# ECON7350: Applied Econometrics for Macroeconomics and Finance

## Tutorial 2: Forecasting Univariate Processes - I

At the end of this tutorial you should be able to:

- derive theoretical properties of ARMA processes;
- compute the theoretical ACF and PACF for a given ARMA processes;
- use R to compute and plot the sample ACF and PACF for time series data.



## Problems with Solutions



1. Derive the expected value, variance, covariance, autocorrelation function (ACF), and partial autocorrelation function (PACF) for the following data generating processes (DGPs):

   (a) AR(1): $y_t = a_0 + a_1 y_{t-1} + \epsilon_t$, $0 \leq |a_1| < 1$;



---

**Solution**

- Expected value:

$$y_t = a_0 + a_1 y_{t-1} + \epsilon_t; \quad 0 \leq |a_1| < 1$$

$$\begin{aligned} \mathrm{E}(y_t) = \mu &= a_0 + a_1 \mathrm{E}(y_{t-1}) + \mathrm{E}(\epsilon_t) \\ \mu &= \frac{a_0}{1 - a_1}; \text{ since } \mathrm{E}(y_{t-1}) = \mu \end{aligned}$$

- Variance:

$$\begin{aligned} \mathrm{Var}(y_t) = \gamma_0 &= a_1^2 \mathrm{Var}(y_{t-1}) + \mathrm{Var}(\epsilon_t) + 2\mathrm{cov}(a_1 y_{t-1}, \epsilon_t) \\ \gamma_0 &= \frac{\sigma^2}{1 - a_1^2}; \text{ since } \mathrm{Var}(y_{t-1}) = \gamma_0, \ \mathrm{cov}(y_{t-1}, \epsilon_t) = 0 \end{aligned}$$

- Covariance:
  - Set $a_0 = 0$ without loss of generality

$$\begin{aligned}\mathrm{cov}(y_t, y_{t-k}) = \gamma_k &= \mathrm{E}(y_t y_{t-k}) \\ &= \mathrm{E}((a_1 y_{t-1} + \epsilon_t) y_{t-k})\end{aligned}$$

  - $\gamma_1$ $(k = 1)$

$$\begin{aligned}\gamma_1 &= \mathrm{E}((a_1 y_{t-1} + \epsilon_t) y_{t-1}) \\ &= a_1 \frac{\sigma^2}{1 - a_1^2} = a_1 \gamma_0\end{aligned}$$

  - $\gamma_2$ $(k = 2)$

$$\begin{aligned}\gamma_2 &= \mathrm{E}((a_1 y_{t-1} + \epsilon_t) y_{t-2}) \\ &= a_1^2 \frac{\sigma^2}{1 - a_1^2} = a_1^2 \gamma_0\end{aligned}$$



  - $\gamma_k$ $(k > 2)$

$$\begin{aligned}\gamma_k &= \mathrm{E}((a_1 y_{t-1} + \epsilon_t) y_{t-k}) \\ &= a_1^k \frac{\sigma^2}{1 - a_1^2} = a_1^k \gamma_0\end{aligned}$$





- Autocorrelation:
  - $\rho_1$ $(k = 1)$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = a_1$$

  - $\rho_2$ $(k = 2)$

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = a_1^2$$

  - $\rho_k$ $(k > 2)$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = a_1^k$$

- Partial autocorrelation:
  - $\phi_{11}$

$$\phi_{11} = \rho_1 = a_1$$

  - $\phi_{22}$

$$\begin{aligned}\phi_{22} &= (\rho_2 - \rho_1^2)/(1 - \rho_1^2) \\ &= (a_1^2 - a_1^2)/(1 - a_1^2) \\ &= 0\end{aligned}$$

– $\phi_{33}$

$$\begin{aligned}
\phi_{33} &= \frac{\rho_3 - \sum_{j=1}^{3-1} \phi_{2,j}\rho_{3-j}}{1 - \sum_{j=1}^{3-1} \phi_{3-1,j}\rho_j} \\
&= \frac{a_1^3 - \phi_{21}\rho_2 - \phi_{22}\rho_1}{1 - \sum_{j=1}^{3-1} \phi_{3-1,j}\rho_j} \\
&= \frac{a_1^3 - a_1 a_1^2 + 0}{1 - \sum_{j=1}^{3-1} \phi_{3-1,j}\rho_j} \\
&= 0
\end{aligned}$$

since

$$\begin{aligned}
\phi_{21} &= \phi_{1,1} - \phi_{22}\phi_{1,1} \\
&= \phi_{1,1}
\end{aligned}$$

---



(b) MA(1): $y_t = b_0 + b_1\epsilon_{t-1} + \epsilon_t$;

---

**Solution**



- Expected value:



$$\begin{aligned}
\mathrm{E}(y_t) &= b_0 + b_1\mathrm{E}(\epsilon_{t-1}) + \mathrm{E}(\epsilon_t) \\
&= \mu
\end{aligned}$$

- Variance:

$$\begin{aligned}
\mathrm{Var}(y_t) = \gamma_0 &= b_1^2\mathrm{Var}(\epsilon_{t-1}) + \mathrm{Var}(\epsilon_t) + 2\mathrm{cov}(\epsilon_t, \epsilon_{t-1}) \\
\gamma_0 &= \sigma^2(1 + b_1^2)
\end{aligned}$$

- Covariance:
  - Set $\mu = 0$ without loss of generality

$$\begin{aligned}
\mathrm{cov}(y_t, y_{t-k}) = \gamma_k &= \mathrm{E}(y_t y_{t-k}) \\
&= \mathrm{E}((b_1\epsilon_{t-1} + \epsilon_t)y_{t-k})
\end{aligned}$$

$\mathrm{cov}(y_t, y_{t-k}) > 0$ for $k = 1$, $\mathrm{cov}(y_t, y_{t-k}) = 0$ for $k > 1$

  - $\gamma_1$ $(k = 1)$

$$
\begin{aligned}
\gamma_1 &= \mathrm{E}((b_1\epsilon_{t-1} + \epsilon_t)y_{t-1}) \\
&= \mathrm{E}(b_1\epsilon_{t-1}(b_1\epsilon_{t-2} + \epsilon_{t-1}) + \epsilon_t y_{t-1}) \\
&= b_1\sigma^2 \\
&= \frac{b_1}{1+b_1^2} \times \gamma_0; \text{ since } \sigma^2 = \gamma_0/(1+b_1^2)
\end{aligned}
$$

– $\gamma_2$ $(k = 2)$

$$
\begin{aligned}
\gamma_2 &= \mathrm{E}((b_1\epsilon_{t-1} + \epsilon_t)y_{t-2}) \\
&= 0; \text{ since } y_{t-2} \text{ is not a function of } \epsilon_t \text{ or } \epsilon_{t-1}
\end{aligned}
$$

– $\gamma_k$ $(k > 2)$

$$
\begin{aligned}
\gamma_k &= \mathrm{E}((b_1\epsilon_{t-1} + \epsilon_t)y_{t-k}) \\
&= 0
\end{aligned}
$$

