QUEEN'S UNIVERSITY AT KINGSTON
Department of Economics

CONFIDENTIAL
turn in exam
question paper

ECONOMICS 351* - Winter Term 2009

Introductory Econometrics

Winter Term 2009 **MID-TERM EXAM: ANSWERS** M.G. Abbott

DATE: **Monday March 2, 2009.**

TIME: **80 minutes; 11:30 a.m. - 12:50 p.m.**

INSTRUCTIONS: The exam consists of **THREE (3)** questions. Students are required to answer **ALL THREE (3)** 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. Your ***name*** is optional.

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.

Tables of percentage points of the t-distribution and F-distribution are given on the last two pages of the exam.

This exam is **CONFIDENTIAL**. This question paper must be submitted in its entirety with your answer booklet(s); otherwise your exam will not be marked.

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 FOUR 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:

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

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).

(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.

ANSWER to Question 1:

(3 marks)

- State the Ordinary Least Squares (OLS) estimation criterion. **(3 marks)**

The OLS coefficient estimators are **those formulas or expressions for $\hat{\beta}_0$ and $\hat{\beta}_1$ that minimize the sum of squared residuals RSS** for any given sample of size N.

The **OLS estimation criterion** is therefore:

$$\underset{\{\hat{\beta}_j\}}{\text{Minimize}}\ \text{RSS}\left(\hat{\beta}_0, \hat{\beta}_1\right) = \sum_{i=1}^{N} \hat{u}_i^2 = \sum_{i=1}^{N} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right)^2$$

Question 1 (continued)

(4 marks)

- State the OLS normal equations. **(4 marks)**

The ***first*** **OLS normal equation** can be written in *any one* of the following forms:

$$\sum_{i=1}^{N} Y_i - N\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{N} X_i = 0$$

$$-N\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{N} X_i = -\sum_{i=1}^{N} Y_i \qquad \textbf{(N1)}$$

$$N\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{N} X_i = \sum_{i=1}^{N} Y_i$$

The ***second*** **OLS normal equation** can be written in *any one* of the following forms:

$$\sum_{i=1}^{N} X_i Y_i - \hat{\beta}_0 \sum_{i=1}^{N} X_i - \hat{\beta}_1 \sum_{i=1}^{N} X_i^2 = 0$$

$$-\hat{\beta}_0 \sum_{i=1}^{N} X_i - \hat{\beta}_1 \sum_{i=1}^{N} X_i^2 = -\sum_{i=1}^{N} X_i Y_i \qquad \textbf{(N2)}$$

$$\hat{\beta}_0 \sum_{i=1}^{N} X_i + \hat{\beta}_1 \sum_{i=1}^{N} X_i^2 = \sum_{i=1}^{N} X_i Y_i$$

(8 marks)

- Derive the OLS normal equations from the OLS estimation criterion.

(4 marks)

Step 1: **Partially differentiate** the $RSS\left(\hat{\beta}_0, \hat{\beta}_1\right)$ function with respect to $\hat{\beta}_0$ and $\hat{\beta}_1$, using

$$\hat{u}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i \qquad \Rightarrow \qquad \frac{\partial \hat{u}_i}{\partial \hat{\beta}_0} = -1 \qquad \text{and} \qquad \frac{\partial \hat{u}_i}{\partial \hat{\beta}_1} = -X_i .$$

$$\frac{\partial RSS}{\partial \hat{\beta}_0} = \sum_{i=1}^{N} 2\hat{u}_i \left(\frac{\partial \hat{u}_i}{\partial \hat{\beta}_0}\right) = \sum_{i=1}^{N} 2\,\hat{u}_i(-1) = -2\sum_{i=1}^{N} \hat{u}_i = -2\sum_{i=1}^{N}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) \qquad \textbf{(2 marks)} \qquad (1)$$

