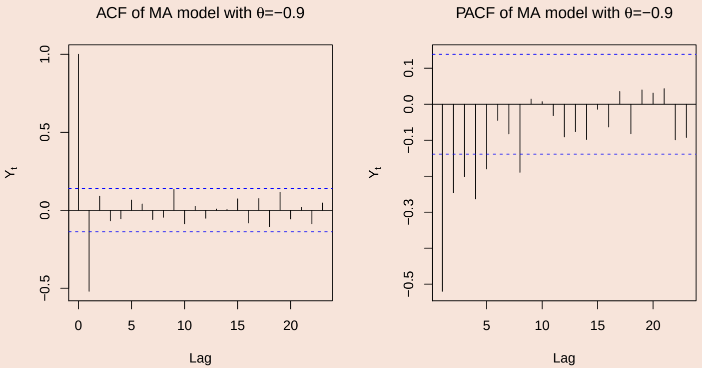


Partial autocorrelation function(PACF)

```
TT <- 200
x= arima.sim(model=list(ma=-0.9), n=TT)

acf(x, ylab=expression(Y[t]), main= expression(paste("ACF of MA model with ", theta, "=-0.9")))

pacf(x, ylab=expression(Y[t]), main= expression(paste("PACF of MA model with ", theta, "=-0.9")))
```



Summary of Results

- An $AR(p)$ process is described by:
    1. an ACF that is infinite in extent (it tails off)
    2. a PACF that is (close to) zero for lags larger than p.
- A $MA(q)$ process is described by:
    1. an ACF that is (close to) zero for lags larger than q.
    2. a PACF that is infinite in extent (it tails off).
- Note that this "identification" strategy only works on stationary time series! For example, if you have quarterly data that trends upwards, you must first remove the seasonal pattern and then remove the long-run trend. After this is done, you can use the acf and pacf functions to identify the de-trended variable.

Outline

Exercise:

1. Generate a white noise series, an $AR(1)$ series and an $AR(2)$ series. Compare the ACFs and PACFs of these 3 series. Do the ACFs and PACFs actually help you to distinguish between the different series? That is, are you able to identify the order of the $AR(p)$ processes by looking at their ACFs and PACFs?

   Note: The stationarity conditions for an AR(2) process are; $\theta_1 + \theta_2$, $\theta_2 - \theta_1 < 1$ and $|\theta_2| < 1$. Write these conditions as a text reminder in your $AR(2)$ .R file, so that you will have them for future reference.

2. Use the ACF and PACF to describe two actual data series. Take two of the series that you used in Exercise $\#1$.

   Note: Remember that this identification strategy only works for stationary time series. You should remove any season and/or long-run trend first.