- Autocorrelation:



– $\rho_1$ $(k = 1)$

$$
\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{b_1}{1+b_1^2}
$$



– $\rho_k$ $(k > 1)$

$$
\rho_k = \frac{\gamma_k}{\gamma_0} = 0
$$



- Partial autocorrelation:

– $\phi_{11}$

$$
\phi_{11} = \rho_1
$$

– $\phi_{22}$

$$
\begin{aligned}
\phi_{22} &= (\rho_2 - \rho_1^2)/(1-\rho_1^2) \\
&= (0 - \rho_1^2)/(1-\rho_1^2) \\
&= -\rho_1^2/(1-\rho_1^2)
\end{aligned}
$$

– $\phi_{33}$

$$
\begin{aligned}
\phi_{33} &= \frac{\rho_3 - \sum_{j=1}^{3-1}\phi_{2,j}\rho_{3-j}}{1 - \sum_{j=1}^{3-1}\phi_{2,j}\rho_j} \\
&= \frac{\rho_3 - \phi_{21}\rho_2 - \phi_{22}\rho_1}{1 - \sum_{j=1}^{3-1}\phi_{2,j}\rho_j} \\
&= \frac{\rho_1^3/(1-\rho_1^2)}{1 - \sum_{j=1}^{3-1}\phi_{2,j}\rho_j}; \text{ since } \rho_2 = \rho_3 = 0
\end{aligned}
$$

---

ARMA(1,1): $y_t = a_0 + a_1 y_{t-1} + b_1 \epsilon_{t-1} + \epsilon_t, 0 \leq |a_1| < 1$.

---

**Solution**

- Expected value:

$$\begin{aligned} \mathrm{E}(y_t) &= a_0 + a_1 \mathrm{E}(y_{t-1}) + b_1 \mathrm{E}(\epsilon_{t-1}) + \mathrm{E}(\epsilon_t) \\ \mu &= \frac{a_0}{1 - a_1}; \text{ since } \mathrm{E}(y_t) = \mathrm{E}(y_{t-1}) = \mu \end{aligned}$$

- Variance:

$$\begin{aligned} \mathrm{Var}(y_t) = \gamma_0 &= \mathrm{Var}(a_0) + a_1^2 \mathrm{Var}(y_{t-1}) + b_1^2 \mathrm{Var}(\epsilon_{t-1}) + \mathrm{Var}(\epsilon_t) \\ &+ 2\mathrm{cov}(a_1 y_{t-1}, b_1 \epsilon_{t-1}) + 2\mathrm{cov}(a_1 y_{t-1}, \epsilon_t) + 2\mathrm{cov}(b_1 \epsilon_{t-1}, \epsilon_t) \\ \gamma_0 &= \frac{1 + b_1^2 + 2a_1 b_1}{1 - a_1^2} \sigma^2, \text{ since } \mathrm{cov}(a_1 y_{t-1}, b_1 \epsilon_{t-1}) = a_1 b_1 E(\epsilon_{t-1}^2) \end{aligned}$$



  - To show $\mathrm{cov}(a_1 y_{t-1}, b_1 \epsilon_{t-1}) = a_1 b_1 E(\epsilon_{t-1}^2)$ you can proceed as follows

$$\begin{aligned} \mathrm{cov}(a_1 y_{t-1}, b_1 \epsilon_{t-1}) &= E[(a_1 y_{t-1})(b_1 \epsilon_{t-1})] \\ &= \mathrm{E}([a_1(a_1 y_{t-2} + b_1 \epsilon_{t-2} + \epsilon_{t-1})](b_1 \epsilon_{t-1})) \\ &= \mathrm{E}(a_1 \epsilon_{t-1} b_1 \epsilon_{t-1}) \end{aligned}$$

  is the only non-zero expected value.

- Covariance:
  - Set $\mu = 0$ without loss of generality

$$\begin{aligned} \mathrm{cov}(y_t, y_{t-k}) = \gamma_k &= \mathrm{E}(y_t y_{t-k}) \\ &= \mathrm{E}((a_1 y_{t-1} + b_1 \epsilon_{t-1} + \epsilon_t) y_{t-k}) \end{aligned}$$

  - $\gamma_1$ $(k = 1)$

$$\gamma_1 = \frac{(1 + a_1 b_1)(a_1 + b_1)}{1 - a_1^2} \sigma^2$$

  - $\gamma_k$ $(k \geq 2)$

$$\gamma_k = a_1 \gamma_{k-1}$$

- Autocorrelation:

- $\rho_1$ $(k = 1)$

$$\rho_1 = \frac{(1 + a_1 b_1)(a_1 + b_1)}{1 + b_1^2 + 2a_1 b_1}$$

- $\rho_k$ $(k \geq 2)$

$$\rho_k = a_1 \rho_{k-1}$$

- Autoregressive pattern dominates for $k > 1$.

• Partial autocorrelation:

- $\phi_{11}$

$$\phi_{11} = \rho_1$$

- $\phi_{22}$

$$\begin{aligned} \phi_{22} &= (\rho_2 - \rho_1^2)/(1 - \rho_1^2) \\ &= (a_1\rho_1 - \rho_1^2)/(1 - \rho_1^2) \end{aligned}$$

- $\phi_{33}$



$$\begin{aligned} \phi_{33} &= \frac{\rho_3 - \sum_{j=1}^{3-1} \phi_{2,j}\rho_{3-j}}{1 - \sum_{j=1}^{3-1} \phi_{2,j}\rho_j} \\ &= \frac{a_1^2\rho_1 - \phi_{21}a_1\rho_1 - \phi_{22}\rho_1}{1 - \sum_{j=1}^{3-1} \phi_{2,j}\rho_j} \end{aligned}$$



where



$$\begin{aligned} \phi_{21} &= \phi_{11} - \phi_{22}\phi_{11} \\ &= \rho_1[1 - (a_1\rho_1\text{-}\rho_1^2)/(1\text{-}\rho_1^2)] \end{aligned}$$

- Moving average pattern dominates for $k > 1$.

---

2. Compute the true ACF values for the following DGPs:

   • DGP1: $y_t = 0.75y_{t-1} + \epsilon_t$;

---

**Solution** $\rho_0 = 1$, $\rho_1 = 0.75$, $\ldots$, $\rho_k = 0.75^k$. The ACF will decay geometrically.

---

   • DGP2: $y_t = -0.75y_{t-1} + \epsilon_t$;

---

**Solution** $\rho_0 = 1$, $\rho_1 = -0.75$, $\ldots$, $\rho_k = (-1)^k 0.75^k$. The ACF will decay in a dampened oscillatory path.

---

- DGP3: $y_t = 0.95 y_{t-1} + \epsilon_t$;

---

**Solution** $\rho_0 = 1$, $\rho_1 = 0.95$, $\ldots$, $\rho_k = 0.95^k$. The ACF will decay geometrically but at a much slower rate than DGP1.

