1045/16

This question paper contains **16+4** printed pages **+7** Pages of Statistical Tables Attached]

Roll No. | | | | | | | | | | | |

S. No. of Question Paper : **3312**

Unique Paper Code : **2271403** **F-6**

Name of the Paper : **Introductory Econometrics**

Name of the Course : **B.A. (Hons.) Economics**

Semester : **IV**

Duration : **3 Hours** Maximum Marks : **75**

(Write your Roll No. on the top immediately on receipt of this question paper.)

(इस प्रश्न-पत्र के मिलते ही ऊपर दिए गए निर्धारित स्थान पर अपना अनुक्रमांक लिखिए ।)

Note : Answers may be written *either* in English *or* in Hindi; but the same medium should be used throughout the paper.

टिप्पणी : इस प्रश्न-पत्र का उत्तर अंग्रेज़ी **या** हिन्दी किसी एक भाषा में दीजिए; लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए ।

The question paper consists of *seven* questions. Attempt any *five* questions.

Each question carries **15** marks. Use of simple non-programmable calculator is allowed.

Statistical tables are attached for your reference.

इस प्रश्न-पत्र में **7** प्रश्न हैं। किन्हीं **पाँच** प्रश्नों के उत्तर दीजिये।

प्रत्येक प्रश्न **15** अंक का है। सरल अप्रोग्रामीय कैलकुलेटर का प्रयोग अनुमोदित है।

संदर्भ के लिए सांख्यिकीय सारणियाँ संलग्न हैं।

P.T.O.

1. Are the following statements correct ? Justify your answers carefully and provide proofs wherever necessary :

   (*a*) The value of $\bar{R}^2$ is always greater than R^2.

   (*b*) If we multiply both Y and X by 100 and re-estimate the regression, the slope coefficient will get multiplied by 100.

   (*c*) In the double log model, the elasticity is variable while the slope $\left(\frac{dy}{dx}\right)$ is constant. But for the linear in variables model, the slope $\left(\frac{dy}{dx}\right)$ is variable and the elasticity is constant.

   (*d*) In case of exact multicollinearity, you would not be able to calculate the regression coefficients of a multiple regression model.

   (*e*) $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ is estimated as $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$. Here β_1 is a random variable, ε_i is a constant and $\hat{\beta}_1$ is unknown. 3×5

   क्या निम्नलिखित कथन सही हैं ? अपने उत्तरों के लिए सावधानीपूर्वक तर्क दीजिए व जहाँ आवश्यक हो वहाँ प्रमाण भी दीजिए :

   (*a*) $\bar{R}^2$ का मान हमेशा R^2 से अधिक होता है।

   (*b*) यदि हम Y व X दोनों को 100 से गुणा करें व समाश्रयण को दुबारा आकलित करें, तो ढाल गुणांक 100 गुना हो जाएगा।

   (*c*) Double log मॉडल में चर की लोच परिवर्तनशील होती है जबकि ढाल $\left(\frac{dy}{dx}\right)$ स्थिर होता है। परन्तु चरों में रेखीय मॉडल में ढाल $\left(\frac{dy}{dx}\right)$ परिवर्तनशील होता है जबकि लोच स्थिर होती है।

(*d*) यथार्थ (exact) बहुसंरेखता की स्थिति में आप एक बहु-समाश्रयण मॉडल (multiple regression model) के समाश्रयण गुणांकों की गणना नहीं कर पाएंगे।

(*e*) $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ को आकलित करने पर $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ प्राप्त होता है। यहाँ β_1 एक यादृच्छिक चर है, ε_i एक स्थिरांक है तथा $\hat{\beta}_1$ अज्ञात है।

2. (*a*) Explain what is meant by dummy variable trap.

(*b*) For a project work in econometrics, suppose two students A and B together have collected data on savings, income and household size for 23 households in Delhi. Then they considered two different regression models using same data set.

