QUEEN'S UNIVERSITY AT KINGSTON
Department of Economics

ECONOMICS 351* - Winter Term 2005

Introductory Econometrics

Winter Term 2005 **MID-TERM EXAM: ANSWERS** M.G. Abbott

DATE: **Thursday March 10, 2005.**

TIME: **80 minutes; 4:00 p.m. - 5:20 p.m.**

INSTRUCTIONS: The exam consists of **FIVE (5)** questions. Students are required to answer **ALL FIVE (5)** questions.
Answer all questions in the exam booklets provided. Be sure your ***student number*** is 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!

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

$$Y_i = \beta_0 + \beta_1 X_i + u_i \qquad \textbf{(1)}$$

where Y_i and X_i are observable variables, β_0 and β_1 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

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i \qquad (i = 1, ..., N) \qquad \textbf{(2)}$$

where $\hat{\beta}_0$ is the OLS estimator of the intercept coefficient β_0, $\hat{\beta}_1$ is the OLS estimator of the slope coefficient β_1, $\hat{u}_i$ is the OLS residual for the i-th sample observation, and N is sample size (the number of observations in the sample).

QUESTIONS: **Answer ALL FIVE questions.**

(15 marks)

1. Show that the OLS slope coefficient estimator $\hat{\beta}_1$ is a linear function of the Y_i sample values. Stating explicitly all required assumptions, prove that the OLS slope coefficient estimator $\hat{\beta}_1$ is an unbiased estimator of the slope coefficient β_1.

ANSWER to Question 1:

(5 marks)

- Show that the OLS slope coefficient estimator $\hat{\beta}_1$ is a *linear* function of the Y_i sample values.

$$\hat{\beta}_1 = \frac{\sum_i x_i y_i}{\sum_i x_i^2} = \frac{\sum_i x_i (Y_i - \bar{Y})}{\sum_i x_i^2} = \frac{\sum_i x_i Y_i}{\sum_i x_i^2} - \frac{\bar{Y}\sum_i x_i}{\sum_i x_i^2}$$

$$= \frac{\sum_i x_i Y_i}{\sum_i x_i^2} \qquad \text{because } \sum_i x_i = 0$$ **(5 marks)**

$$= \Sigma_i k_i Y_i \qquad \text{where } k_i \equiv \frac{x_i}{\Sigma_i x_i^2}.$$

(10 marks)

- Stating explicitly all required assumptions, prove that the OLS slope coefficient estimator $\hat{\beta}_1$ is an unbiased estimator of the slope coefficient β_1.

(1) **Substitute for Y_i** the expression $Y_i = \beta_0 + \beta_1 X_i + u_i$ from the population regression equation (or PRE). **(5 marks)**

$$\begin{aligned}
\hat{\beta}_1 &= \sum_i k_i Y_i \\
&= \sum_i k_i (\beta_0 + \beta_1 X_i + u_i) && \text{since } Y_i = \beta_0 + \beta_1 X_i + u_i \text{ by assumption (A1)} \\
&= \sum_i (\beta_0 k_i + \beta_1 k_i X_i + k_i u_i) \\
&= \beta_0 \sum_i k_i + \beta_1 \sum_i k_i X_i + \sum_i k_i u_i \\
&= \beta_1 + \sum_i k_i u_i, && \text{since } \sum_i k_i = 0 \text{ and } \sum_i k_i X_i = 1
\end{aligned}$$

(2) Now **take expectations** of the above expression for $\hat{\beta}_1$ conditional on the value X_i of X: **(5 marks)**

$$\begin{aligned}
E(\hat{\beta}_1) &= E(\beta_1) + E[\sum_i k_i u_i] \\
&= \beta_1 + \sum_i k_i E(u_i | X_i) && \text{since } \beta_1 \text{ is a constant and the } k_i \text{ are nonstochastic} \\
&= \beta_1 + \sum_i k_i 0 && \text{since } E(u_i | X_i) = E(u_i) = 0 \text{ by assumption (A2)} \\
&= \beta_1
\end{aligned}$$

(15 marks)

2. Give a general definition of the t-distribution. Starting from this definition, derive the t-statistic for the OLS slope coefficient estimator $\hat{\beta}_1$. State all assumptions required for the derivation.