---

- DGP4: $y_t = 0.5 y_{t-1} + 0.25 y_{t-2} + \epsilon_t$;

---

**Solution** For the AR(2) model $y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \epsilon_t$, $\rho_0 = 1$, $\rho_1 = a_1/(1 - a_2)$, $\ldots$, $\rho_k = a_1 \rho_{k-1} + a_2 \rho_{k-2}$. Thus, $\rho_0 = 1$, $\rho_1 = 2/3$, $\rho_2 = 7/12$, $\ldots$, $\rho_k = a_1 \rho_{k-1} + a_2 \rho_{k-2}$ for $k \geq 2$.



---

- DGP5: $y_t = 0.25 y_{t-1} - 0.5 y_{t-2} + \epsilon_t$;

---

**Solution** $\rho_0 = 1$, $\rho_1 = 1/6$, $\rho_2 = -11/24$, $\ldots$, $\rho_k = a_1 \rho_{k-1} + a_2 \rho_{k-2}$ for $k \geq 2$.

---

- DGP6: $y_t = 0.75 \epsilon_{t-1} + \epsilon_t$;

---

**Solution** For the MA($q$) model $y_t = b_0 + b_1 \epsilon_{t-1} + \cdots + b_q \epsilon_{t-q} + \epsilon_t$, the ACF cuts off at $k = q$—i.e., $\rho_k = 0$ for all $k > q$. Thus, $\rho_0 = 1$, $\rho_1 = b_1/(1 + b_1^2) = 12/25$, $\rho_k = 0$ for $k \geq 2$.

---

- DGP7: $y_t = 0.75 \epsilon_{t-1} - 0.5 \epsilon_{t-2} + \epsilon_t$;

---

**Solution** $\rho_0 = 1$ and $\rho_k = 0$ for $k \geq 3$. $\rho_1 = b_1(1+b_2)/(1+b_1^2+b_2^2) = 6/29$ and $\rho_2 = b_2/(1+b_1^2+b_2^2) = -8/29$.

---

- DGP8: $y_t = 0.75y_{t-1} + 0.5\epsilon_{t-1} + \epsilon_t$.

---

**Solution** For the ARMA(1,1) model $y_t = a_0 + a_1 y_{t-1} + b_1\epsilon_{t-1} + \epsilon_t$, $\rho_0 = 1$, $\rho_1 = (1+a_1b_1)(a_1+b_1)/(1+b_1^2+2a_1b_1)$, $\rho_k = a_1\rho_{k-1}$ for all $k \geq 2$. Thus, $\rho_0 = 1, \rho_1 = 0.859$, $\rho_2 = 0.645$, ...

---

3. The data file `arma.csv` contains (simulated) data for each of the DGPs in Question 2. Import the data into R. Compute, plot, and describe the behaviour of the ACF and PACF for each DGP. Discuss the effects of parameter signs. Hint: use the `acf` and `pacf` commands, respectively.



---

**Solution** Load the data using the `read.delim` command with the `sep = ","` option as it is comma delimited.

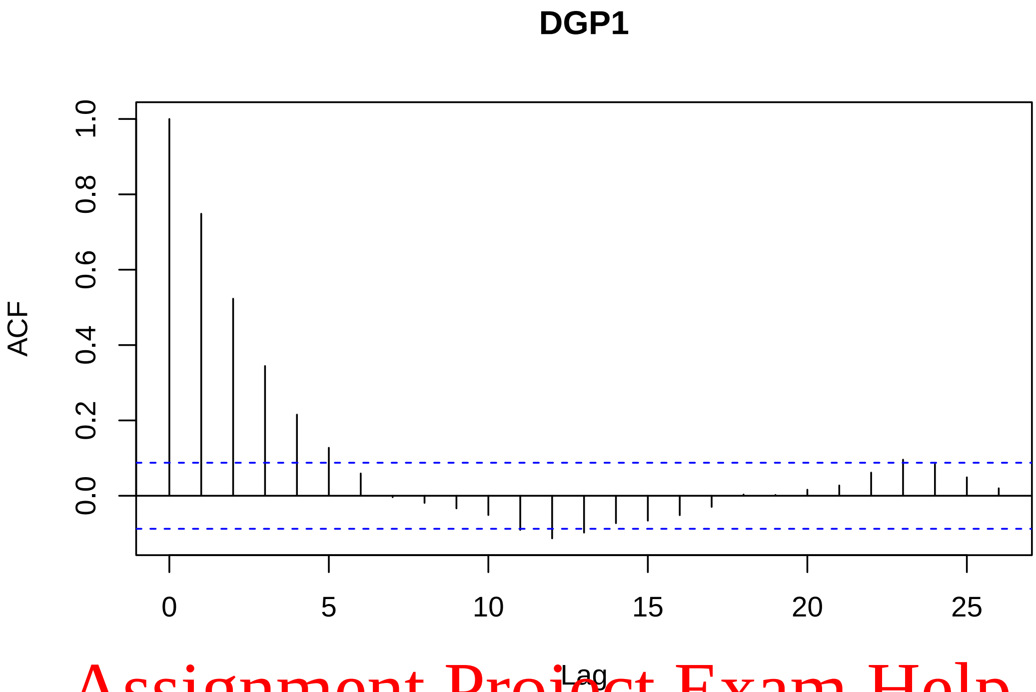


```
mydata <- read.delim("arma.csv", header = TRUE,  sep = ",")
```