Question 1 (continued)

$$\frac{\partial RSS}{\partial \hat{\beta}_1} = \sum_{i=1}^{N} 2\hat{u}_i\left(\frac{\partial \hat{u}_i}{\partial \hat{\beta}_1}\right) = \sum_{i=1}^{N} 2\hat{u}_i(-X_i) = -2\sum_{i=1}^{N} X_i\hat{u}_i$$

$$= -2\sum_{i=1}^{N} X_i\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) \quad \text{since } \hat{u}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i \qquad \textbf{(2 marks)} \qquad (2)$$

$$= -2\sum_{i=1}^{N}\left(X_iY_i - \hat{\beta}_0 X_i - \hat{\beta}_1 X_i^2\right).$$

(4 marks)

Step 2: **Obtain the first-order conditions (FOCs)** for a minimum of the RSS function by **setting the partial derivatives (1) and (2) *equal to zero*** and then dividing each equation by −2 and re-arranging:

$$\frac{\partial RSS}{\partial \hat{\beta}_0} = 0 \quad \Rightarrow \quad -2\sum_{i=1}^{N}\hat{u}_i = 0 \Rightarrow \quad -2\sum_{i=1}^{N}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) = 0$$

$$\Rightarrow \quad \sum_{i=1}^{N}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) = 0$$

$$\Rightarrow \quad \sum_{i=1}^{N} Y_i - N\hat{\beta}_0 - \hat{\beta}_1\sum_{i=1}^{N} X_i = 0$$

$$\Rightarrow \quad \sum_{i=1}^{N} Y_i = N\hat{\beta}_0 + \hat{\beta}_1\sum_{i=1}^{N} X_i \qquad \textbf{(N1)} \qquad \textbf{(2 marks)}$$

$$\frac{\partial RSS}{\partial \hat{\beta}_1} = 0 \quad \Rightarrow \quad -2\sum_{i=1}^{N} X_i\hat{u}_i = 0 \quad \Rightarrow \quad -2\sum_{i=1}^{N} X_i\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) = 0$$

$$\Rightarrow \quad \sum_{i=1}^{N} X_i\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) = 0$$

$$\Rightarrow \quad \sum_{i=1}^{N}\left(X_iY_i - \hat{\beta}_0 X_i - \hat{\beta}_1 X_i^2\right) = 0$$

$$\Rightarrow \quad \sum_{i=1}^{N} X_iY_i - \hat{\beta}_0\sum_{i=1}^{N} X_i - \hat{\beta}_1\sum_{i=1}^{N} X_i^2 = 0$$

$$\Rightarrow \quad \sum_{i=1}^{N} X_iY_i = \hat{\beta}_0\sum_{i=1}^{N} X_i + \hat{\beta}_1\sum_{i=1}^{N} X_i^2 \quad \textbf{(N2)} \qquad \textbf{(2 marks)}$$

(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**:

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

(3 marks) – one for each of the three implications

- ***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).

ANSWER to Question 2 (continued)

(1 mark)

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

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.

(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).

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

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

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

ANSWER to Question 2 (continued)

(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.

(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

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

(70 marks)

3. You have been commissioned to investigate the relationship between the net financial wealth of single adults 25-64 years of age and their annual income. The dependent variable is ***netfa*** $netfa_i$, the net financial wealth (or net financial assets) of the i-th single adult person, measured in *thousands of dollars*. The explanatory variable is ***income***$_i$, the annual income of the i-th single adult person, measured in *thousands of dollars per years*. The model you propose to estimate is given by the population regression equation

$$\text{netfa}_i = \beta_0 + \beta_1 \text{income}_i + u_i .$$

Your research assistant has used 2017 sample observations on ***netfa***$_i$ and ***income***$_i$ 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:

$$\underset{}{\text{netfa}_i} = \underset{(2.0607)}{-10.571} + \underset{(0.060900)}{0.82068}\,\text{income}_i + \hat{u}_i \qquad (i = 1, ..., N) \qquad N = 2017 \qquad \textbf{(3)}$$

