HW4

Charles Dotson

October 12, 2022

[2]:
```python
'''Importing Packages'''
import pandas as pd
import numpy as np
import seaborn as sb
import matplotlib.pyplot as plt
from matplotlib import dates
from IPython.display import Markdown as md
import statsmodels.tsa.stattools as ts
import statsmodels.api as sm
import datetime
from statsmodels.api import stats as sm
from loess import loess_1d
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from openpyxl import Workbook, load_workbook
from sklearn import linear_model
from statsmodels.tsa.ar_model import AutoReg, ar_select_order
from scipy.linalg import toeplitz
import math
import scipy.stats as stats
from statsmodels.tsa.api import ExponentialSmoothing, SimpleExpSmoothing, Holt
from statsmodels.tsa.exponential_smoothing.ets import ETSModel
import calendar
from statsmodels.tsa.seasonal import seasonal_decompose
%matplotlib inline
import random
```

1 Prove equation (5.14) in the textbook for MA(1) process.

For a MA(1) process of the form $y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$, where $\varepsilon_t \overset{iid}{\sim} N(0,1)$ we can see the following:

The expectation of the process:

$$\mathbb{E}[y_t] = \mathbb{E}[\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}]$$

$$= \mathbb{E}[\mu] + \mathbb{E}[\varepsilon_t] - \mathbb{E}[\theta_1 \varepsilon_{t-1}]$$

$$= \mu + \mathbb{E}[\varepsilon_t] - \theta_1 \mathbb{E}[\varepsilon_{t-1}]$$

$$= \mu$$

The variance of the process:

$$\mathrm{Var}(y_t) = \mathbb{E}[y_t^2] - \mathbb{E}[y_t]^2$$

$$= \mathbb{E}[(\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1})^2] - \mathbb{E}[\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}]^2$$

$$= \mathbb{E}[\mu^2 + 2\mu\varepsilon_t - 2\mu\theta_1\varepsilon_{t-1} - 2\theta_1\varepsilon_t\varepsilon_{t-1} + \varepsilon_t^2 + \theta_1^2\varepsilon_{t-1}^2] - \mu^2$$

$$= \mathbb{E}[\mu^2] + \mathbb{E}[2\mu\varepsilon_t] - \mathbb{E}[2\mu\theta_1\varepsilon_{t-1}] - \mathbb{E}[2\theta_1\varepsilon_t\varepsilon_{t-1}] + \mathbb{E}[\varepsilon_t^2] + \mathbb{E}[\theta_1^2\varepsilon_{t-1}^2] - \mu^2$$

$$= \mu^2 + \mathbb{E}[\varepsilon_t^2] + \mathbb{E}[\theta_1^2\varepsilon_{t-1}^2] - \mu^2$$

$$= 1 + \theta_1^2$$

The covariance of the process:

$$\mathrm{Cov}(y_t, y_{t+k}) = \mathbb{E}[y_t y_{t+k}] - \mathbb{E}[y_t]\mathbb{E}[y_{t+k}]$$

$$= \mathbb{E}[(\mu + \varepsilon_t - \theta_1\varepsilon_{t-1})(\mu + \varepsilon_{t+k} - \theta_1\varepsilon_{t+k-1})] - \mu^2$$

$$= \mathbb{E}[\mu^2 + \mu\varepsilon_{t+k} - \mu\theta_1\varepsilon_{t+k-1} + \mu\varepsilon_t + \varepsilon_t\varepsilon_{t+k} - \theta_1\varepsilon_t\varepsilon_{t+k-1} - \mu\theta_1\varepsilon_{t-1} - \theta_1\varepsilon_{t-1}\varepsilon_{t+k} + \theta_1^2\varepsilon_{t-1}\varepsilon_{t+k-1}] - \mu^2$$

After seperating expectations since it is a linear operator, seperating expectations of iid $N(0,1)$ variables, and using the fact that $\mathbb{E}[N(0,1)] = 0$, we arrive at,

$$= \mathbb{E}[\varepsilon_t \varepsilon_{t+k}] - \theta_1 \mathbb{E}[\varepsilon_t \varepsilon_{t+k-1}] - \theta_1 \mathbb{E}[\varepsilon_{t-1} \varepsilon_{t+k}] + \theta_1^2 \mathbb{E}[\varepsilon_{t-1} \varepsilon_{t+k-1}]$$