<u>ANSWER to Question 2</u>:

(2 marks)

- **General Definition of the t-Distribution**

  A random variable has the **t-distribution with *m* degrees of freedom** if it can be constructed by dividing

  (1) a **standard normal random variable Z ~ N(0, 1)**

  by

  (2) the **square root** of an ***independent*** **chi-square random variable V** that has been divided by its degrees of freedom ***m*** .

  <u>Formally</u>: Consider the two random variables Z and V.

  If (1) $Z \sim N(0,1)$
  (2) $V \sim \chi^2[m]$
  (3) Z and V are ***independent***,

  then the random variable

  $$t = \frac{Z}{\sqrt{V/m}} \sim t[m],$$ the **t-distribution** with ***m*** **degrees of freedom**.

(1 mark)

- ***Error Normality Assumption*****:** The random error terms u_i are **independently and identically distributed (iid)** as the **normal distribution** with **zero mean** and **constant variance** σ^2:

  $$u_i \mid X_i \sim N(0, \sigma^2) \text{ for all i} \qquad \text{OR} \qquad u_i \text{ is iid as } N(0, \sigma^2)$$

(3 marks)

- ***Three implications of error normality assumption (A9)***: (follow from ***linearity property*** **of the normal distribution** whereby any *linear* function of a normally distributed random variable is itself normally distributed).

  (1 mark)

  1. The OLS slope coefficient estimator $\hat{\beta}_1$ **is normally distributed:** $\hat{\beta}_1 \sim N(\beta_1, Var(\hat{\beta}_1))$.

  Why? Because $\hat{\beta}_1$ **can be written as a *linear* function of the Y_i values** $\hat{\beta}_1 = \sum_i k_i Y_i$; and the **$Y_i$ values** are normally distributed because they are linear functions of the random error terms u_i.

ANSWER to Question 2 (continued)

(1 mark)

2. The statistic $(N-2)\hat{\sigma}^2/\sigma^2$ has a chi-square distribution with (N – 2) degrees of freedom**:**

$$\frac{(N-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2[N-2].$$

(1 mark)

3. The estimators $\hat{\beta}_1$ and $\hat{\sigma}^2$ are **statistically independent**.

(2 marks)

- ***Numerator* of the t-statistic for $\hat{\beta}_1$:** the $Z(\hat{\beta}_1)$ statistic.

  The normality of the sampling distribution of $\hat{\beta}_1$ implies that $\hat{\beta}_1$ can be written in the form of a **standard normal variable** with mean zero and variance one, denoted as N(0,1).

$$\hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{\sum_i x_i^2}\right) \quad \Rightarrow \quad Z(\hat{\beta}_1) = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\text{Var}(\hat{\beta}_1)}} = \frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} \sim N(0,1)$$

  where the **Z-statistic for** $\hat{\beta}_1$ can be written as

$$Z(\hat{\beta}_1) = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\text{Var}(\hat{\beta}_1)}} = \frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} = \frac{\hat{\beta}_1 - \beta_1}{\sigma\big/\sqrt{\sum_i x_i^2}} = \frac{\left(\hat{\beta}_1 - \beta_1\right)\sqrt{\sum_i x_i^2}}{\sigma}. \qquad (1)$$

(2 marks)

- ***Denominator* of the t-statistic for $\hat{\beta}_1$:**

  The error normality assumption implies that the statistic $\hat{\sigma}^2/\sigma^2$ has a degrees-of-freedom-adjusted chi-square distribution with (N – 2) degrees of freedom; that is

$$\frac{(N-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2[N-2] \quad \Rightarrow \quad \frac{\hat{\sigma}^2}{\sigma^2} \sim \frac{\chi^2[N-2]}{(N-2)} \quad \Rightarrow \quad \frac{\hat{\sigma}}{\sigma} \sim \sqrt{\frac{\chi^2[N-2]}{(N-2)}}. \qquad (2)$$

  The last term $\hat{\sigma}/\sigma$ in (2) is the denominator of the t-statistic for $\hat{\beta}_1$: it is distributed as the square root of a degrees-of-freedom-adjusted chi-square variable with (N – 2) degrees of freedom.

ANSWER to Question 2 (continued)

(5 marks)

- **The t-statistic for** $\hat{\beta}_1$.