$\leftarrow$ (standard errors)

sample mean of ***netfa***$_i$ = $\overline{\text{netfa}} = 13.595$

sample mean of ***income***$_i$ = $\overline{\text{income}} = 29.446$

$$\text{TSS} = \sum_{i=1}^{N} (\text{netfa}_i - \overline{\text{netfa}})^2 = 4{,}565{,}965.05 \qquad \text{RSS} = \sum_{i=1}^{N} \hat{u}_i^2 = 4{,}188{,}482.98$$

(4 marks)

(a) State the formula used to compute the OLS estimate of the slope coefficient on ***income***$_i$ in sample regression equation (3), i.e., give the formula for computing the OLS estimate $\hat{\beta}_1$ of β_1.

- $$\hat{\beta}_1 = \frac{\sum_{i=1}^{N} x_i y_i}{\sum_{i=1}^{N} x_i^2} = \frac{\sum_{i=1}^{N} (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{N} (X_i - \overline{X})^2}$$

where $X_i \equiv \text{income}_i$, $Y_i \equiv \text{netfa}_i$,
$\overline{X} \equiv \overline{\text{income}}$ = the sample mean of the $X_i \equiv \text{income}_i$ values,
and $\overline{Y} \equiv \overline{\text{netfa}}$ = the sample mean of the $Y_i \equiv \text{netfa}_i$ values. **(4 marks)**

OR

- $$\hat{\beta}_1 = \frac{\sum_{i=1}^{N}(\text{income}_i - \overline{\text{income}})(\text{netfa}_i - \overline{\text{netfa}})}{\sum_{i=1}^{N}(\text{income}_i - \overline{\text{income}})^2}$$

$\overline{\text{income}}$ = the sample mean of the $X_i \equiv \text{income}_i$ values;

$\overline{\text{netfa}}$ = the sample mean of the $Y_i \equiv \text{netfa}_i$ values. **(4 marks)**

(4 marks)

(b) State the formula used to compute the OLS estimate of the intercept coefficient in sample regression equation (3), i.e., give the formula for computing the OLS estimate $\hat{\beta}_0$ of β_0.

- $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{X} = \overline{\text{netfa}} - \hat{\beta}_1\overline{\text{income}}$ **(4 marks)**

(4 marks)

(c) Interpret the slope coefficient estimate on ***income$_i$*** in sample regression equation (3) – i.e., explain in words what the numeric value of $\hat{\beta}_1$ means.

Note: $\hat{\beta}_1$ = **0.82068**. ***netfa$_i$*** is measured in ***thousands of dollars***, and ***income$_i$*** is measured in ***thousands of dollars per year***.

The estimate **0.82068** of β_1 means that a $**1,000 per year *increase* (decrease) in annual income** is associated on average with **an *increase* (decrease) in net financial wealth equal to $820.68 (or 0.82068 thousands of dollars)**.

(4 marks)

(d) Use the above estimation results for sample regression equation (3) to calculate an estimate of σ^2, the error variance.

$$\text{RSS} = \sum_{i=1}^{N}\hat{u}_i^2 = 4{,}188{,}482.98; \qquad N-2 = 2017-2 = 2015$$

$$\hat{\sigma}^2 = \frac{\text{RSS}}{N-2} = \frac{\sum_{i=1}^{N}\hat{u}_i^2}{N-2} = \frac{4188482.98}{2017-2} = \frac{4188482.98}{2015} = \underline{\mathbf{2078.6516}}$$

$= \underline{\mathbf{2078.65}}$ **(4 marks)**

(6 marks)

(e) Compute the value of R^2, the coefficient of determination for the estimated OLS sample regression equation (3). Briefly explain in words what the value you have calculated for R^2 means.