Since this is an MA(1) process, we then set $k = 1$ which yields,

$$= \mathbb{E}[\varepsilon_t \varepsilon_{t+1}] - \theta_1 \mathbb{E}[\varepsilon_t \varepsilon_t] - \theta_1 \mathbb{E}[\varepsilon_{t-1} \varepsilon_{t+1}] + \theta_1^2 \mathbb{E}[\varepsilon_{t-1} \varepsilon_t]$$

Using independence we arrive at,

$$= \mathbb{E}[\varepsilon_t]\mathbb{E}[\varepsilon_{t+1}] - \theta_1 \mathbb{E}[\varepsilon_t^2] - \theta_1 \mathbb{E}[\varepsilon_{t-1}]\mathbb{E}[\varepsilon_{t+1}] + \theta_1^2 \mathbb{E}[\varepsilon_{t-1}]\mathbb{E}[\varepsilon_t]$$

Since $\mathbb{E}[N(0,1)^2] = 1$ and $\mathbb{E}[N(0,1)] = 0$, we arrive at the solution,

$$\mathrm{Cov}(y_t, y_{t+k}) = -\theta_1$$

We then set our autocovariance function $\gamma_y(k)$ as:

$$\gamma_y(0) = \mathrm{Var}(y_t) = 1 + \theta_1^2$$

$$\gamma_y(1) = \mathrm{Cov}(y_t, y_{t+1}) = -\theta_1$$

Then our autocorrelation function, $\rho_y(k)$, is

$$\rho_y(1) = \frac{\gamma_y(1)}{\gamma_y(0)} = \frac{-\theta_1}{1 + \theta_1^2}$$

Since θ_q in a MA(q) process is a weight, we can assum that all $\theta_q \in [-1, 1]$. Thus we say $f(\theta_1) = \rho_y(1)$. We then choose -1, 0, and 1 for values to test. We see the following.

$$f(-1) = \frac{1}{1+1} = \frac{1}{2}$$

$$f(0) = \frac{0}{1} = 0$$

$$f(1) = \frac{-1}{1+1} = -\frac{1}{2}$$

To now see where the global maximum and minimum are, we find the same values for $f'(\theta_1)$

$$f'(\theta_1) = \frac{\theta_1^2 - 1}{(1 + \theta_1^2)^2}$$

$$f'(-1) = \frac{0}{(1+1)^2} = 0$$

$$f'(0) = \frac{-1}{1} = -1$$

$$f'(1) = \frac{0}{(1+1)^2} = 0$$

Thus we can clearly see there is a global maximum at -1 and global minimum as 1 since $f'(\theta_1)$ is zero at these values. Since these values are the additive inverse of each other, the following holds with two global maximum of $\frac{1}{2}$ when $\theta_1 = -1, 1$ and new global minimum of 0 when $\theta_1 = 0$. Thus,

$$|\rho_y(1)| = \frac{|\theta_1|}{1 + \theta_1^2} \leq \frac{1}{2}$$

2 Prove for AR(2) process that:

$$\phi_1 + \phi_2 < 1$$
$$\phi_2 - \phi_1 < 1$$
$$|\phi_2| < 1$$

We start by showing the AR(2) model $y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$

We start by rewriting the equation in the infinite MA form as

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

$$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} = \delta + \varepsilon_t$$

$$y_t(1 - \phi_1 B - \phi_2 B^2) = \delta + \varepsilon_t$$

$$y_t \Phi(B) = \delta + \varepsilon_t$$

$$y_t = \Phi(B)^{-1}\delta + \Phi(B)^{-1}\varepsilon_t$$

We then set $\Phi(B)^{-1}\delta = \mu$ and $\Phi(B)^{-1} = \Psi(B)$. Thus,

$$y_t = \mu + \Psi(B)\varepsilon_t$$

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i B^i \varepsilon_t$$

To Get the weights ϕ_1, ϕ_2, we start by setting $\Phi(B)\Psi(B) = 1$. Thus,

$$\Phi(B)\Psi(B) = 1$$

$$(1 - \phi_1 B - \phi_2 B^2)(\psi_0 + \psi_1 B + \psi_2 B^2 + \dots) = 1$$

$$\psi_0 + (\psi_1 - \phi_1\psi_0)B + (\psi_2 - \phi_1\psi_1 - \phi_2\psi_0)B^2 + \cdots + (\psi_j - \phi_1\psi_{j-1} - \phi_2\psi_{j-2})B^j = 1$$