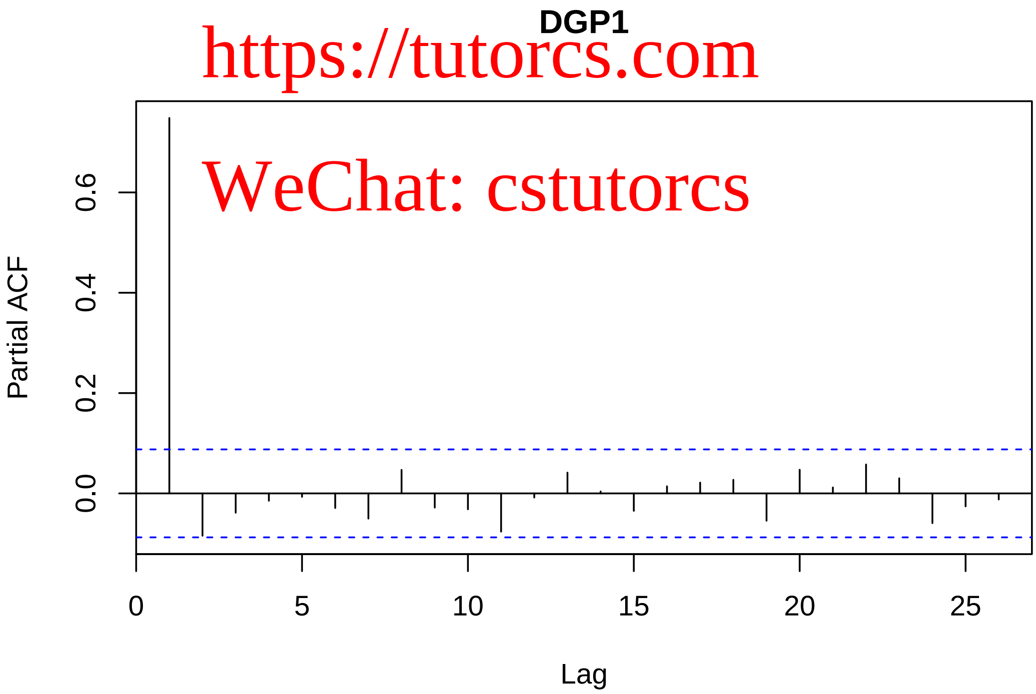


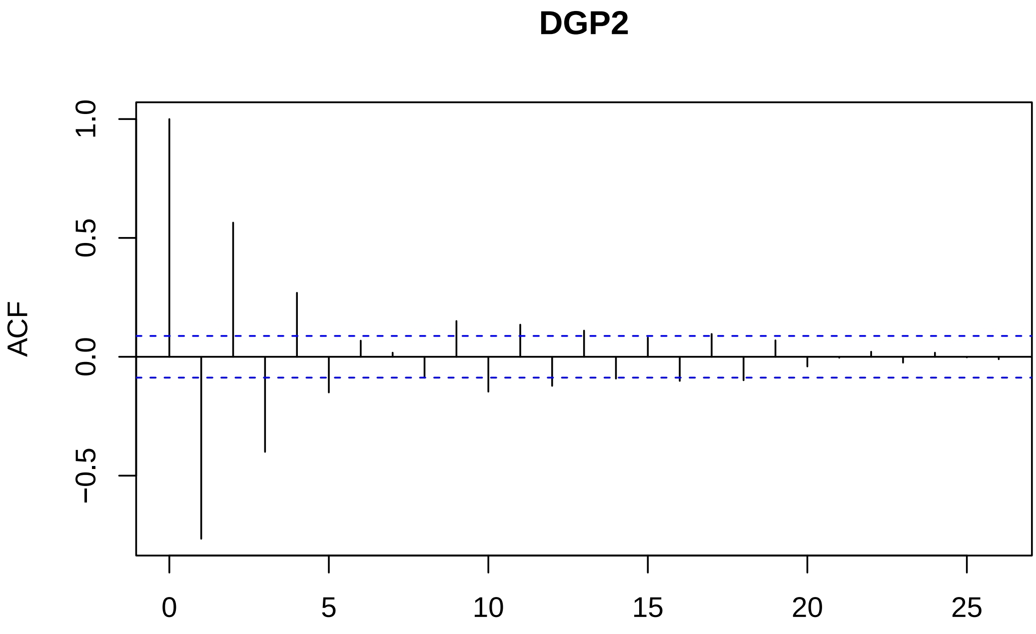
The ACF and PACF plots can be generated for all eight GDPs quickly using the `for` loop. Note that we index each column in `mydata` as `1 + i` because the first column contains the time variable `t`. The option `main` is passed to `plot`—we assign it the name of a given column, which corresponds to the GDP in the loop that the sample ACF/PACF are being computed for.

```
for (i in 1:8)
{
  acf(mydata[1 + i], main = colnames(mydata[1 + i]))
  pacf(mydata[1 + i], main = colnames(mydata[1 + i]))
}
```