(4 marks)

$$R^2 = \frac{ESS}{TSS} = \frac{4565965.05 - 4188482.98}{4565965.05} = \frac{377482.07}{4565965.05} = \underline{\mathbf{0.082673}} = \underline{\mathbf{0.0827}}$$

OR

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{4188482.98}{4565965.05} = 1 - 0.917327 = \underline{\mathbf{0.082673}} = \underline{\mathbf{0.0827}}$$

(2 marks)

Interpretation of R^2 = 0.0827: The value of 0.0827 indicates that **8.27 percent of the total sample (or observed) variation in *netfa$_i$*** (net financial wealth) **is *attributable to*, or *explained by*, the regressor *income$_i$*** (annual income) **or the sample regression function**.

(8 marks)

(f) Perform a t-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 α = 0.01). Show how you calculated the test statistic. State the decision rule you use, and the inference you would draw from the test. What would you conclude from the results of the test?

H$_0$: $\beta_1 = 0$
H$_1$: $\beta_1 \neq 0$ a ***two-sided* alternative hypothesis** ⇒ a ***two-tailed* t-test**

- Test statistic is $t(\hat{\beta}_1) = \dfrac{\hat{\beta}_1 - \beta_1}{\hat{se}(\hat{\beta}_1)} \sim t[N-2]$.
- $\hat{\beta}_1$ = **0.82068**; $\hat{se}(\hat{\beta}_1)$ = **0.060900**
- Calculate the ***sample value* of the t-statistic** under H$_0$: set β_1 = 0, $\hat{\beta}_1$ = **0.82068**, and $\hat{se}(\hat{\beta}_1)$ = **0.060900**.

$$t_0(\hat{\beta}_1) = \frac{\hat{\beta}_1 - \beta_1}{\hat{se}(\hat{\beta}_1)} = \frac{0.82068 - 0.0}{0.060900} = \frac{0.82068}{0.060900} = \mathbf{13.4759} = \underline{\mathbf{13.48}}$$

(3 marks)

ANSWER to Question 4(f) -- continued:

- **Null distribution** of $t_0(\hat{\beta}_1)$ is **t[N – 2] = t[2017 – 2] = t[2015]**

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

- ***reject* H₀** if $\left|t_0(\hat{\beta}_1)\right| > t_{\alpha/2}[2015]$,

  i.e., if either (1) $t_0(\hat{\beta}_1) > t_{\alpha/2}[2015]$ or (2) $t_0(\hat{\beta}_1) < -t_{\alpha/2}[2015]$;

- ***retain* H₀** if $\left|t_0(\hat{\beta}_1)\right| \le t_{\alpha/2}[2015]$,

  i.e., if $-t_{\alpha/2}[2015] \le t_0(\hat{\beta}_1) \le t_{\alpha/2}[2015]$.

Critical values of t[2015]-distribution: in t-table, use **df = ∞**.

- ***two-tailed*** **1 percent critical value** = $t_{\alpha/2}[2015] = t_{0.005}[\infty] = \underline{\mathbf{2.576}}$ **(1 mark)**

- **Inference:** **(1 mark)**

  - At **1 percent significance level**, i.e., for α = 0.01,

    $\left|t_0(\hat{\beta}_1)\right| = \mathbf{13.48} > \mathbf{2.576} = t_{0.005}[2015] \Rightarrow$ ***reject* H_0** vs. H_1 **at 1 percent level**.

- **Inference:** **At the 1% significance level**, the null hypothesis H_0: $\beta_1 = 0$ is ***rejected*** in favour of the alternative hypothesis H_1: $\beta_1 \ne 0$.

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

Rejection of the null hypothesis $\beta_1 = 0$ against the alternative hypothesis $\beta_1 \ne 0$ means that **the sample evidence favours the existence of a relationship between the *net financial wealth* and *annual income*** for single adults.

(8 marks)

(g) Perform an F-test of the proposition that annual income is unrelated to net financial wealth for single persons. 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.