Since we can see from the following for this to be true, we must have the following conditions met,

$$\psi_0 = 1$$

$$(\psi_1 - \phi_1\psi_0) = 0$$

$$(\psi_j - \phi_1\psi_{j-1} - \phi_2\psi_{j-2}) = 0 \quad \text{for all } j = 2, 3 \ldots$$

Since there are infinite solutions to this problem. We notice that the ψ_j takes the form of second-order linear difference equation and thus, the roots to $m^2 - \phi_1 m - \phi_2$ will give us our results. Thus is the roots m_1, m_2 obtained from:

$$m_1, m_2 = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

and $|m_1|, |m_2| < 1$, then we will have $\sum_{i=0}^{+\infty} |\psi_i| < \infty$. We now obtain the parameters for ϕ_1, ϕ_2 for stationarity. We set the interval based on the absolute value constraint.

$$-1 < \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2} < 1$$

$$-2 < \phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2} < 2$$

We now notice that the larger of the ϕ will be bounded by and thus

$$\phi_1 + \sqrt{\phi_1^2 + 4\phi_2} < 2$$

$$\sqrt{\phi_1^2 + 4\phi_2} < 2 - \phi_1$$

$$\phi_1^2 + 4\phi_2 < 4 - 4\phi_1 + \phi_1^2$$

$$4\phi_2 < 4 - 4\phi_1$$

$$\phi_2 < 1 - \phi_1 \rightarrow \phi_1 + \phi_2 < 1$$

For the inverse and thus smaller of the ϕ we find.

$$\phi_1 - \sqrt{\phi_1^2 + 4\phi_2} > -2$$

$$-\sqrt{\phi_1^2 + 4\phi_2} > -2 - \phi_1$$

$$\phi_1^2 + 4\phi_2 < 4 + 4\phi_1 + \phi_1^2$$

$$4\phi_2 < 4 + 4\phi_1$$

$$\phi_2 < 1 + \phi_1 \rightarrow \phi_2 - \phi_1 < 1$$

Now we can add the two prior conditions $\phi_2 - \phi_1 < 1$ and $\phi_2 + \phi_1 < 1$ and see that we also get the condition $\phi_2 < 1$.

Now when we take the complex roots. We will start by finding the conjugate where our complex roots are in the form $a \pm bi$. Thus we set the following.

$$a = \frac{1}{2}\phi_1 \qquad b = \frac{1}{2}\sqrt{-(\phi_1^2 + 4\phi_2)}$$

Thus we get,

$$|m| = \sqrt{a^2 + b^2}$$

$$= \sqrt{\frac{1}{4}\phi_1^2 - \frac{1}{4}(\phi_1^2 + 4\phi_2)}$$

$$= \sqrt{-\phi_2}$$

Since the conjugate of a complex number is a distance, $|m|$ must be a real number. This will only occur when $-1 < \phi_2 < 0$. Combing two constraints $\phi_2 < 1$ and $-1 < \phi_2 < 0$. we can see that $-1 < \phi_2 < 1$ which is equivalent to saying $|\phi_2| < 1$. Thus we show that the following constraints for a stationary AR(2) process are:

$$\phi_2 + \phi_1 < 1$$

$$\phi_2 - \phi_1 < 1$$

$$|\phi_2| < 1$$

3 Consider the time series model: $y_t = 20 + \epsilon_t + 0.2\epsilon_{t-1}$

Excercise 5.7 in the book. Problems (a), (c), (d), and (e). Where problem (e) is augmented by professor

(a) Is this a stationary time series process?