  Since $\hat{\beta}_1$ and $\hat{\sigma}^2$ are **statistically independent**, the t-statistic for $\hat{\beta}_1$ is the ratio of (1) to (2): i.e.,

$$t(\hat{\beta}_1) = \frac{Z(\hat{\beta}_1)}{\hat{\sigma}/\sigma} = \frac{\left(\hat{\beta}_1 - \beta_1\right)\sqrt{\sum_i x_i^2}\,/\sigma}{\hat{\sigma}/\sigma} = \frac{\left(\hat{\beta}_1 - \beta_1\right)\sqrt{\sum_i x_i^2}}{\hat{\sigma}}. \qquad (3)$$

  - Dividing the numerator and denominator of (3) by $\sqrt{\sum_i x_i^2}$ yields

$$t(\hat{\beta}_1) = \frac{\left(\hat{\beta}_1 - \beta_1\right)}{\hat{\sigma}\Big/\sqrt{\sum_i x_i^2}}. \qquad (4)$$

  - But the denominator of (4) is simply the ***estimated*** **standard error of** $\hat{\beta}_1$; i.e.,

$$\frac{\hat{\sigma}}{\sqrt{\sum_i x_i^2}} = \sqrt{\hat{Var}(\hat{\beta}_1)} = \hat{se}(\hat{\beta}_1).$$

■ **<u>Result</u>:** The **t-statistic for** $\hat{\beta}_1$ thus takes the form

$$t(\hat{\beta}_1) = \frac{\left(\hat{\beta}_1 - \beta_1\right)}{\hat{\sigma}\Big/\sqrt{\sum_i x_i^2}} = \frac{\left(\hat{\beta}_1 - \beta_1\right)}{\sqrt{\hat{Var}(\hat{\beta}_1)}} = \frac{\left(\hat{\beta}_1 - \beta_1\right)}{\hat{se}(\hat{\beta}_1)} \sim t[N-2]. \qquad \mathbf{(5)}$$

(10 marks)

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

(a) $\hat{\theta}$ is a minimum variance estimator of θ.

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

What is the difference between the minimum variance and efficiency properties of the estimator $\hat{\theta}$?

ANSWER to Question 3:

(4 marks)

- (a) $\hat{\theta}$ is a minimum variance estimator of θ. **(4 marks)**

  The variance of the estimator $\hat{\theta}$ is ***smaller than*** the variance of *any other* estimator of the parameter θ.

  If $\tilde{\theta}$ is any other estimator of θ, then $\hat{\theta}$ is a ***minimum variance*** **estimator** of θ if

  $$\text{Var}(\hat{\theta}) \le \text{Var}(\tilde{\theta}).$$

(4 marks)

- (b) $\hat{\theta}$ is an efficient estimator of θ. **(4 marks)**

  The estimator $\hat{\theta}$ is an efficient estimator if it is ***unbiased*** and has ***smaller variance*** than *any other* ***unbiased*** estimator of the parameter θ.

  If $\tilde{\theta}$ is *any other* ***unbiased*** *estimator* of θ, then $\hat{\theta}$ is an ***efficient*** **estimator** of θ if

  $$\text{Var}(\hat{\theta}) \le \text{Var}(\tilde{\theta}) \qquad \text{where } E(\hat{\theta}) = \theta \text{ and } E(\tilde{\theta}) = \theta.$$

(2 marks)

- The important difference between statements (a) and (b) is that **an *efficient* estimator must be *unbiased*** whereas a minimum variance estimator may be biased or unbiased.

  An ***efficient*** **estimator** is the ***minimum variance*** **estimator in the class of all *unbiased* estimators** of the parameter θ.

(36 marks)

4. A researcher is using data for a sample of 121 students in an Introductory Econometrics course to investigate the relationship between students' grades on the final exam Y_i (measured in *percentage points*) and their grades on the mid-term exam X_i (measured in *percentage points*). The population regression equation takes the form of equation (1): $Y_i = \beta_0 + \beta_1 X_i + u_i$.
Preliminary analysis of the sample data produces the following sample information:

$$N = 121 \qquad \sum_{i=1}^{N} Y_i = 8344.0 \qquad \sum_{i=1}^{N} X_i = 8957.0 \qquad \sum_{i=1}^{N} Y_i^2 = 616433.0$$

$$\sum_{i=1}^{N} X_i^2 = 713631.0 \qquad \sum_{i=1}^{N} X_i Y_i = 641136.5 \qquad \sum_{i=1}^{N} x_i y_i = 23473.62$$

$$\sum_{i=1}^{N} y_i^2 = 41041.79 \qquad \sum_{i=1}^{N} x_i^2 = 50590.93 \qquad \sum_{i=1}^{N} \hat{u}_i^2 = 30150.294$$

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

<u>ANSWERS to Question 4</u>:

(10 marks)

(a) Use the above information to compute OLS estimates of the intercept coefficient β_0 and the slope coefficient β_1.