DGP1
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
Lag

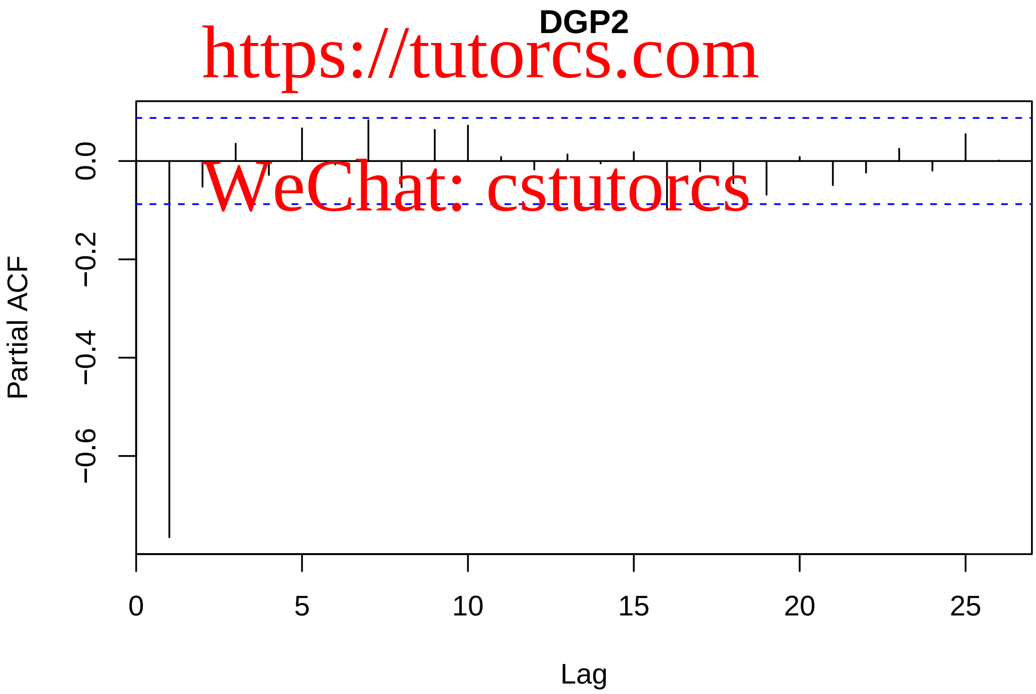


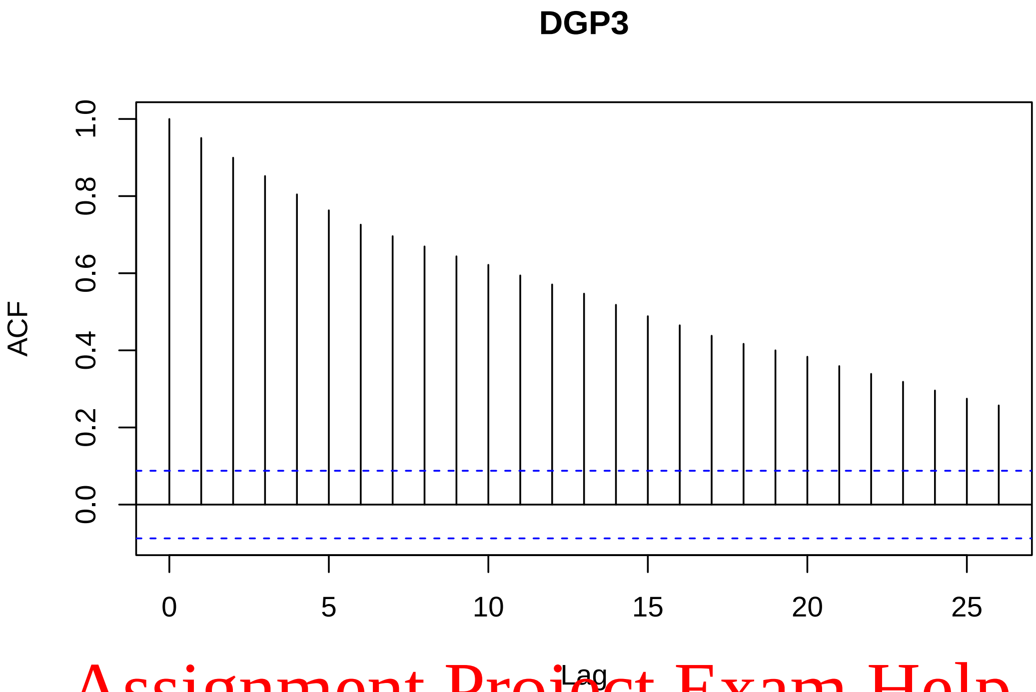

DGP1
Partial ACF
0.6
0.4
0.2
0.0
0
5
10
15
20
25
Lag


DGP2
ACF
1.0
0.5
0.0
−0.5
0
5
10
15
20
25
Lag

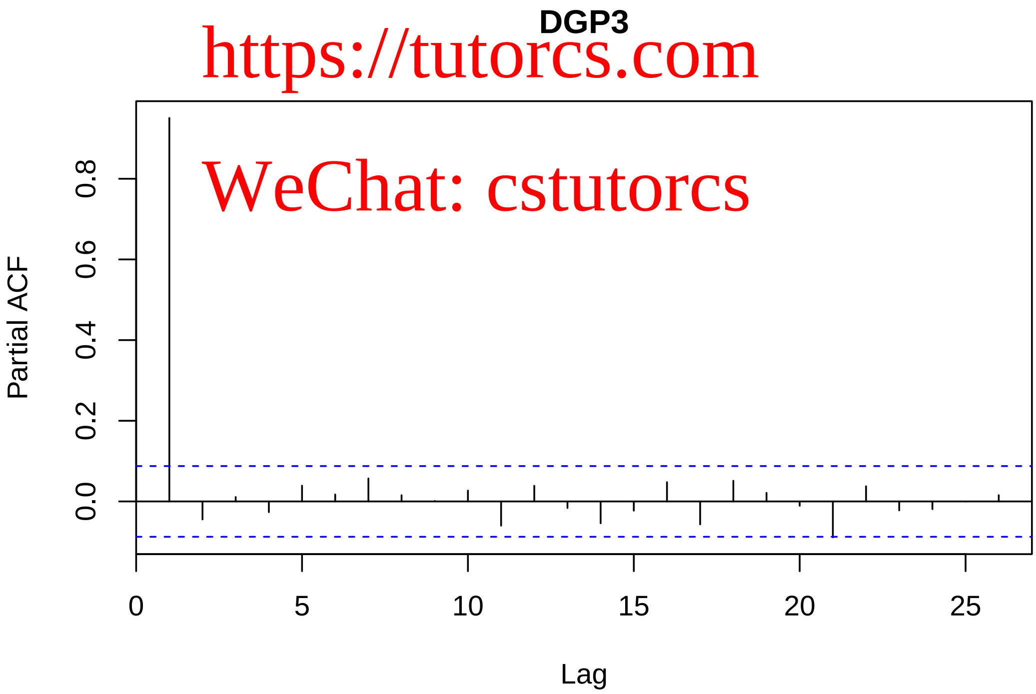


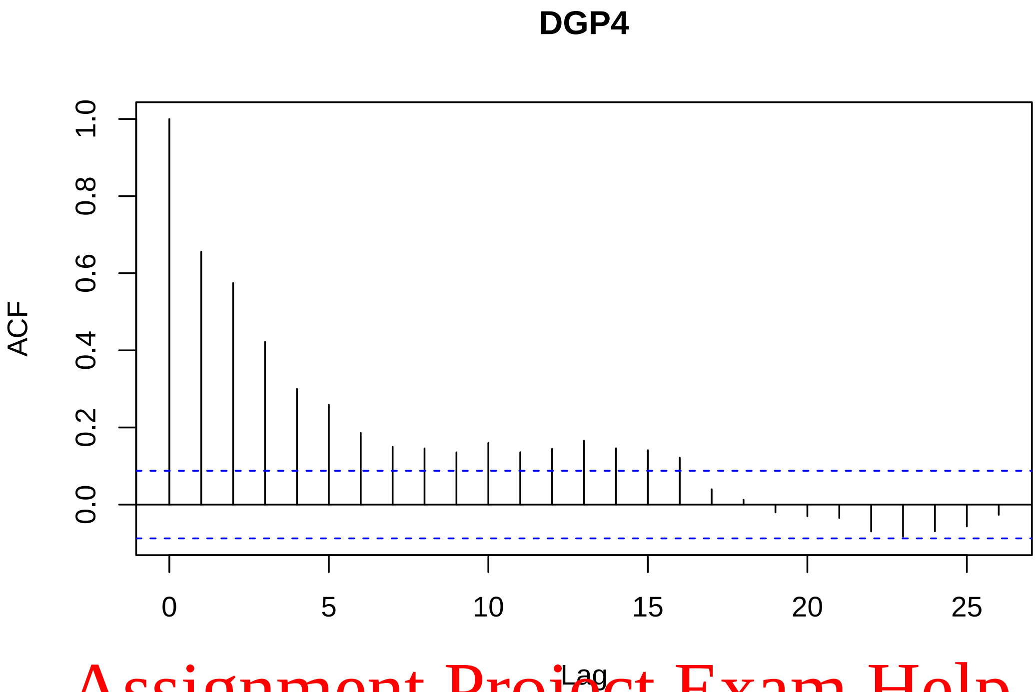



DGP2
Partial ACF
0.0
−0.2
−0.4
−0.6
0
5
10
15
20
25
Lag

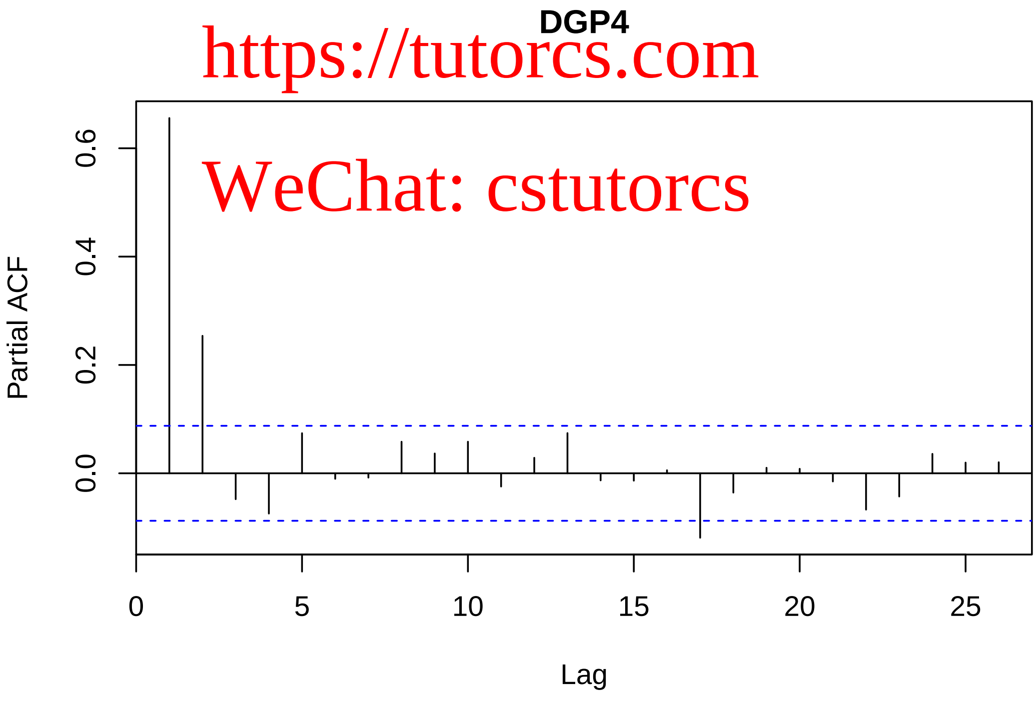

DGP3
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
Lag

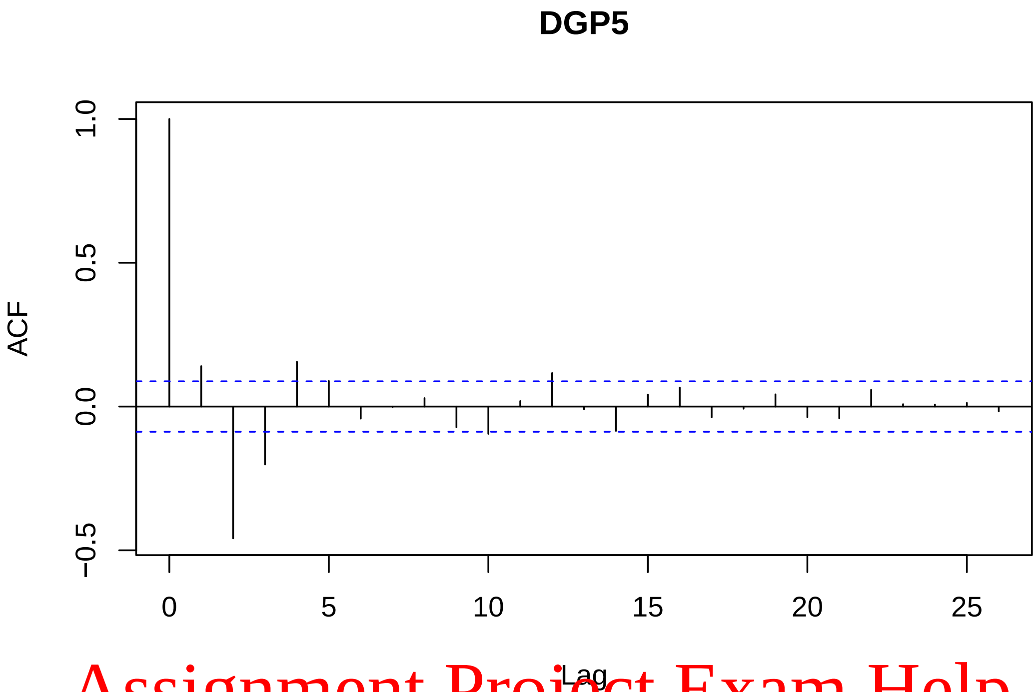



DGP3
Partial ACF