This is a stationary timeseries. We can clearly see that the process is an MA(1) process of the form, $y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$. Because this, we know that all MA(1) time series processes are stationary.

(c) What is the mean of the time series?

Building off the prior question, since the process is of the form $y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$, we can see that the mean of the process is $\mu = 20$.

(d) If the current observation is $y_{100} = 23$, would you expect the next observation to be above or below the mean? Explain your answer.

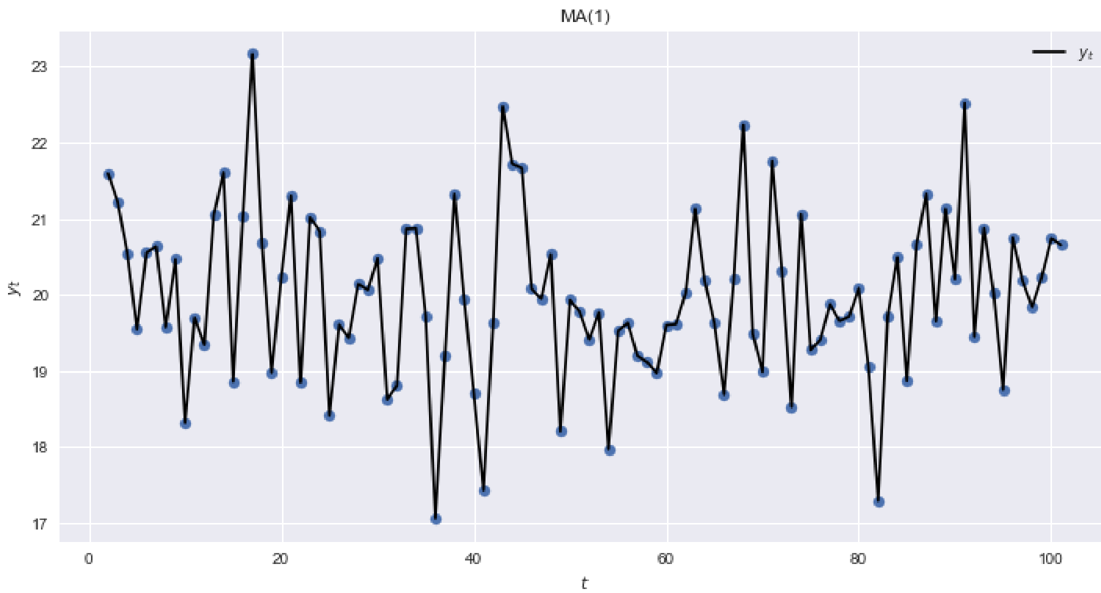
We can see that $\rho_y(1) = 0.1923$ which shows a postive auto correlation and thus we can expect y_{101} to be above the mean.

(e) Generate a series of length 100 with using your LIN as the random seed, plot the time-series plot and ACF plot.

