Name ______________ **Surname** ______________ **Student ID.** ______________

All questions are related to the simple (two-variable) linear regression model for which the population regression equation can be written in conventional notation as:

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

where Y_i and X_i are observable variables, β_1 and β_2 are true parameters and u_i is a disturbance term.
The Ordinary Least Squares (OLS) sample regression equation corresponding to the above equation is

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i \qquad (i = 1, 2, ..., N)$$

where $\hat{\beta}_1$ is the OLS estimator of the intercept coefficient β_1, and $\hat{\beta}_2$ is the OLS estimator of the slope coefficient β_2, $\hat{u}_i$ is the OLS residual for the i-th sample observation.

(Green Booklet)

1. **(10 Marks)** When we adopt the ordinary least square criterion (OLS), why do we choose the criterion of minimizing $\sum \hat{u_i}^2$ instead of the criterion of minimizing $\sum \hat{u_i}$? Logically explain with support of a diagram.

(Green Booklet)

2. **(15 Marks)** Show that $\sum x_i^2 = \sum X_i^2 - n\bar{X}^2$ and that $\sum x_i y_i = \sum x_i Y_i$. In other words, show that:

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\sum x_i Y_i}{\sum X_i^2 - n\bar{X}^2}$$

(Green Booklet)

3.**(20 Marks)**
(a). What is the Gauss-Markov Theorem?
(b). Show that the OLS intercept coefficient estimator $\hat{\beta}_1$ is a linear function of the sample observations Y_i and X_i.
(c). Stating explicitly all required assumptions, prove that the OLS intercept coefficient estimator $\hat{\beta}_1$ is an unbiased estimator of the true parameter β_1.

(Yellow Booklet)

4. **(45 Marks)** A researcher is using data for a sample of 16 B.E students to investigate the relationship between the GPA (grade point average) Y_i where the grade point average is based on a four-point scale and the SAT scores X_i. Preliminary analysis of the sample data produces the following sample information:

$$\sum Y_i = 50.90 \qquad \sum X_i = 388.00 \qquad \sum Y_i^2 = 164.51$$

$$\sum X_i^2 = 9620.00 \qquad \sum X_i Y_i = 1254.90 \qquad \sum x_i y_i = 20.58$$

$$\sum y_i^2 = 2.5844 \qquad \sum x_i^2 = 211.00 \qquad \sum \hat{y_i}^2 = 2.0063$$

where $x_i = X_i - \bar{X}$, $y_i = Y_i - \bar{Y}$, and $\hat{y}_i = \hat{Y}_i - \bar{Y}$
Use the above sample information to answer all the following questions. **Show explicitly all formulas and calculations**

(a) **(5 Marks)** Use the above information to compute OLS estimates of the intercept coefficient β_1 and that of the slope coefficient β_2.

(b) **(5 Marks)** Interpret the slope coefficient estimate you calculated in part(a)–i.e.,explain in words what the numeric value you calculated for $\hat{\beta}_2$ means.

(c) **(10 Marks)** Calculate an estimate of σ^2.

(d) **(5 Marks)** Compute the value of r^2. Briefly explain what the calculated value of r^2 means.

(e) **(5 Marks)** Compute the estimated variance of $\hat{\beta}_1$ and the estimated variance of $\hat{\beta}_2$.

(f)**(5 Marks)** Compute the two-sided 90% confidence interval for the intercept coefficient β_1. Briefly explain what the two-sided 90% confidence interval means.

(g)**(5 Marks)** Perform a test of the null hypothesis $H_0 : \beta_1 = 0$ against the alternative hypothesis $H_1 : \beta_1 \neq 0$ at the 1 % significance level (i.e.,for significance level $\alpha = 0.01$). Show how you calculated the test statistic. State the decision rule you use, and the inference you would draw from the test. Briefly explain what the test outcome means.

(h) **(5 Marks)** Perform a test of the null hypothesis $H_0 : \sigma^2 \geq 0.025$ against the alternative hypothesis $H_1 : \sigma^2 < 0.025$ at the 5 % significance level (i.e.,for significance level $\alpha = 0.05$). Show how you calculated the test statistic. State the decision rule you use, and the inference you would draw from the test. Briefly explain what the test outcome means.

——————————————————The End of Exam——————————————————