- $\hat{\beta}_1 = \dfrac{\sum_i x_i y_i}{\sum_i x_i^2} = \dfrac{23473.62}{50590.93} = \mathbf{0.46399} = \underline{\mathbf{0.464}}$ **(5 marks)**

- $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$

  $$\overline{Y} = \frac{\sum_{i=1}^{N} Y_i}{N} = \frac{8344}{121} = 68.9587 \quad \text{and} \quad \overline{X} = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{8957}{121} = 74.0248$$

  Therefore

  $$\hat{\beta}_0 = \overline{Y} - \beta_1 \overline{X} = 68.9587 - (0.46399)(74.0248) = 68.9587 - 34.3467 = \underline{\mathbf{34.612}}$$

  (5 marks)

ANSWERS to Question 4 (continued):

(5 marks)

(b) Interpret the slope coefficient estimate you calculated in part (a) -- i.e., explain in words what the numeric value you calculated for $\hat{\beta}_1$ means.

Note: $\hat{\beta}_1$ = **0.46399**. Y_i is measured in ***percentage points***, and X_i is measured in ***percentage points***.

The estimate **0.46399** of β_1 means that an ***increase*** **(decrease) in mid-term exam grade** X_i of ***1 percentage point*** is associated on average with **an *increase* (decrease) in final exam grade equal to *0.46399* *percentage points***, or ***0.464*** ***percentage points***.

(5 marks)

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

$$\text{RSS} = \sum_{i=1}^{N} \hat{u}_i^2 = 30150.294; \qquad N - 2 = 121 - 2 = 119$$

$$\hat{\sigma}^2 = \frac{\text{RSS}}{N-2} = \frac{\sum_{i=1}^{N} \hat{u}_i^2}{N-2} = \frac{30{,}150.294}{121-2} = \frac{30{,}150.294}{119} = \underline{\mathbf{253.3638}}$$

(5 marks)

(5 marks)

(d) Calculate an estimate of $\text{Var}(\hat{\beta}_1)$, the variance of $\hat{\beta}_1$.

$$\hat{\text{Var}}(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{\sum_{i=1}^{N} x_i^2} = \frac{253.3638}{50590.93} = \underline{\mathbf{0.005008087}} = \underline{\mathbf{0.0050081}}$$

(5 marks)

(6 marks)

(e) 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.

(4 marks)

$$R^2 = \frac{\text{ESS}}{\text{TSS}} = \frac{\sum_{i=1}^{N} y_i^2 - \sum_{i=1}^{N} \hat{u}_i^2}{\sum_{i=1}^{N} y_i^2} = \frac{41041.79 - 30150.294}{41041.79} = \frac{10891.496}{41041.79} = \underline{\mathbf{0.2654}}$$

OR

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{N} \hat{u}_i^2}{\sum_{i=1}^{N} y_i^2} = 1 - \frac{30150.294}{41041.79} = 1 - 0.7346 = \underline{\mathbf{0.2654}}$$

(2 marks)
Interpretation of R^2 = 0.2654: The value of 0.2654 indicates that **26.54 percent of the total sample (or observed) variation in Y_i** (final exam grades) **is *attributable to*, or *explained by*, the sample regression function or the regressor X_i** (mid-term exam grades).

(5 marks)
(f) Calculate the sample value of the F-statistic for testing the null hypothesis H_0: $\beta_1 = 0$ against the alternative hypothesis H_1: $\beta_1 \neq 0$. (Note: You are not required to obtain or state the inference of this test.)

- F-statistic for $\hat{\beta}_1$ is $F(\hat{\beta}_1) = \dfrac{(\hat{\beta}_1 - \beta_1)^2}{\hat{Var}(\hat{\beta}_1)}$ (1) **(2 marks)**

- From part (a), $\hat{\beta}_1$ = 0.46399; from part (d), $\hat{Var}(\hat{\beta}_1) = 0.0050081$.