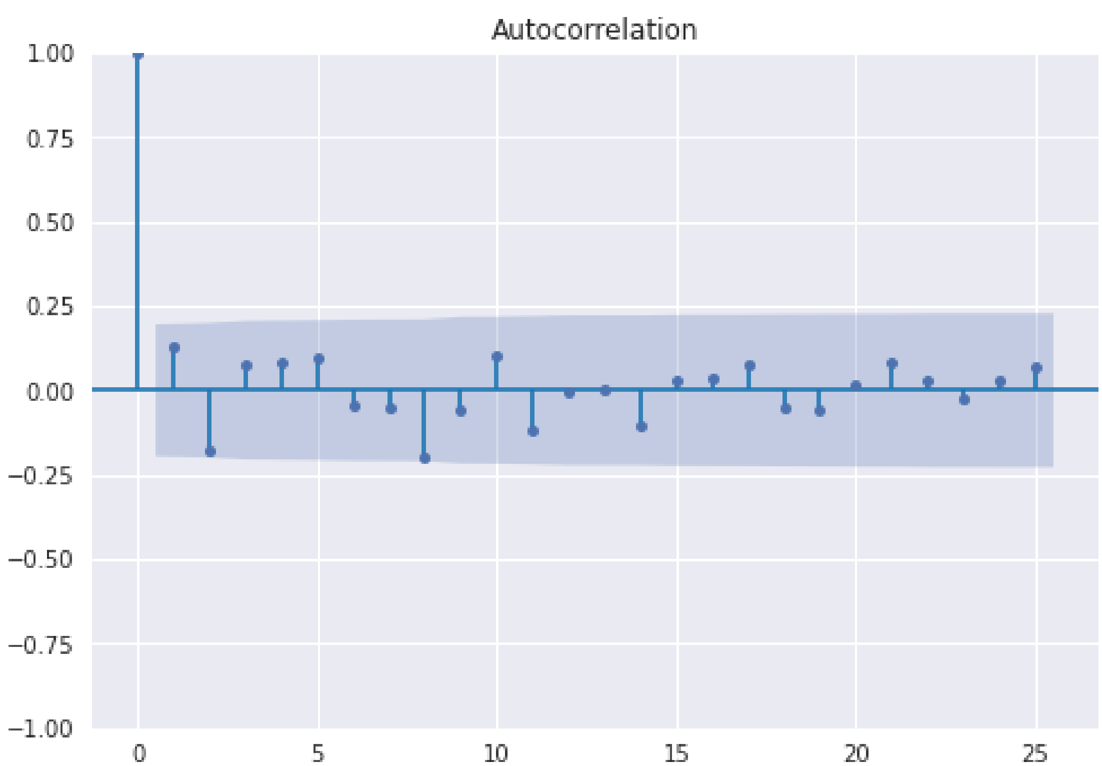
[3]:
```
random.seed(867002105)
```

[4]:
```
'''Generating the series'''
mean = 0
std = 1
samps = 101
mu = 20
theta_1 = -0.2

epsilons = list(np.random.normal(mean, std, size = samps))
y_t = [mu + epsilons[i] - theta_1*epsilons[i-1] for i in range(1,␣
 ↪len(epsilons))]
yt_df = pd.DataFrame(
    y_t,
    index = [i for i in range(2, 102)],
    columns=['y_t']
)

with plt.style.context('seaborn'):
    fig = plt.figure(figsize=(12,6))
    ax = plt.axes()
    plt.plot(yt_df, c = 'Black', label = '$y_t$')
    plt.scatter(yt_df.index, yt_df['y_t'])
    ax.set_xlabel('$t$')
    ax.set_ylabel('$y_t$')
    plt.legend()
```

```
    plt.title('MA(1)')
    plt.show()
```



[5]:
```
with plt.style.context('seaborn'):
    fig = plot_acf(yt_df, lags=25)
```



4 Consider the time series model

$$y_t = 200 + 0.7y_{t-1} + \epsilon_t$$

Excercise 5.4 in the book. Problem (e) where (e) is augmented by professor

(a) Is This a Stationary Process?

It is sationary because $\phi = 0.7$. Thus, holding true that it is also causal.

(b) What is the Mean of the Process?

$\mu = 666.67$

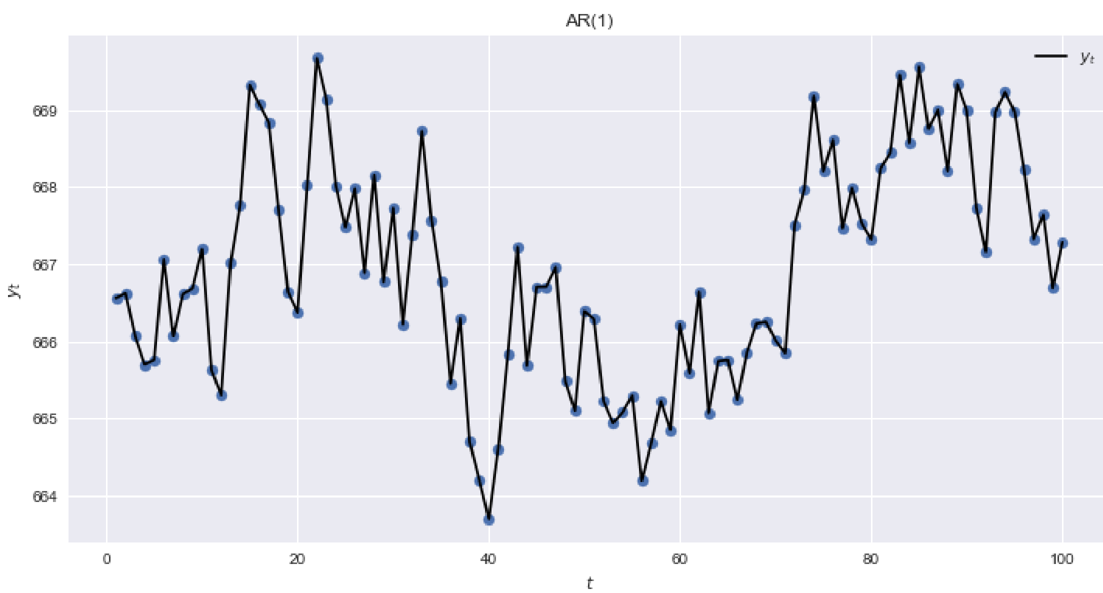
(c) If the current observation is $y_{100} = 750$, would you expect the next observation to be above or below the mean?