- State the null hypothesis H_0 and the alternative hypothesis H_1. **(2 marks)**

  H_0: $\beta_1 = 0$

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

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

- $\hat{\beta}_1 = \mathbf{0.82068}$; $\hat{se}(\hat{\beta}_1) = \mathbf{0.060900}$; $\hat{Var}(\hat{\beta}_1) = \left(\hat{se}(\hat{\beta}_1)\right)^2 = (0.0609)^2 = \mathbf{0.00370881}$

- Calculate the ***sample value*** **of the F-statistic** under H_0: set $\beta_1 = 0$, $\hat{\beta}_1 = \mathbf{0.82068}$, and $\hat{Var}(\hat{\beta}_1) = \mathbf{0.00370881}$.

$$F_0(\hat{\beta}_1) = \frac{(\hat{\beta}_1 - \beta_1)^2}{\hat{Var}(\hat{\beta}_1)} = \frac{(0.82068 - 0.0)^2}{0.00370881} = \frac{0.6735157}{0.00370881} = \mathbf{181.5989} = \underline{\mathbf{181.60}}$$

(3 marks)

- **Null distribution** of $F_0(\hat{\beta}_1)$ is **F[1, N – 2] = F[1, 2017 – 2] = F[1, 2015]**

<u>Decision Rule</u>: At significance level α, **(1 mark)**

- ***reject* H_0** if $F_0(\hat{\beta}_1) > F_\alpha[1, 2015]$;

- ***retain* H_0** if $F_0(\hat{\beta}_1) \leq F_\alpha[1, 2015]$.

Critical values of F[1, 2015]-distribution: in F-table, use **denominator df = ∞**.

- **<u>1 percent</u> critical value** = $F_\alpha[1, 2015] = F_{0.01}[1, 2015] = F_{0.01}[1, \infty] = \underline{\mathbf{6.63}}$ **(1 mark)**

- **<u>Inference</u>:** At **1 percent significance level**, i.e., for $\alpha = 0.01$, **(1 mark)**

  $F_0(\hat{\beta}_1) = \mathbf{181.60 > 6.63} = F_{0.01}[1, 2015]$ $\Rightarrow$ ***reject* H_0** vs. H_1 **at 1 percent level**.

<u>Inference</u>: At the 1% significance level, the null hypothesis H_0: $\beta_1 = 0$ is ***rejected*** in favour of the alternative hypothesis H_1: $\beta_1 \neq 0$.

(8 marks)

(h) Perform a test of the proposition that single persons' net financial wealth is positively related to their annual income, i.e., that an increase in annual income *increases* the net financial wealth of single persons. 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.

<u>ANSWER to Question 4(h)</u>:

<u>Null and Alternative Hypotheses</u>:

H_0: $\beta_1 = 0$ **(1 mark)**

H_1: $\beta_1 > 0$ $\Rightarrow$ a ***right-tailed* t-test** **(1 mark)**

- Test statistic is $t(\hat{\beta}_1) = \dfrac{\hat{\beta}_1 - \beta_1}{\hat{se}(\hat{\beta}_1)} \sim t[N-2]$; $\hat{\beta}_1 = \mathbf{14.1219}$; $\hat{se}(\hat{\beta}_1) = \mathbf{5.36347}$.

- Calculate the ***sample value* of the t-statistic** under H_0: set $\beta_1 = 0$, $\hat{\beta}_1 = \mathbf{0.82068}$, and $\hat{se}(\hat{\beta}_1) = \mathbf{0.060900}$.

$$t_0(\hat{\beta}_1) = \frac{\hat{\beta}_1 - \beta_1}{\hat{se}(\hat{\beta}_1)} = \frac{0.82068 - 0.0}{0.060900} = \frac{0.82068}{0.060900} = \mathbf{13.4759} = \underline{\mathbf{13.48}}$$ **(2 marks)**

- **Null distribution** of $t_0(\hat{\beta}_1)$ is **t[N – 2] = t[2017– 2] = t[2015]**

<u>Decision Rule</u>: At significance level α, **(2 marks)**

- ***reject* $\mathbf{H_0}$** if $t_0(\hat{\beta}_1) > t_\alpha[2015]$,
- ***retain* $\mathbf{H_0}$** if $t_0(\hat{\beta}_1) \le t_\alpha[2015]$.