- Calculate the ***sample value* of the F-statistic** (1) under H_0: set $\beta_1 = 0$, $\hat{\beta}_1$ = 0.46399 and $\hat{Var}(\hat{\beta}_1) = 0.0050081$ in (1).

$$F_0(\hat{\beta}_1) = \frac{(\hat{\beta}_1 - \beta_1)^2}{\hat{Var}(\hat{\beta}_1)} = \frac{(0.46399 - 0)^2}{0.0050081} = \frac{0.215287}{0.0050081} = \underline{\mathbf{42.99}}$$ **(3 marks)**

Alternative Answer to 4(f): use the ANOVA F-statistic

- ANOVA F-statistic is: $ANOVA - F_0 = \dfrac{ESS/1}{RSS/N-2} = \dfrac{\sum_i y_i^2 - \sum_i \hat{u}_i^2}{\hat{\sigma}^2}$ **(2 marks)**

- $\sum_{i=1}^{N} y_i^2 = 41041.79$; $\sum_{i=1}^{N} \hat{u}_i^2 = 30150.294$; from part (c), $\hat{\sigma}^2 = \sum_{i=1}^{N} \hat{u}_i^2 \Big/ (N-2) = \underline{\mathbf{253.3638}}$

- Calculate the ***sample value* of the ANOVA F-statistic**. **(3 marks)**

$$ANOVA - F_0 = \frac{\sum_i y_i^2 - \sum_i \hat{u}_i^2}{\hat{\sigma}^2} = \frac{41041.79 - 30150.294}{253.3638} = \frac{10891.496}{253.3638} = \underline{\mathbf{42.99}}$$

(24 marks)

5. You have been commissioned to investigate the relationship between the birth weight of newborn babies and the average number of cigarettes women smoked per day during pregnancy. The dependent variable is ***bwght$_i$***, the birth weight of the baby born to the i-th mother, measured in *grams*. The explanatory variable is ***cigs$_i$***, the average number of cigarettes per day smoked by the i-th mother during pregnancy, measured in *cigarettes per day*. The model you propose to estimate is given by the population regression equation

$$\text{bwght}_i = \beta_0 + \beta_1 \text{cigs}_i + u_i .$$

Your research assistant has used 1,722 sample observations on bwght_i and cigs_i 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:

$$\underset{}{\text{bwght}_i} = \underset{(14.145)}{3421.71} - \underset{(3.2447)}{11.4783}\text{cigs}_i + \hat{u}_i \qquad (i = 1, ..., N) \qquad N = 1,722 \qquad \textbf{(3)}$$

$\leftarrow$ (standard errors)

ANSWERS to Question 5:

(8 marks)

(a) 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 state the conclusion you would draw from the test.

H_0: $\beta_1 = 0$

H_1: $\beta_1 \neq 0$ a ***two-sided* alternative hypothesis** $\Rightarrow$ a ***two-tailed* test**

- Test statistic is either $t(\hat{\beta}_1) = \dfrac{\hat{\beta}_1 - \beta_1}{\hat{se}(\hat{\beta}_1)} \sim t[N-2]$ *or* $F(\hat{\beta}_1) = \dfrac{(\hat{\beta}_1 - \beta_1)^2}{\hat{Var}(\hat{\beta}_1)} \sim F[1, N-2]$.

- $\hat{\beta}_1 = \mathbf{-11.4783}$; $\hat{se}(\hat{\beta}_1) = \mathbf{3.2447}$; $\hat{Var}(\hat{\beta}_1) = \left(\hat{se}(\hat{\beta}_1)\right)^2 = \mathbf{10.528}$

- Calculate the ***sample value* of *either* the t-statistic *or* the F-statistic** under H_0: set $\beta_1 = 0$, $\hat{\beta}_1 = \mathbf{-11.4783}$, $\hat{se}(\hat{\beta}_1) = \mathbf{3.2447}$, and $\hat{Var}(\hat{\beta}_1) = \mathbf{10.528}$.