$\rho_y(1) = \phi = 0.7$, Thus we have a postive autocorrelation so we can expect for the y_{101} to be above the mean.

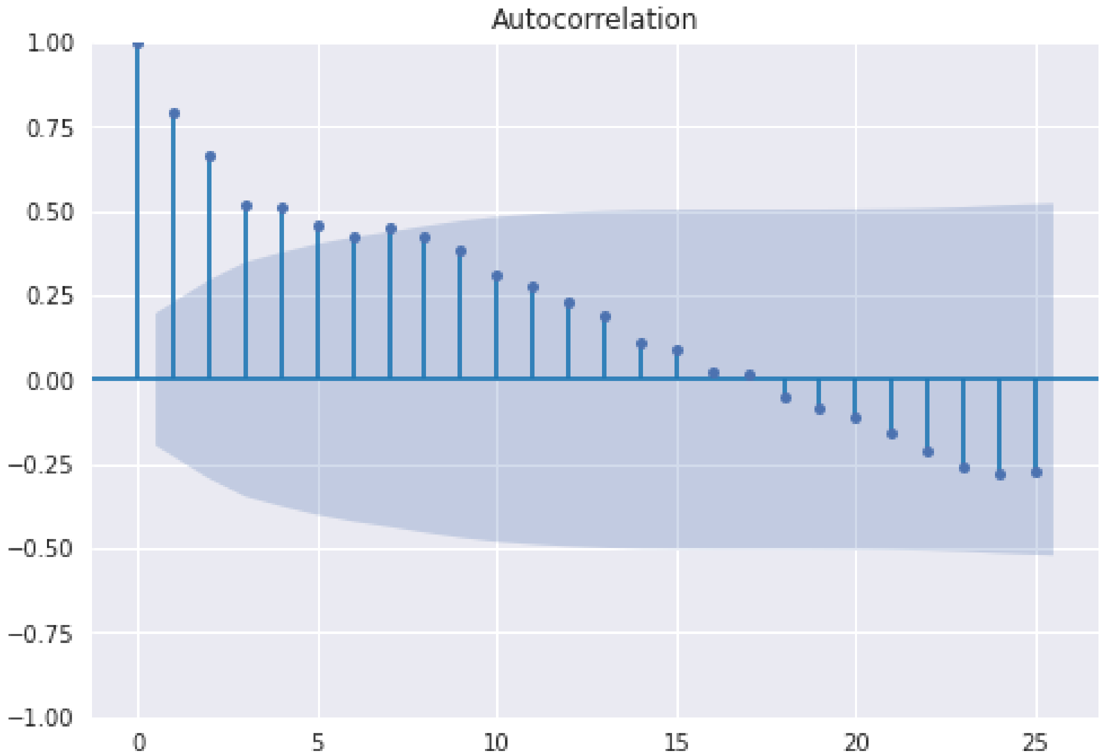
(d) Generate a series of length 100 with using your LIN as the random seed, plot the time-series plot, ACF plot and PACF plot (use type="partial" in acf function).

