Problems on the two-variable econometric model

1. What is meant by and what is the function of (a) Simple regression analysis? (b) Linear regression analysis? (c) a scatter diagram? (d) An error term?

2. The data in the given Table (1) reports the aggregate consumption (Y in billions of US dollars) and disposable income (X, also in billions of US dollars) for a developing economy for the 12 years from 1988 to 1999. Draw a scatter diagram for the data and determine by inspection if there exists an approximate linear relationship between Y and X.

Table 1: Aggregate Consumption (Y) and Disposable Income (X)

Year	n	Y_i	X_i
1988	1	102	114
1989	2	106	118
1990	3	108	126
1991	4	110	130
1992	5	122	136
1993	6	124	140
1994	7	128	148
1995	8	130	156
1996	9	142	160
1997	10	148	164
1998	11	150	170
1999	12	154	178

3. State the general relationship between consumption Y and income X in (a) exact linear form and (b) stochastic form. (c) Why would you expect most observed values of Y not to fall exactly on a straight line?

4. State five assumptions of the classical regression model (OLS) and give an intuitive explanation of the meaning and the need for each.

5. What is meant by Ordinary Least Squares (OLS) method of estimating the *best* straight line that fits the sample of XY observations? Why do we take vertical deviations? Why do we not take the sum of the deviations without squaring them? Why do we not take the sum of the absolute deviations?

6. Starting from the minimization of the sum of the squared deviations or residuals, (a) derive the normal equations, (b) derive the formulae to find $\hat{\beta_1}$ and $\hat{\beta_2}$.

7. State the differences between (a) β_1 and β_2, (b) $\hat{\beta_1}$ and $\hat{\beta_2}$, (c) error term and residuals.

8. Write the equations for (a) true and estimated relationships between X and Y (b) for the true and estimated regression lines between X and Y.

9. Find the regression equation for the consumption schedule in Table (1) and also plot the regression line and show the deviations of each Y_i from the corresponding $\hat{Y_i}$.

10. In the context of the estimated regression based on Table (1), what is the meaning of
(a) estimator $\hat{\beta_1}$ and (b) estimator $\hat{\beta_2}$

11. Define the following: σ^2, $\sigma^{\wedge 2}$, var $(\beta_1\wedge)$, var $(\beta_2\wedge)$, se $((\beta_2\wedge))$, se $((\beta_1\wedge))$

12. Prove that

Mean $\hat{\beta_1} = \beta_1$

var $(\beta_1\wedge) = \dfrac{\sigma^2 \Sigma X_i^2}{n \, \Sigma x_i^2}$

Mean $\hat{\beta_2} = \beta_2$

var $(\beta_2\wedge) = \dfrac{\sigma^2}{\Sigma x_i^2}$

13. Find the following on the basis of data of Table (1),
  $\sigma^{\wedge 2}$, var $(\beta_1\wedge)$, var $(\beta_2\wedge)$, se $((\beta_2\wedge))$, se $((\beta_1\wedge))$

14. State the null and alternative hypothesis to test the statistical significance of the parameters of the regression equation. What is the form of sampling distribution of $\hat{\beta_1}$ and $\hat{\beta_2}$? Which distribution must we use to test the statistical significance of β_1 and β_2? What are the degrees of freedom? [Hint: t = 2.228 at the 5% level (two-tail test) with 10 df]

15. Construct the 95% confidence interval for β_1 and β_2.

16. Derive the formula for R^2.

17. What does R^2 mean? What is it used to measure? What is its range of values?

18. What does the correlation coefficient measure? What is its range of values?

19. What is the relationship between correlation and regression analysis?

20. Derive the equation of r from R^2.

N.B. All the questions have been taken from the following:

References:

Salvatore, D., & Reagle, D. (2011). *Schaum's outline of statistics and econometrics* (2nd ed.).

   McGraw-Hill Education.