$$t_0(\hat{\beta}_1) = \frac{\hat{\beta}_1 - \beta_1}{\hat{se}(\hat{\beta}_1)} = \frac{-11.4783 - 0.0}{3.2447} = \frac{-11.4783}{3.2447} = -3.5376 = \underline{\mathbf{-3.54}}$$

or **(3 marks)**

$$F_0(\hat{\beta}_1) = \frac{(\hat{\beta}_1 - \beta_1)^2}{\hat{Var}(\hat{\beta}_1)} = \frac{(-11.4783 - 0.0)^2}{10.528} = \frac{131.7514}{10.528} = 12.5144 = \underline{\mathbf{12.51}}$$

ANSWER to Question 5(a) -- continued:

- **Null distribution** of $t_0(\hat{\beta}_1)$ is **t[N – 2] = t[1722 – 2] = t[1720]**
- **Null distribution** of $F_0(\hat{\beta}_1)$ is **F[1, N – 2] = F[1, 1722 – 2] = F[1, 1720]**

Decision Rule: At significance level α, **(2 marks)**

- ***reject*** $\mathbf{H_0}$ if $F_0(\hat{\beta}_1) > F_\alpha[1, 1720]$ *or* $|t_0(\hat{\beta}_1)| > t_{\alpha/2}[1720]$,

  i.e., if either (1) $t_0(\hat{\beta}_1) > t_{\alpha/2}[1720]$ or (2) $t_0(\hat{\beta}_1) < -t_{\alpha/2}[1720]$;

- ***retain*** $\mathbf{H_0}$ if $F_0(\hat{\beta}_1) \le F_\alpha[1, 1720]$ *or* $|t_0(\hat{\beta}_1)| \le t_{\alpha/2}[1720]$,

  i.e., if $-t_{\alpha/2}[1720] \le t_0(\hat{\beta}_1) \le t_{\alpha/2}[1720]$.

Critical values of t[1720]-distribution or F[1, 1720]-distribution: in t-table, use **df = ∞**.

- ***two-tailed*** **1 percent critical value** = $t_{\alpha/2}[1720] = t_{0.005}[1720] = 2.576 =$ **2.58** **(1 mark)**
- **1 percent critical value** = $F_\alpha[1, 1720] = F_{0.01}[1, 1720] =$ **6.65**

Inference: **(1 mark)**

- At **1 percent significance level**, i.e., for $\alpha = 0.01$,

  $|t_0(\hat{\beta}_1)| =$ **3.54 > 2.58** $= t_{0.005}[1720] \Rightarrow$ ***reject*** $\mathbf{H_0}$ vs. H_1 **at 1 percent level**.

  $F_0(\hat{\beta}_1) =$ **12.51 > 6.65** $= F_{0.01}[1, 1720] \Rightarrow$ ***reject*** $\mathbf{H_0}$ vs. H_1 **at 1 percent level**.

- **Inference: At the 1% significance level**, the null hypothesis $\beta_1 = 0$ is ***rejected*** in favour of the alternative hypothesis $\beta_1 \neq 0$.

Conclusion implied by test outcome: **(1 mark)**

Rejection of the null hypothesis $\beta_1 = 0$ against the alternative hypothesis $\beta_1 \neq 0$ means that **the sample evidence favours the existence of a relationship between** students' ***final exam grades*** and their ***mid-term exam grades***.

Question 5(a) – Alternative Answer -- uses confidence interval approach

- The **two-sided (1 – α)-level,** or **100(1 – α) percent, confidence interval for β_1** is:

$$\hat{\beta}_1 - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1) \le \beta_1 \le \hat{\beta}_1 + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1)$$
$$\hat{\beta}_{1L} \le \beta_1 \le \hat{\beta}_{1U}$$

- Required results and intermediate calculations:

$N - K = 1722 - 2 = 1720$; $\hat{\beta}_1 = \mathbf{-11.4783}$; $\hat{se}(\hat{\beta}_1) = \mathbf{3.2447}$

$1 - \alpha = 0.99 \Rightarrow \alpha = 0.01 \Rightarrow \boldsymbol{\alpha/2 = 0.005}$: $t_{\alpha/2}[N-2] = \mathbf{t_{0.005}[1720]} = \underline{\mathbf{2.58}}$ **(1 mark)**

$$t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1) = t_{0.005}[1720]\hat{se}(\hat{\beta}_1) = 2.58(3.2447) = \mathbf{8.37133}$$

