Heteroscedasticity Analysis

Dataset Background:

This dataset contains information on the Annual Expenditure on Food by households. The goal is to predict the Annual Expenditure on Food based on the Annual Income of households.

Dataset Glimpse:

Expenditure	Income
9.46	25.83
10.56	34.31
14.81	42.5
21.71	46.75
22.79	48.29
18.19	48.77
22	49.65
18.12	51.94
23.13	54.33
19	54.87

Total Number of Rows: 40

Total Number of Columns: 2

Column Details:

- Expenditure: Annual Expenditure on Food by households (in thousands of dollars).
- Income: Annual Income of households (in thousands of dollars).

Main Dependent Variable: Expenditure (in thousands of dollars).

Using SPSS Software EViews, we have analysed the data:

Descriptive Statistics:

	EXPENDIT...	INCOME
Mean	23.59450	69.80000
Median	21.85500	71.23000
Maximum	48.71000	115.4600
Minimum	9.460000	25.83000
Std. Dev.	8.176025	19.82269
Skewness	1.034804	0.218662
Kurtosis	4.129972	2.985860
Jarque-Bera	9.266858	0.319087
Probability	0.009721	0.852533
Sum	943.7800	2792.000
Sum Sq. Dev.	2607.048	15324.63
Observations	40	40

Inferences:

- The variable expenditure is right skewed, ranging between 9.46 to 48.71 thousand dollars.
- The variable income is slightly right skewed, ranging between 25.83 to 115.46 thousand dollars.
- There is no missing data.

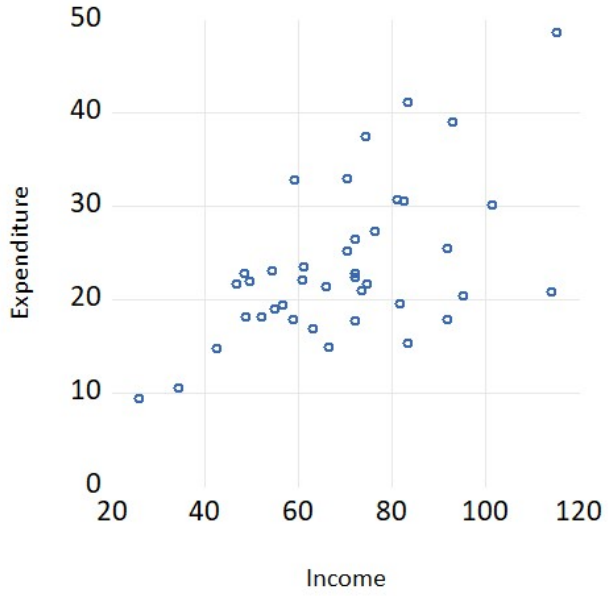
Correlation Analysis:

	EXPENDIT...	INCOME
EXPE...	1.000000	0.563096
INCOME	0.563096	1.000000

Inferences:

- Variables expenditure and income have only the possibility of positive linear correlation, having correlation coefficient 0.56.

Scatter Plot:



Inference: the variables income and expenditure have only the possibility of positive linear correlation.

Regression Model:

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.383218	4.008356	1.841956	0.0733
INCOME	0.232253	0.055293	4.200378	0.0002

R-squared	0.317077	Mean dependent var	23.59450
Adjusted R-squared	0.299105	S.D. dependent var	8.176025
S.E. of regression	6.844922	Akaike info criterion	6.733598
Sum squared resid	1780.413	Schwarz criterion	6.818042
Log likelihood	-132.6720	Hannan-Quinn criter.	6.764130
F-statistic	17.64318	Durbin-Watson stat	2.370272
Prob(F-statistic)	0.000155		

From the above model, we get the estimate equation:

Expenditure = 7.382 + (0.232)(Income)

Inferences:

- The model has low explanatory power, having R^2 value 0.32.
- The independent variable income has p-value 0.0002 and is statistically significant.
- The intercept has a p-value of 0.0733 and is statistically insignificant.
- The Std. Error of the slope is very low, having value of 0.055. Hence, the variance value will also be very low.

From the inferences made using the low value of R^2, low value of Std. Error of slope and statistically significant p-value of the independent variable income, we may infer the presence of heteroscedasticity. To confirm the same, we need to perform the white test.

White Test Results:

Heteroskedasticity Test: White
Null hypothesis: Homoskedasticity

F-statistic	10.61267	Prob. F(2,37)	0.0002
Obs*R-squared	14.58151	Prob. Chi-Square(2)	0.0001
Scaled explained SS	12.33032	Prob. Chi-Square(2)	0.0021

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	63.10781	77.54055	0.813869	0.4209
INCOME^2	0.028786	0.014972	1.922686	0.0622
INCOME	-2.433686	2.192411	-1.110050	0.2741

R-squared	0.364538	Mean dependent var	44.51031
Adjusted R-squared	0.330188	S.D. dependent var	61.70718
S.E. of regression	50.50240	Akaike info criterion	10.75396
Sum squared resid	94368.20	Schwarz criterion	10.88062
Log likelihood	-212.0791	Hannan-Quinn criter.	10.79976
F-statistic	10.61267	Durbin-Watson stat	2.009572
Prob(F-statistic)	0.000228		

Inference:

- Since the probability value of chi-square is less than the f-critical value, we reject the null hypothesis and confirm the presence of heteroscedasticity.