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
Lag

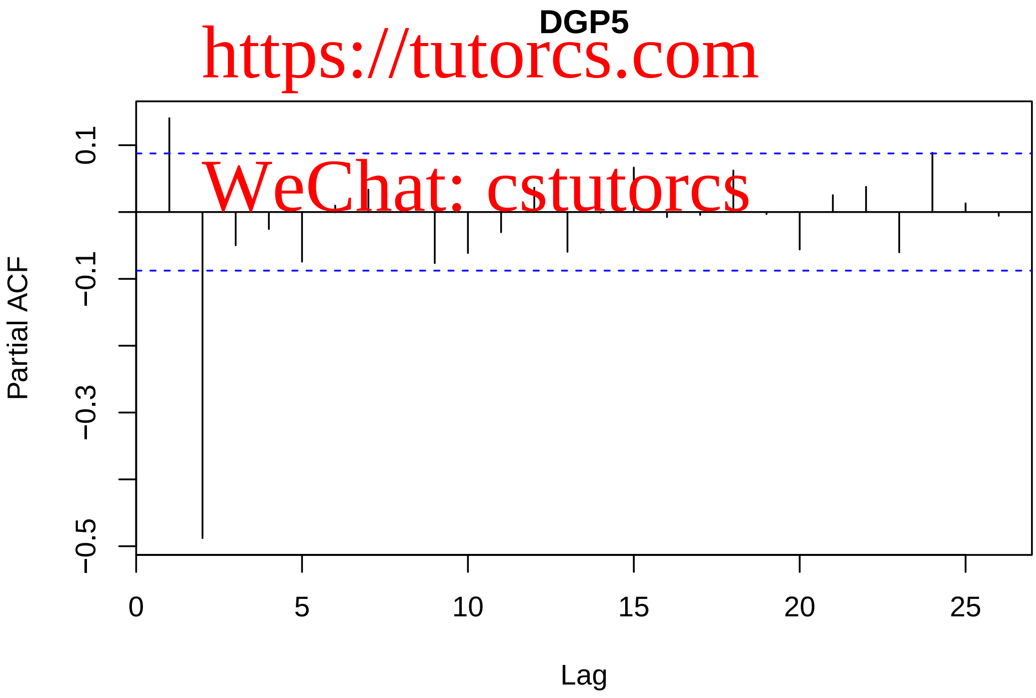

DGP4
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
Lag

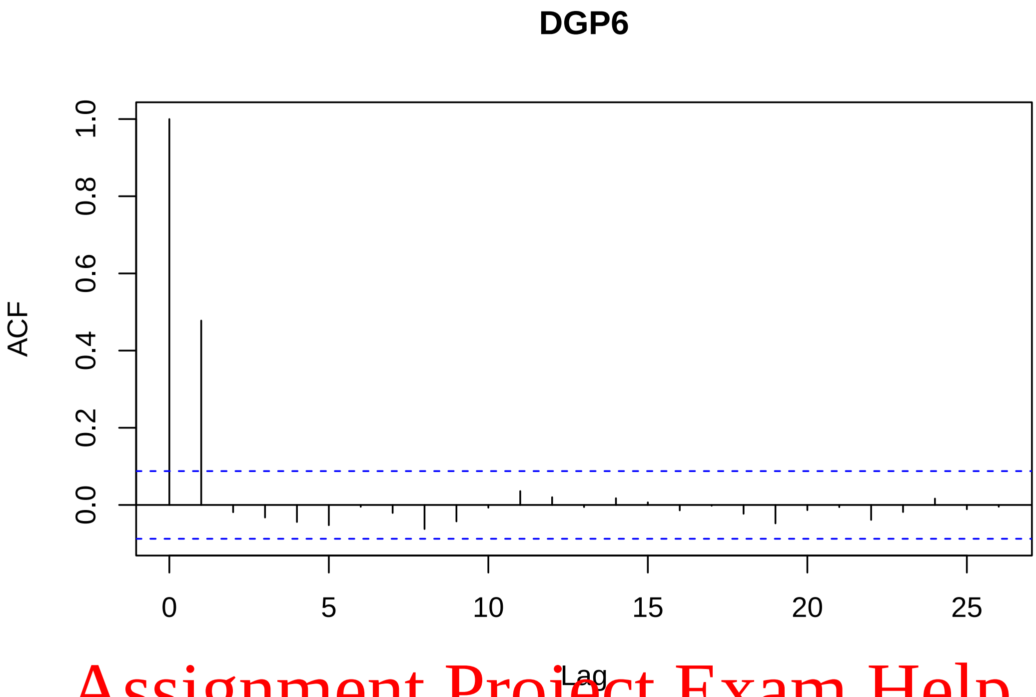



DGP4
Partial ACF
0.6
0.4
0.2
0.0
0
5
10
15
20
25
Lag

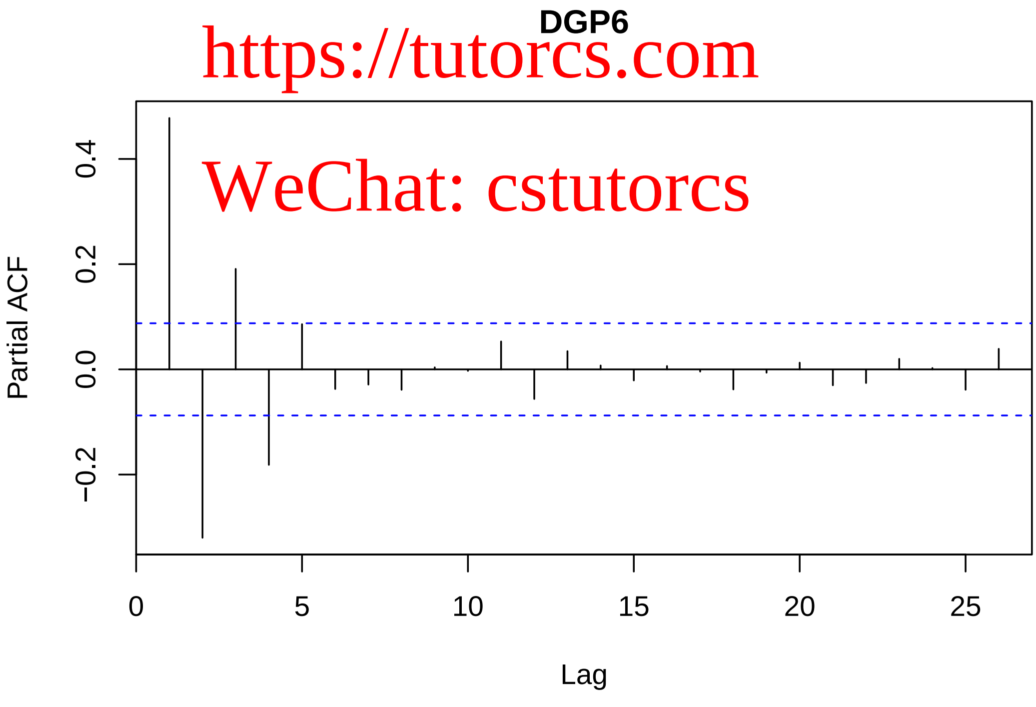

DGP5
ACF
1.0
0.5
0.0
−0.5
0
5
10
15
20
25
Lag

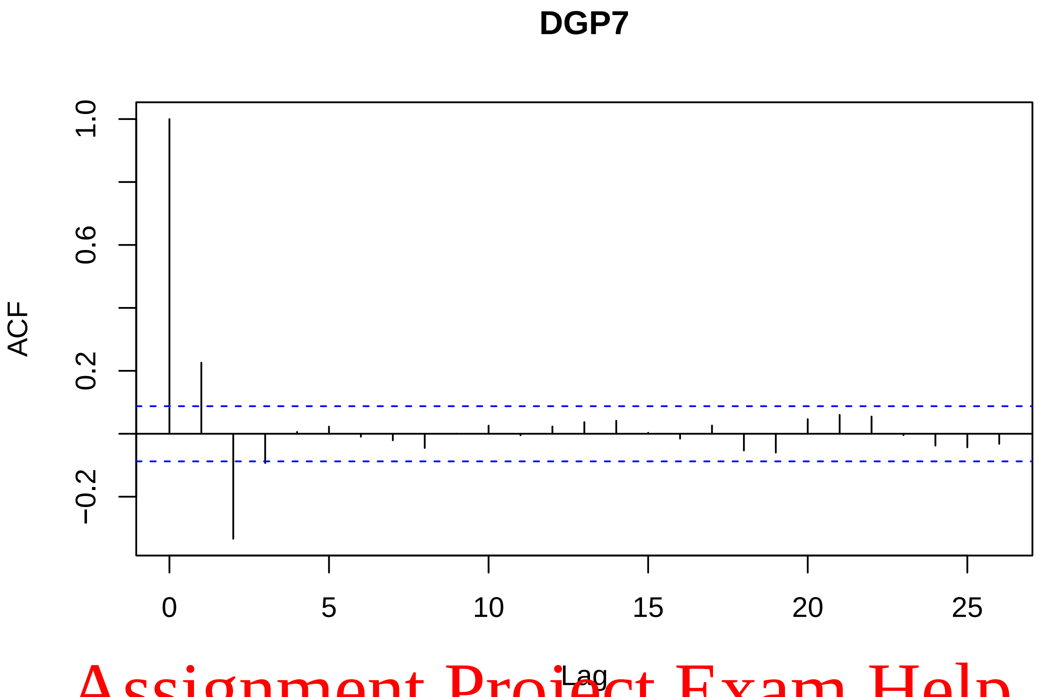



DGP5
Partial ACF
0.1
−0.1
−0.3
−0.5
0
5
10
15
20
25
Lag

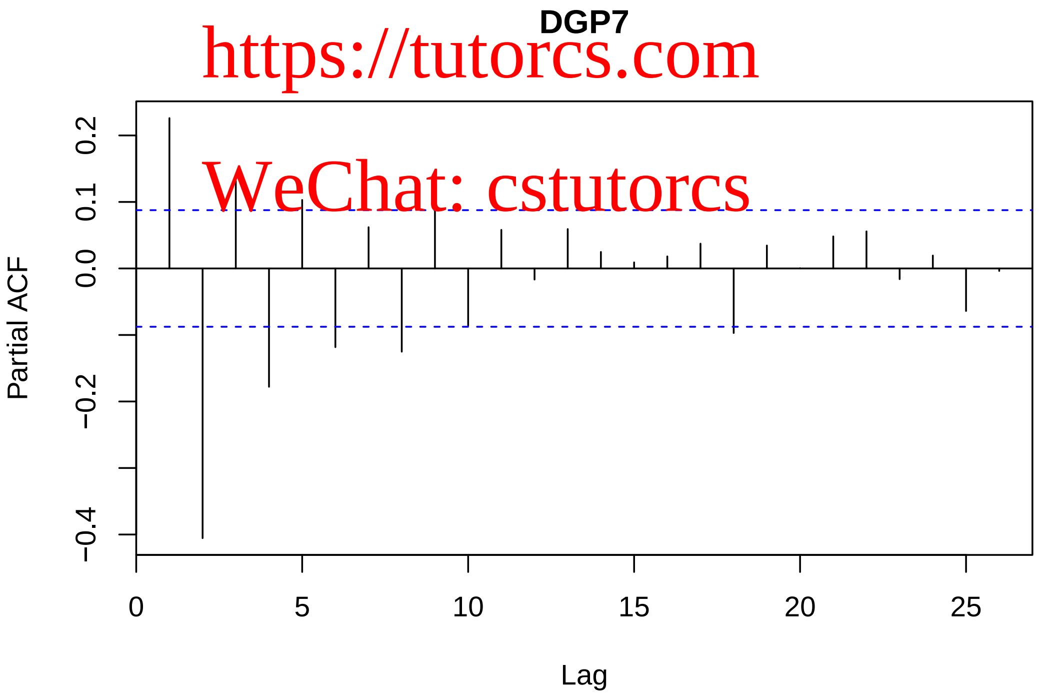

DGP6
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
Lag



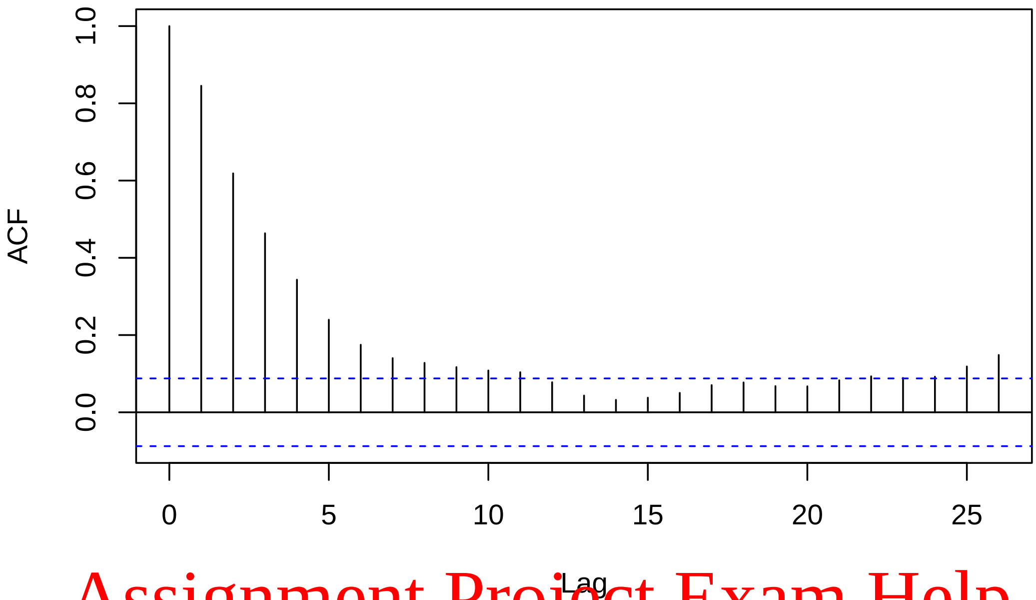