Student A : $\ln Sav_i = \beta_1 + \beta_2 \ln Inc_i + \beta_3 Hhs_i + \beta_4 \ln Mexp_i + u_i$

But Student B imposed the restrictions $\beta_2 = \beta_3 = 0$ and estimated the following model

Student B : $\ln Sav_i = \alpha_1 + \alpha_2 \ln Mexp_i + v_i$

where

ln : natural log

Sav : Annual savings in lakhs of rupees

Inc : Annual disposable income in lakhs of rupees

Hhs : Size of the household (number of family members)

Mexp : Annual medical expenditure in lakhs of rupees.

The regression results are given below :

Dependent Variable	Intercept coefficient	lnInc coefficient	Hhs coefficient	lnMexp coefficient	R^2	TSS
lnSav	1.422 (.876)	0.231 (.058)	–0.379 (.14)	–0.021 (.03)	0.8752	37.949
lnSav	4.442 (0.561)			–0.23 (.04)	0.6123	37.949

Standard errors are written in parentheses

(*i*) Draw the ANOVA table for the multiple regression equation of student A.

(*ii*) Test the overall significance of the multiple regression equation of student A at 5% level of significance.

(*iii*) Test the validity of the restrictions imposed by student B on student A's model at 5% level of significance. 3,4+4+4

(*a*) डमी चर जाल (dummy variable trap) क्या है ? समझाइए।

(*b*) मान लीजिए कि दो छात्रों ने मिलकर अर्थमिति के परियोजना कार्य के लिए दिल्ली के 23 परिवारों के लिए बचत, आय व परिवार के आकार पर आँकड़े एकत्र किए। इसके बाद उन्होंने उन्हीं आँकड़ों की सहायता से दो अलग-अलग समाश्रयण मॉडलों पर विचार किया।

छात्र A : $\ln Sav_i = \beta_1 + \beta_2 \ln Inc_i + \beta_3 Hhs_i + \beta_4 \ln Mexp_i + u_i$

परन्तु छात्र B ने प्रतिबन्ध $\beta_2 = \beta_3 = 0$ लगाए व निम्नलिखित मॉडल आकलित किया :

छात्र B : $\ln Sav_i = \alpha_1 + \alpha_2 \ln Mexp_i + v_i$

जहाँ

ln : प्राकृतिक लघुगणक

Sav : वार्षिक बचत लाख रुपयों में

Inc : वार्षिक प्रयोज्य आय लाख रुपयों में

Hhs : परिवार का आकार (परिवार के सदस्यों की संख्या)

Mexp : स्वास्थ्य पर वार्षिक व्यय लाख रुपयों में

समाश्रयण परिणाम नीचे दिए गए हैं :

निर्भर चर	अन्तःखण्ड गुणांक	lnInc गुणांक	Hhs गुणांक	lnMexp गुणांक	R^2	TSS
lnSav	1.422 (.876)	0.231 (.058)	–0.379 (.14)	–0.021 (.03)	0.8752	37.949
lnSav	4.442 (0.561)			–0.23 (.04)	0.6123	37.949

मानक त्रुटियाँ कोष्ठकों में लिखी हुई हैं।

(*i*) छात्र A के बहु-समाश्रयण मॉडल हेतु ANOVA सारणी बनाइए।

(*ii*) छात्र A के बहु-समाश्रयण समीकरण की सम्पूर्ण सार्थकता का 5% सार्थकता स्तर पर परीक्षण कीजिए।

(*iii*) छात्र B द्वारा छात्र A के मॉडल पर लगाए गए प्रतिबन्धों की वैधता का परीक्षण 5% सार्थकता स्तर पर कीजिए।

3. (*a*) Write the assumptions of CLRM for a two variable case. Which assumption of the CLRM is required for performing hypothesis testing ? Discuss the rationale of this assumption.

(*b*) Consider the following estimated regression equation based on data collected in 2014 for 365 consumers in a country :

$$\widehat{\text{ExpCar}_i} = 26 - 260\left(1/\text{Inc}_i\right)$$

Where

ExpCar : Expenditure (in lakhs) on Cars

Inc : Disposable Income in lakhs

(*i*) Interpret the intercept value

(*ii*) Find out the threshold level of disposable income below which a car is not purchased. Also draw the estimated regression function on the income-expenditure plane.

