Instruction:

1) Use 4 decimal places for all numerical answers.
2) Manage your time spending on each question carefully.
3) All relevant statistical table and necessary formula are provided at the back page of this exam paper.
5) 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.

1. (10 points) 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.

2. (15 points)
Let

$$\hat{\beta}_1 = \sum(\frac{1}{n} - \bar{X}k_i)Y_i$$

$$\hat{\beta}_2 = \sum k_i Y_i$$

where $k_i = \frac{x_i}{\sum x_i}$, $x_i = X_i - \bar{X}$
and $Y_i = \beta_1 + \beta_2 X_i + u_i$.
Show that $\hat{\beta}_1$ is an unbiased estimator of β_1.

3. (20 points)
Consider the regression model through the origin (A regression model without the intercept):

$$Y_i = \beta_2 X_i + u_i$$

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

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

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

3.1 (10 points) Apply the ordinary least square (OLS) criterion as you state in Question 1 to obtain the formula for $\hat{\beta}_2$ (Hint: Calculate for the first-order condition, there is one normal equation for this model)

3.2 (10 points) Stating explicitly all required assumptions, prove that the OLS coefficient estimator $\hat{\beta}_2$ you get in 3.1 is an unbiased estimator of the true parameter β_2.

4. **(45 points)** A researcher is using data for a sample of 20 Top Business Schools in the United States to investigate the relationship between GMAT scores X_i and the average starting daily pay in U.S. dollars Y_i. Preliminary analysis of the sample data products the following sample information:

$$\sum Y_i = 4,893 \qquad \sum X_i = 13,666 \qquad \sum Y_i^2 = 1,224,645$$

$$\sum X_i^2 = 9,352,298 \qquad \sum X_iY_i = 3,360,050 \qquad \sum x_iy_i = 16,663.10$$

$$\sum y_i^2 = 27,572.55 \qquad \sum x_i^2 = 14,320.20 \qquad \sum \hat{y_i}^2 = 19,389.0496$$

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**

4.1 **(5 points)** Use the above information to compute OLS estimates of the intercept coefficient β_1 and that of the slope coefficient β_2.

4.2 **(5 points)** 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.

4.3 **(5 points)** Calculate an estimate of σ^2.

4.4 **(5 points)** Compute the value of r^2. Briefly explain what the calculated value of r^2 means.

4.5 **(5 points)** Compute the estimated variance of $\hat{\beta}_1$ and the estimated variance of $\hat{\beta}_2$.

4.6 **(5 points)** Compute the two-sided 99% confidence interval for the σ^2. Briefly explain what the two-sided 99% confidence interval means.

4.7 **(5 points)** Perform a test of the null hypothesis $H_0 : \beta_1 = 0$ against the alternative hypothesis $H_1 : \beta_1 \neq 0$ 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.

4.8 **(10 points)** Perform a test of the null hypothesis $H_0 : \beta_2 = 0$ against the alternative hypothesis $H_1 : \beta_2 \neq 0$ at the 1 % significance level (i.e., for significance level $\alpha = 0.01$). Fill in all information in table1. Show how you calculated the analysis of variance or F test. State the decision rule you use, and the inference you would draw from the test. Briefly explain what the test outcome means.

Table 1. ANOVA Table for the two-variable regression model

Source of variation	Sum of Square **SS**	**df**	Mean Sum of Square **MSS**
Due to regression (ESS)			
Due to residuals (RSS)			
TSS			

—The End of Exam—