QUEEN'S UNIVERSITY AT KINGSTON
Department of Economics

ECONOMICS 351* - Section A

Introductory Econometrics

Fall Term 2002 **MID-TERM EXAM** M.G. Abbott

DATE: **Monday October 28, 2002.**

TIME: **80 minutes; 2:30 p.m. - 3:50 p.m.**

INSTRUCTIONS: The exam consists of **SIX (6)** questions. Students are required to answer **ALL SIX (6)** questions.

Answer all questions in the exam booklets provided. Be sure your ***name*** and ***student number*** are printed clearly on the front of all exam booklets used.

Do not write answers to questions on the front page of the first exam booklet.

Please label clearly each of your answers in the exam booklets with the appropriate number and letter.

Please write legibly.

A table of percentage points of the t-distribution is given on the last page of the exam.

MARKING: The marks for each question are indicated in parentheses immediately above each question. **Total marks** for the exam **equal 100**.

GOOD LUCK!

QUESTIONS: **Answer ALL SIX questions.**

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

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \textbf{(1)}$$

where Y_i and X_i are observable variables, β_1 and β_2 are unknown (constant) regression coefficients, and u_i is an unobservable random error term. The Ordinary Least Squares (OLS) sample regression equation corresponding to regression equation (1) is

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

where $\hat{\beta}_1$ is the OLS estimator of the intercept coefficient β_1, $\hat{\beta}_2$ is the OLS estimator of the slope coefficient β_2, $\hat{u}_i$ is the OLS residual for the i-th sample observation, $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ is the OLS estimated value of Y for the i-th sample observation, and N is sample size (the number of observations in the sample).

(15 marks)

1. State the Ordinary Least Squares (OLS) estimation criterion. State the OLS normal equations. Derive the OLS normal equations from the OLS estimation criterion.

(15 marks)

2. Show that the OLS slope coefficient estimator $\hat{\beta}_2$ is a linear function of the Y_i sample values. Stating explicitly all required assumptions, prove that the OLS slope coefficient estimator $\hat{\beta}_2$ is an unbiased estimator of the slope coefficient β_2.

(10 marks)

3. Answer parts (a) and (b) below.

(a) Write the expression (or formula) for $\text{Var}(\hat{\beta}_2)$, the variance of $\hat{\beta}_2$.

(b) Which of the following factors makes $\text{Var}(\hat{\beta}_2)$ *smaller*?

(1) a smaller value of N, sample size

(2) smaller values of $x_i^2 = (X_i - \overline{X})^2$, i = 1, …, N

(3) a larger value of σ^2, the error variance

(4) a smaller value of σ^2, the error variance

(5) a larger value of N, sample size

(6) larger values of $x_i^2 = (X_i - \overline{X})^2$, i = 1, …, N.

(10 marks)

4. Explain what is meant by each of the following statements about the estimator $\hat{\theta}$ of the population parameter θ.

(a) $\hat{\theta}$ is an unbiased estimator of θ.
(b) $\hat{\theta}$ is an efficient estimator of θ.

(34 marks)

5. A researcher is using data for a sample of 25 business schools that offer MBA degrees to investigate the relationship between the annual salary gain of graduates Y_i (measured in *thousands* of dollars per year) and annual tuition fees X_i (measured in *thousands* of dollars per year). Preliminary analysis of the sample data produces the following sample information:

$$N = 25 \qquad \sum_{i=1}^{N} Y_i = 1{,}034.97 \qquad \sum_{i=1}^{N} X_i = 528.599 \qquad \sum_{i=1}^{N} Y_i^2 = 45{,}237.19$$

$$\sum_{i=1}^{N} X_i^2 = 11{,}432.92 \qquad \sum_{i=1}^{N} X_i Y_i = 22{,}250.54 \qquad \sum_{i=1}^{N} x_i y_i = 367.179$$

$$\sum_{i=1}^{N} y_i^2 = 2{,}390.67 \qquad \sum_{i=1}^{N} x_i^2 = 256.241 \qquad \sum_{i=1}^{N} \hat{y}_i^2 = 526.147$$

where $x_i \equiv X_i - \overline{X}$, $y_i \equiv Y_i - \overline{Y}$ and $\hat{y}_i \equiv \hat{Y}_i - \overline{Y}$ for i = 1, ..., N. Use the above sample information to answer all the following questions. **Show explicitly all formulas and calculations**.

(10 marks)

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

(6 marks)

(b) 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.

(6 marks)

(c) Calculate an estimate of σ^2, the error variance.

(6 marks)

(d) Compute the value of R^2, the coefficient of determination for the estimated OLS sample regression equation. Briefly explain what the calculated value of R^2 means.

(6 marks)

(e) Compute the estimated variance of $\hat{\beta}_2$ and the estimated standard error of $\hat{\beta}_2$.

(16 marks)

6. You have been commissioned to investigate the relationship between annual R&D expenditures (Y) and total annual sales revenues (X) for chemical firms. You have assembled data for a sample of 32 chemical firms, where Y_i is annual R&D expenditures of the i-th firm (measured in *millions* of dollars per year) and X_i is total annual sales revenues of the i-th firm (measured in *millions* of dollars per year). Your research assistant has used the sample data to estimate the following OLS sample regression equation, where the figures in parentheses below the coefficient estimates are the *estimated standard errors* of the coefficient estimates:

$$\underset{}{Y_i} = \underset{(20.515)}{-0.5772} + \underset{(0.0024487)}{0.04063}X_i + \hat{u}_i \qquad (i = 1, ..., N) \qquad N = 32 \qquad \textbf{(3)}$$

(8 marks)

(a) Compute the two-sided 95% confidence interval for the slope coefficient β_2.

(8 marks)

(b) 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 $\alpha = 0.01$). 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.

Percentage Points of the t-Distribution

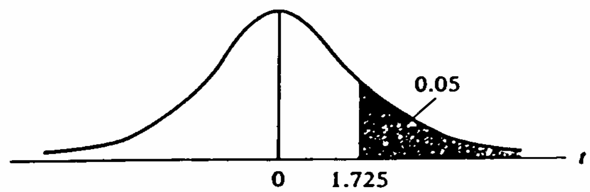
TABLE D.2
Percentage points of the *t* distribution

Example

Pr(t > 2.086) = 0.025

Pr(t > 1.725) = 0.05 for df = 20

Pr($|t|$ > 1.725) = 0.10



Pr / df	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.010	0.001 0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
40	0.681	1.303	1.684	2.021	2.423	2.704	3.307
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232
120	0.677	1.289	1.658	1.980	2.358	2.617	3.160
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090

Note: The smaller probability shown at the head of each column is the area in one tail; the larger probability is the area in both tails.

Source: From E. S. Pearson and H. O. Hartley, eds., *Biometrika Tables for Statisticians*, vol. 1, 3d ed., table 12, Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika*.