- **Lower 99% confidence limit for β_1** is: **(2 marks)**

$$\hat{\beta}_{1L} = \hat{\beta}_1 - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1) = \hat{\beta}_1 - t_{0.005}[1720]\hat{se}(\hat{\beta}_1)$$
$$= -11.4783 - 2.58(3.2447) = -11.4783 - 8.37133 = -19.8496 = \underline{\mathbf{-19.85}}$$

- **Upper 99% confidence limit for β_1** is: **(2 marks)**

$$\hat{\beta}_{1U} = \hat{\beta}_1 + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1) = \hat{\beta}_1 + t_{0.005}[1720]\hat{se}(\hat{\beta}_1)$$
$$= -11.4783 + 2.58(3.2447) = -11.4783 + 8.37133 = -3.10697 = \underline{\mathbf{-3.107}}$$

- **Decision Rule:** At significance level α, **(1 mark)**

  - ***reject*** **H_0** if the ***hypothesized*** **value b_1 of β_1** specified by H_0 **lies *outside*** the two-sided (1−α)-level confidence interval for β_1, i.e., if either
  (1) $b_1 < \hat{\beta}_1 - t_{\alpha/2}[1720]\hat{se}(\hat{\beta}_1)$ or (2) $b_1 > \hat{\beta}_1 + t_{\alpha/2}[1720]\hat{se}(\hat{\beta}_1)$.

  - ***retain*** **H_0** if the ***hypothesized*** **value b_1 of β_1** specified by H_0 **lies *inside*** the two-sided (1−α)-level confidence interval for β_1, i.e., if
  $\hat{\beta}_1 - t_{\alpha/2}[1720]\hat{se}(\hat{\beta}_1) \le b_1 \le \hat{\beta}_1 + t_{\alpha/2}[1720]\hat{se}(\hat{\beta}_1)$.

Question 5(a) – Alternative Answer (continued)

<u>Inference</u>: **(1 mark)**

- At **1 percent significance level**, i.e., for $\alpha = 0.01$,

  $$b_1 = \mathbf{0} > \mathbf{-3.107} = \hat{\beta}_{1U} = \hat{\beta}_1 + t_{0.005}[1720]\hat{se}(\hat{\beta}_1)$$

  $\Rightarrow$ ***reject*** $\mathbf{H_0}$ vs. H_1 **at 1 percent level**.

- **<u>Inference</u>: At the 1% significance level**, the null hypothesis $\beta_1 = 0$ is ***rejected*** in favour of the alternative hypothesis $\beta_1 \neq 0$.

<u>Conclusion implied by test outcome</u>: **(1 mark)**

Rejection of the null hypothesis $\beta_1 = 0$ against the alternative hypothesis $\beta_1 \neq 0$ means that **the sample evidence favours the existence of a relationship between** students' ***final exam grades*** and their ***mid-term exam grades***.

(8 marks)

(b) Perform a test of the proposition that an increase in women's cigarette consumption during pregnancy of one cigarette per day *decreases* the average birth weight of babies by *less than* 20 grams. Use the 1 percent significance level (i.e., $\alpha = 0.01$). State the null hypothesis H_0 and the alternative hypothesis H_1. Show how you calculated the test statistic. State the decision rule you use, and the inference you would draw from the test.

ANSWER to Question 5(b):

Null and Alternative Hypotheses:

H_0: $\beta_1 = -20$

H_1: $\beta_1 > -20$ $\Rightarrow$ a ***right-tailed*** **t-test** **(1 mark)**

- Test statistic is $t(\hat{\beta}_1) = \dfrac{\hat{\beta}_1 - \beta_1}{\hat{se}(\hat{\beta}_1)} \sim t[N-2]$; $\hat{\beta}_1 = \mathbf{-11.4783}$ and $\hat{se}(\hat{\beta}_1) = \mathbf{3.2447}$

- Calculate the ***sample value*** **of the t-statistic** (1) under H_0: set $\beta_1 = -20$, $\hat{\beta}_1 = \mathbf{-11.4783}$ and $\hat{se}(\hat{\beta}_1) = \mathbf{3.2447}$ in (1).

$$t_0(\hat{\beta}_1) = \frac{\hat{\beta}_1 - \beta_1}{\hat{se}(\hat{\beta}_1)} = \frac{-11.4783 - (-20.0)}{3.2447} = \frac{8.5217}{3.2447} = 2.6263445 = \underline{\mathbf{2.63}}$$ **(3 marks)**

- **Null distribution** of $t_0(\hat{\beta}_1)$ is $\mathbf{t[N-2] = t[1722-2] = t[1720]}$

Decision Rule: At significance level α, **(1 mark)**

- ***reject*** $\mathbf{H_0}$ if $t_0(\hat{\beta}_1) > t_\alpha[1720]$,
- ***retain*** $\mathbf{H_0}$ if $t_0(\hat{\beta}_1) \le t_\alpha[1720]$.