Critical value of t[2015]-distribution: from t-table, use **df = ∞**.

- ***right-tail* <u>1 percent</u> critical value** = $t_{0.01}[2015] = t_{0.01}[\infty]$ = <u>**2.326**</u> **(1 mark)**

<u>Inference</u>: At **1 percent significance level**, i.e., for $\alpha = 0.01$, **(1 mark)**

$t_0(\hat{\beta}_1) = \mathbf{13.48} > \mathbf{2.326} = t_{0.01}[2015] \Rightarrow$ ***reject* $\mathbf{H_0}$** vs. H_1 **at 1 percent level**.

♦ **<u>Inference</u>: At the 1% significance level**, the null hypothesis **H_0: $\beta_1 = 0$ is *rejected*** in favour of the alternative hypothesis **H_1: $\beta_1 > 0$**.

(12 marks)

(i) Compute the two-sided 95% confidence interval for the intercept coefficient β_0. Use this two-sided 95% confidence interval for β_0 to test the hypothesis that the mean net financial wealth of single adults whose annual income is zero dollars per year equals zero. State the null hypothesis H_0 and the alternative hypothesis H_1. State the decision rule you use, and the inference you would draw from the test.

ANSWER to Question 4(i):

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

$$\hat{\beta}_0 - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_0) \le \beta_0 \le \hat{\beta}_0 + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_0)$$ **(2 marks)**

where

- $\hat{\beta}_{0L} = \hat{\beta}_0 - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_0)$ = the ***lower*** **100(1 – α)% confidence limit for β_0**
- $\hat{\beta}_{0U} = \hat{\beta}_0 + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_0)$ = the ***upper*** **100(1 – α)% confidence limit for β_0**
- $t_{\alpha/2}[N-2]$ = the **α/2 critical value of the t-distribution with N–2 degrees of freedom**.

- Required results and intermediate calculations:

$N - 2 = 2017 - 2 = 2015$; $\hat{\beta}_0 = -10.571$; $\hat{se}(\hat{\beta}_0) = 2.0607$

$1 - \alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow$ **$\alpha/2 = 0.025$**: $t_{\alpha/2}[N-2] = \mathbf{t_{0.025}[2015] = 1.960}$
$= \mathbf{t_{0.025}[\infty] = 1.960}$

$t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_0) = t_{0.025}[855]\hat{se}(\hat{\beta}_0) = 1.960(2.0607) = \mathbf{4.038972}$

- **Lower 95% confidence limit for β_0** is: **(2 marks)**

$$\hat{\beta}_{0L} = \hat{\beta}_0 - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_0) = \hat{\beta}_0 - t_{0.025}[2015]\hat{se}(\hat{\beta}_0)$$
$$= -10.571 - \mathbf{1.960}(2.0607) = -10.571 - 4.038972 = -14.609972 = \underline{\mathbf{-14.610}}$$

- **Upper 95% confidence limit for β_0** is: **(2 marks)**

$$\hat{\beta}_{0U} = \hat{\beta}_0 + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_0) = \hat{\beta}_0 + t_{0.025}[2015]\hat{se}(\hat{\beta}_0)$$
$$= -10.571 + \mathbf{1.960}(2.0607) = -10.571 + 4.038972 = -6.532028 = \underline{\mathbf{-6.5320}}$$

ANSWER to Question 4(i) -- continued:

- **Result:** The **two-sided 95% confidence interval for β_0** is: **[–14.610, –6.5320]**

  $$-14.610 \le \beta_0 \le -6.5320$$

- **Null and Alternative Hypotheses:** **(2 marks)**

  H_0: $\beta_0 = 0$
  H_1: $\beta_0 \neq 0$ a ***two-sided* alternative hypothesis** $\Rightarrow$ a ***two-tailed* test**

