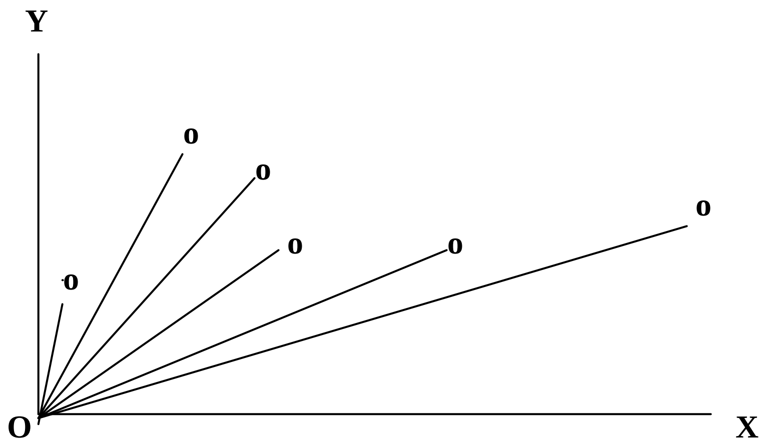
Econ 120B **Ramu Ramanathan**
Fall 2003 **First Midterm Answers**

I. Suppose the true model is $Y_t = b X_t + u_t$.

I.a (3 points)

An estimate of b is derived as follows: In the scatter diagram for X and Y given below, draw a straignt line from the origin to each of the points (X_1, Y_1), (X_2, Y_2), . . . , (X_n, Y_n).



I.b (5 points)

Then compute the average (b*) of the slopes of these lines. Write down an algebraic expression for b*. Do this carefully because subsequent answers depend on your getting this right.

The slope of a straight line from the origin to a typical point (X_t, Y_t) is Y_t/X_t.

The average of these points is $b^* = \frac{1}{n}\sum_{t=1}^{n}\frac{Y_t}{X_t}$.

I.c (7 points)

Compute the expected value of b* and state whether or not it is an unbiased estimator of b. Be sure to state any assumptions you made.

$E(b^*) = E\left[\frac{1}{n}\sum_{t=1}^{n}\frac{Y_t}{X_t}\right] = \frac{1}{n}\sum_{1}^{n}E\left(\frac{Y_t}{X_t}\right)$. $E\left(\frac{Y_t}{X_t}\right) = \frac{1}{X_t}E(Y_t)$ *by the assumption that the Xs are given and nonrandom.* $E(Y_t) = E(bX_t + u_t) = b X_t$ *by the assumption that* $E(u_t) = 0$. *Substituting this in the expression for the expected value, we get,* $E(b^*) = \frac{1}{n}\Sigma\ b = b$. *Hence* b^* *is unbiased.*

I.d (5 points)

Without formal derivations, argue why b^* is inferior to the OLS estimator for the above model (you need not apply OLS). State any properties that enable you to make the assertion.

By the Gauss-Markov Theorem, OLS gives linear estimators that are BLUE. This means that any other estimator, such as b^*, has a higher variance and is hence inferior.

II.

Consider the following two models of the expenditures for maintenance of a certain automobile:

$$E_t = a_t + b_1 \text{ Miles}_t + u_t$$
$$E_t = a_2 + b_2 \text{ Age}_t + u_t$$

where E is the cumulative expenditure on maintenance (excluding gasoline), in dollars, Miles is the cumulative number of miles driven (in thousands), and Age is the age in weeks. Using 57 time series observations, the two models were estimated and the partial computer output is reproduced here (data in DATA3-7).

<u>Model A</u>

Variable	Coefficient	Standard Error
Constant	-625.935025	104.149581
Age	7.343478	0.32958

Error Sum of Sq (ESS) 7.401653e + 06 Std Err of Resid. (sgmahat) 366.845346
R-squared 0.900

<u>Model B</u>

Variable	Coefficient	Standard Error
Constant	-796.074573	134.74494
Miles	53.450724	2.926144

Error Sum of Sq (ESS) 1.050175e + 07 Std Err of Resid. (sgmahat) 436.96796
R-squared 0.858

IIa. (5 points)

What signs would you expect for b_1 and b_2? Do the observed signs agree with your expectation?

As a car becomes older, expenses for maintenance will increase. Hence b_1 will be greater than zero. As a car is driven a lot, expenses for maintenance will increase. Hence b_1 will be greater than zero also.

IIb. (2 points)
Which of the two models do you think is "better"? Clearly state the criteria you used.

Model A is better because it has a higher R^2.

IIc. (13 points)
In the better model you chose, perform appropriate tests for the significance (two -sided) of each of the regression coefficients at the 1 percent level. Be sure to state the null and alternative hypotheses for each coefficient, the distribution of the test statistic including d.f., and your criteria for rejecting or not rejecting the null. What do you conclude?

First, H_0: $a_1 = 0$ $\qquad$ H_1: $a_1 \neq 0$

Compute $|t_c| = \dfrac{625.035025}{104.149581} = 6.01$. *Under the null hypothesis, this has the t-distribution with n-k = 57-2 = 55 d.f. From the t-table, look up the critical value (t*) for 55 d.f. and 1 percent level. It is in the range (2.66 – 2.704). Since* $|t_c| > t^*$, *we reject the null hypothesis and conclude that* a_1 *is significantly different from zero.*

Next, H_0: $b_1 = 0$ $\qquad$ H_1: $b_1 \neq 0$

Compute $|t_c| = \dfrac{7.343478}{0.32958} = 22.28$. *Under the null hypothesis, this also has the t-distribution with n-k = 57-2 = 55 d.f. As before, from the t-table, t* is in the range (2.66 – 2.704). Since* $|t_c| > t^*$, *we reject the null hypothesis and conclude that* b_1 *is also significantly different from zero.*

IId. (10 points) In Model A suppose Age is measured in days (call it AGE*) rather than weeks. Rewrite the table. In the side, show your work.

Variable	Coefficient	Standard Error
Constant	**- 625.935025**	**104.149581**
AGE*	**1.049068**	**0.04708**

Let AGE be age in days.*
AGE = 7 AGE. AGE = AGE*/7.*
The model, in symbols, is,

$$EXPENSE = a_1 + b_1 \frac{AGE^*}{7} + u$$

$$= a_1 + b_1^* AGE^* + u$$

Only change is in $\hat{b}_1$ *and its standard error.*

Error Sum of Sq (ESS) **7.401653e+06**

***R*-squared** **0.900** **Std Err of Resid. (sgmahat)** **366.845346**