DGP6
Partial ACF
0.4
0.2
0.0
−0.2
0
5
10
15
20
25
Lag


DGP7
ACF
1.0
0.6
0.2
−0.2
0
5
10
15
20
25
Lag



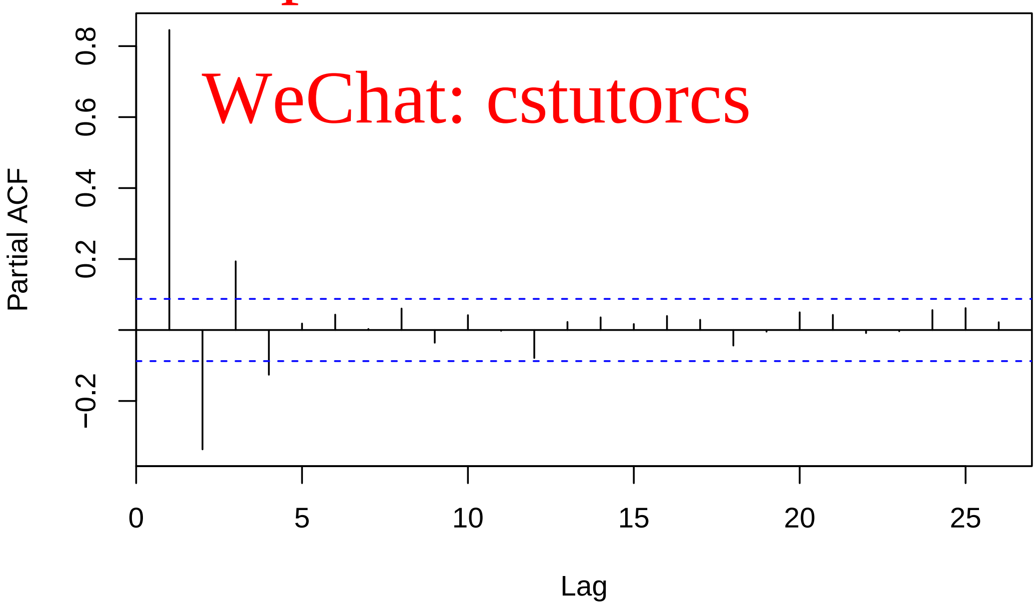

DGP7
Partial ACF
0.2
0.1
0.0
−0.2
−0.4
0
5
10
15
20
25
Lag

DGP8


ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
Lag



DGP8


Partial ACF
0.8
0.6
0.4
0.2
−0.2
0
5
10
15
20
25
Lag

Interpretation:

DGP1
- ACF: Decays geometrically as parameter is positive.
- PACF: One non-zero peak.

DGP2
- ACF: Decays in a dampened oscillatory path as parameter is negative.
- PACF: One non-zero peak.

DGP3
- ACF: Decays geometrically but slower than DGP1.
- PACF: One non-zero peak.

DGP4
- ACF: Decays geometrically as parameter is positive.
- PACF: Two non-zero peaks.

DGP5
- ACF: Decays in an oscillatory path as one parameter is negative (and large in absolute value).
- PACF: Two non-zero peaks.

DGP6
- ACF: One non-zero peak.
- PACF: Decays in an oscillatory path.

DGP7
- ACF: Two non-zero peak.
- PACF: Decays in an oscillatory path.

DGP8
- ACF: Decays geometrically from $k = 2$ onwards as the AR(1) component dominates.
- PACF: Decays in an oscillatory path from $k = 2$ as the MA(1) component dominates.

---