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

- ***reject* H_0** if the ***hypothesized* value b_0 of β_0** specified by H_0 **lies *outside*** the two-sided (1–α)-level confidence interval for β_0, i.e., if either
  (1) $b_0 < \hat{\beta}_0 - t_{\alpha/2}[2015]\hat{se}(\hat{\beta}_0)$ or (2) $b_0 > \hat{\beta}_0 + t_{\alpha/2}[2015]\hat{se}(\hat{\beta}_0)$.

- ***retain* H_0** if the ***hypothesized* value b_0 of β_0** specified by H_0 **lies *inside*** the two-sided (1–α)-level confidence interval for β_0, i.e., if
  $\hat{\beta}_0 - t_{\alpha/2}[2015]\hat{se}(\hat{\beta}_0) \le b_0 \le \hat{\beta}_0 + t_{\alpha/2}[2015]\hat{se}(\hat{\beta}_0)$.

Inference: **(2 marks)**

- At **5 percent significance level**, i.e., for $\alpha = 0.05$,

  $$b_0 = \mathbf{0} > \mathbf{-6.5320} = \hat{\beta}_{0U} = \hat{\beta}_0 + t_{0.025}[2015]\hat{se}(\hat{\beta}_0)$$

  $\Rightarrow$ ***reject* H_0** vs. H_1 **at 5 percent level**.

- **Inference: At the 5% significance level**, the **null hypothesis H_0: $\beta_0 = 0$** is ***rejected*** in favour of the alternative hypothesis H_1: $\beta_0 \neq 0$.

(8 marks)

(j) Perform a test of the proposition that an increase in annual income of $1,000 per year is associated on average with an increase in net financial wealth of less than $1,000. Use the 5 percent significance level (i.e., $\alpha = 0.05$). 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 4(j):

Null and Alternative Hypotheses:

H_0: $\beta_1 = 1$ **(1 mark)**

H_1: $\beta_1 < 1$ ⇒ a ***left-tailed*** **t-test** **(1 mark)**

- Test statistic is $t(\hat{\beta}_1) = \dfrac{\hat{\beta}_1 - \beta_1}{\widehat{se}(\hat{\beta}_1)} \sim t[N-2]$; $\hat{\beta}_1 = \mathbf{14.1219}$; $\widehat{se}(\hat{\beta}_1) = \mathbf{5.36347}$.

- Calculate the ***sample value*** **of the t-statistic** under H_0: set $\beta_1 = 1$, $\hat{\beta}_1 = \mathbf{0.82068}$, and $\widehat{se}(\hat{\beta}_1) = \mathbf{0.060900}$.

$$t_0(\hat{\beta}_1) = \frac{\hat{\beta}_1 - \beta_1}{\widehat{se}(\hat{\beta}_1)} = \frac{0.82068 - 1.0}{0.0609} = \frac{-0.17932}{0.0609} = \mathbf{-2.94450} = \underline{\mathbf{-2.945}}$$ **(2 marks)**

- **Null distribution** of $t_0(\hat{\beta}_1)$ is $\mathbf{t[N-2] = t[2017-2] = t[2015]}$

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

- ***reject*** $\mathbf{H_0}$ if $t_0(\hat{\beta}_1) < -t_\alpha[2015]$,
- ***retain*** $\mathbf{H_0}$ if $t_0(\hat{\beta}_1) \geq -t_\alpha[2015]$.

Critical value of t[2015]-distribution: from t-table, use **df = ∞**.