[20]:
```python
'''Generating the series'''
mean = 0
std = 1
samps = 101

delta = 200
phi = 0.7
mu = delta/(1-phi)


epsilons = list(np.random.normal(mean, std, size = samps))
y_t = []

for i in range(len(epsilons)):
    if i == 0:
        y_t_ent = delta + phi*mu+epsilons[i]
        #y_t_ent = mu
        y_t.append(y_t_ent)
    else:
        y_t_ent = delta + phi*y_t[i-1]+epsilons[i]
        y_t.append(y_t_ent)

yt_df = pd.DataFrame(
    y_t[:-1],
    index = [i for i in range(1, 101)],
```

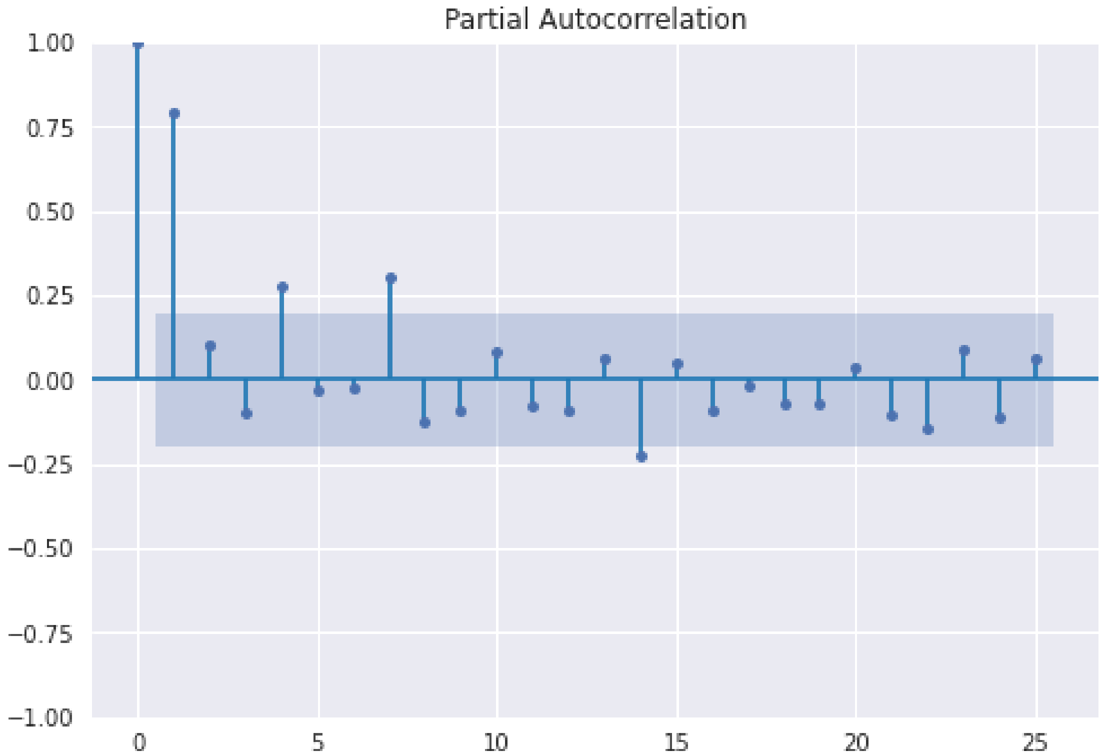
```
    columns=['y_t']
)

with plt.style.context('seaborn'):
    fig = plt.figure(figsize=(12,6))
    ax = plt.axes()
    plt.plot(yt_df, c = 'Black', label = '$y_t$')
    plt.scatter(yt_df.index, yt_df['y_t'])
    ax.set_xlabel('$t$')
    ax.set_ylabel('$y_t$')
    plt.legend()
    plt.title('AR(1)')
    plt.show()
```



[24]:
```
with plt.style.context('seaborn'):
    fig = plot_acf(yt_df.values.tolist(), lags=25)
```



[26]:
```
with plt.style.context('seaborn'):
    fig = plot_pacf(yt_df.values.tolist(), lags=25, method='ywm')
```



Partial Autocorrelation
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
5
10
15
20
25