Assignment 1 (Econometrics)

Name : Biswajit Rana
Roll no: B2330026

A suitable regression line of expenditure on income has to be fitted and the following assumptions for the fitted line has to be checked:
1. E(u) = 0
2. Var(u) = 0
3. Cov(u_i. u_j) = 0

Code:

```
# Importing Libraries
import pandas as pd
import numpy as np
import statsmodels.api as sm
import matplotlib.pyplot as plt
import seaborn as sns
from statsmodels.stats.diagnostic import het_breuschpagan

# Reading the data
data = pd.read_csv('/mnt/data/income_exp_data.csv')

# Fitting a regression line
X = data['Income']
y = data['Expenditure']
X = sm.add_constant(X)
model = sm.OLS(y, X).fit()
print(model.summary())

# Plotting the regression line on the data
plt.figure(figsize=(10, 6))
sns.regplot(x='Income', y='Expenditure', data=data, ci=None,
line_kws={"color": "red"})
plt.title('Regression Line of Expenditure on Income')
```

```
plt.xlabel('Income (in thousands)')
plt.ylabel('Expenditure (in thousands)')
plt.show()

# Finding the residuals
residuals = model.resid

# Checking First Assumption is true or not
mean_residuals = np.mean(residuals)
print(f'Mean of Residuals (E(u)): {mean_residuals}')

# Checking Second Assumption is true or not
_, pval, __, f_pval = het_breuschpagan(residuals, model.model.exog)
print(f'P-value for Breusch-Pagan test (Var(u) = 0 assumption): {pval}')

# Checking Third Assumption is true or not
dw = sm.stats.stattools.durbin_watson(residuals)
print(f'Durbin-Watson statistic for autocorrelation (Cov(u_i, u_j) = 0
assumption): {dw}')
```

Output:

```
Mean of Residuals (E(u)): 2.8848035071860065e-14
P-value for Breusch-Pagan test (Var(u) = 0 assumption):
3.3464863589453733e-16
Durbin-Watson statistic for autocorrelation (Cov(u_i, u_j) = 0 assumption):
2.167365723538409
```

Conclusion:

1. We can see the mean of teh residuals are very much closer to 0. So, first assumption is met. So, E(u)=0.

2. Using Breusch-Pagan test to check for heteroscedasticity, as p value is very small we also can assure that heteroscedasticity holds which means var(u) = 0.

3. Using Durbin-Watson statistic for checking the Correlation. As the value of the statistic is closer to 2 so it indicates no autocorrelation. So, $Cov(u_i. u_j) = 0$.