- ***left-tail*** **<u>1 percent</u> critical value** $= -t_{0.05}[2015] = -t_{0.05}[\infty] = \underline{\mathbf{-1.645}}$ **(1 mark)**

Inference: At **5 percent significance level**, i.e., for $\alpha = 0.05$, **(1 mark)**

At **5 percent significance level**, i.e., for $\alpha = 0.05$,

$t_0(\hat{\beta}_1) = \mathbf{-2.03} < \mathbf{-1.65} = -t_{0.05}[855]$ ⇒ ***reject*** $\mathbf{H_0}$ vs. H_1 **at 5 percent level**.

♦ **Inference: At the 5% significance level**, the null hypothesis $\mathbf{H_0: \beta_1 = 1}$ **is *rejected*** in favour of the alternative hypothesis $\mathbf{H_1: \beta_1 < 1}$.

(4 marks)

(k) Use the estimation results given above for sample regression equation (3) to an estimate of the mean net financial wealth of single adults whose annual income is $50,000 per year. Show how you computed your estimate.

ANSWER to Question 3(k):

For ***income_i*** = 50, estimated mean net financial wealth is: **(2 marks)**

$$\hat{E}\left(\text{netfa}_i \middle| \text{income}_i = 50\right) = -10.571 + 0.82068(50) = -10.571 + 41.034 = \mathbf{30.463}$$

Estimated mean net financial wealth of single adults whose annual income is $50,000 per year equals **$30,463**. **(2 marks)**

Percentage Points of the t-Distribution

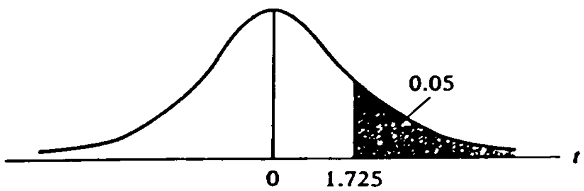
TABLE D.2
Percentage points of the t distribution

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*.

Source: Damodar N. Gujarati, *Basic Econometrics*, Third Edition. New York: McGraw-Hill, 1995, p. 809.

Selected Upper Percentage Points of the F-Distribution

TABLE D.3
Upper percentage points of the *F* distribution (*continued*)

df for denominator N_2	df for numerator N_1												
	Pr	1	2	3	4	5	6	7	8	9	10	11	12
	.25	1.40	1.48	1.47	1.45	1.44	1.42	1.41	1.40	1.39	1.39	1.38	1.37
22	.10	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.88	1.86
	.05	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23
	.01	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12
	.25	1.39	1.47	1.46	1.44	1.43	1.41	1.40	1.39	1.38	1.38	1.37	1.36
24	.10	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.85	1.83
	.05	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.21	2.18
	.01	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03
	.25	1.38	1.46	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.37	1.36	1.35
26	.10	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.84	1.81
	.05	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
	.01	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	3.02	2.96
	.25	1.38	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34
28	.10	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.81	1.79
	.05	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12
	.01	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.96	2.90
	.25	1.38	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.36	1.35	1.35	1.34
30	.10	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.79	1.77
	.05	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09
	.01	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84
	.25	1.36	1.44	1.42	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.32	1.31
40	.10	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.73	1.71
	.05	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00
	.01	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66
	.25	1.35	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.29
60	.10	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.68	1.66
	.05	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
	.01	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50
	.25	1.34	1.40	1.39	1.37	1.35	1.33	1.31	1.30	1.29	1.28	1.27	1.26
120	.10	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.62	1.60
	.05	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.87	1.83
	.01	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.40	2.34
	.25	1.33	1.39	1.38	1.36	1.34	1.32	1.31	1.29	1.28	1.27	1.26	1.25
200	.10	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.60	1.57
	.05	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80
	.01	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.27
	.25	1.32	1.39	1.37	1.35	1.33	1.31	1.29	1.28	1.27	1.25	1.24	1.24
∞	.10	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.57	1.55
	.05	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75
	.01	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18

Source: Damodar N. Gujarati, *Basic Econometrics*, Third Edition. New York: McGraw-Hill, 1995, p. 814.