Critical value of t[1720]-distribution: from t-table, use **df =** ∞.

- ***right-tail*** **1 percent critical value** = $t_{0.01}[1720]$ = **2.326** = **2.33** **(1 mark)**

Inference: **(2 marks)**

- At **1 percent significance level**, i.e., for $\alpha = 0.01$,

  $t_0(\hat{\beta}_1) = \mathbf{2.63} > \mathbf{2.33} = t_{0.01}[1720]$ $\Rightarrow$ ***reject*** $\mathbf{H_0}$ vs. H_1 **at 1 percent level**.

- **Inference: At the 1% significance level**, the null hypothesis $\boldsymbol{\beta_1 = -20}$ **is** ***rejected*** in favour of the alternative hypothesis $\boldsymbol{\beta_1 > -20}$.

(8 marks)

(c) Compute the two-sided 95% confidence interval for the slope coefficient β_1.

<u>ANSWER to Question 5(c)</u>:

- The **two-sided (1 – α)-level,** or **100(1 – α) percent, confidence interval for β_1** is computed as

$$\hat{\beta}_1 - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1) \le \beta_1 \le \hat{\beta}_1 + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1)$$ **(2 marks)**

where

- $\hat{\beta}_{1L} = \hat{\beta}_1 - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1)$ = the ***lower*** **100(1 – α)% confidence limit for β_1**
- $\hat{\beta}_{1U} = \hat{\beta}_1 + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1)$ = the ***upper*** **100(1 – α)% confidence limit for β_1**
- $t_{\alpha/2}[N-2]$ = the **α/2 critical value of the t-distribution with N–2 degrees of freedom**.

- Required results and intermediate calculations:

N – K = 1722 – 2 = 1720; $\hat{\beta}_1$ = **– 11.4783**; $\hat{se}(\hat{\beta}_1)$ = **3.2447**

1 – α = 0.95 ⇒ α = 0.05 ⇒ **α/2 = 0.025**: $t_{\alpha/2}[N-2]$ = $\mathbf{t_{0.025}[1720] = 1.96}$

$t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1) = t_{0.025}[1720]\hat{se}(\hat{\beta}_1)$ = 1.96(3.2447) = **6.359612**

- **Lower 95% confidence limit for β_1** is: **(3 marks)**

$$\hat{\beta}_{1L} = \hat{\beta}_1 - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1) = \hat{\beta}_1 - t_{0.025}[1720]\hat{se}(\hat{\beta}_1)$$

= – 11.4783 – 1.96(3.2447) = – 11.4783 – 6.359612 = – 17.8379 **= <u>– 17.84</u>**

- **Upper 95% confidence limit for β_1** is: **(3 marks)**

$$\hat{\beta}_{1U} = \hat{\beta}_1 + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_1) = \hat{\beta}_1 + t_{0.025}[1720]\hat{se}(\hat{\beta}_1)$$

= – 11.4783 + 1.96(3.2447) = – 11.4783 + 6.359612 = – 5.11869 **= <u>– 5.119</u>**

- **<u>Result</u>:** The **two-sided 95% confidence interval for β_1** is: **[– 17.84, – 5.119]**

Percentage Points of the t-Distribution

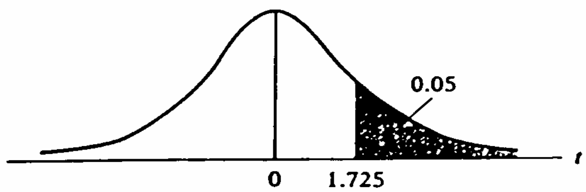
TABLE D.2
Percentage points of the *t* distribution

Example

Pr(t > 2.086) = 0.025

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

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



df \ Pr	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*.