(*iii*) What is the elasticity of car expenditure with respect to disposable income, at the point ? (Inc = 15, ExpCar = 9)

(*c*) Explain how the Park test for heteroscedasticity formalises the graphical test of examining a residual plot. For a particular case the null hypothesis of the Park test was not rejected. Does this imply that heteroscedasticity is not a problem ? 5,1+3+2,4

(*a*) दो चर वाली स्थिति के लिए CLRM की मान्यताएँ लिखिए। परिकल्पनाओं के परीक्षण हेतु CLRM की किस मान्यता की आवश्यकता होती है ? इस मान्यता के औचित्य का विवेचन कीजिए।

(*b*) 2014 में किसी देश के 365 उपभोक्ताओं हेतु संग्रहित आँकड़ों के आधार पर आकलित निम्नलिखित समाश्रयण समीकरण पर विचार कीजिए :

$$\widehat{\text{ExpCar}_i} = 26 - 260\left(1/\text{Inc}_i\right)$$

जहाँ

ExpCar : कारों पर व्यय (लाखों में)

Inc : प्रयोज्य आय लाखों में

(*i*) अन्तःखण्ड के मान की व्याख्या कीजिए।

(*ii*) प्रयोज्य आय का वह स्तर ज्ञात कीजिए जिसके नीचे कार नहीं खरीदी जाती है। आकलित समाश्रयण फलन का आय-व्यय समतल में आरेख भी बनाइए।

(*iii*) कार पर व्यय की प्रयोज्य आय के सापेक्ष लोच क्या है ? (Inc = 15, ExpCar = 9)

(*c*) प्रसरण-विषमता (heteroscedasticity) का पार्क परीक्षण (Park's test) अवशिष्टों के चित्र (residual plot) के परीक्षण का औपचारिक रूप कैसे है, समझाइए। किसी दी गई स्थिति में पार्क के परीक्षण ने शून्य परिकल्पना को अस्वीकार नहीं किया। क्या इसका तात्पर्य यह है कि प्रसरण-विषमता की समस्या विद्यमान नहीं थी ?

4. (*a*) Data were collected on 344 corporate executives to find out the effect of MBA degree and work experience on their salary. The following model was estimated :

$$\widehat{\ln Sal}_i = 2.3501 + 3.6306\ D_{2i} + 2.6354\ D_{3i} + 0.1527\ WE_i + 0.234\ (D_{3i} * WE_i)$$

se = (1.86) (1.67) (0.762) (0.02) (0.09)

$R^2 = 0.8968$

ln : Natural log

Sal : Annual gross salary in Lakhs of Rupees

WE : Work exprience in years

D_{2i} = 1 if one has MBA degree

= 0 otherwise

D_{3i} = 1 for a male executive

= 0 for a female executive

(*i*) Write the regression equations for female MBA executives and male non MBA executives separately.

(*ii*) Find the relative change in salary due to one year increase in work experience for male and female executives. Do they differ statistically at 5% level of significance ?

(*iii*) Now suppose out of this sample of executives, 48 are female MBA executives and 156 are male MBA executives. To find out the relation between income earned and work experience, we run three regressions.

Regression A : for 156 male MBA executives :

$$\ln \widehat{Inc_i} = \hat{a}_i + \hat{b}_i \ WE_i \qquad RSS^A = 3.701$$

Regression B : for 48 female MBA executives :

$$\ln \widehat{Inc_i} = \hat{\alpha}_i + \hat{\beta}_i \ WE_i \qquad RSS^B = 4.803$$

Regression C : with 204 (156 male + 48 female) MBA executives :

$$\ln \widehat{Inc_i} = \hat{\gamma}_i + \hat{\delta}_i \ WE_i \qquad RSS^C = 9.7602$$

Using the above data, at 10% level of significance, check whether statistically there is an improvement in doing a Regression C as compared to other two subsample regressions.

(*b*) Illustrate with an example how with the use of extraneous information one can control the problem of multicollinearity. 2+6+3,4

P.T.O.

(*a*) MBA व कार्यानुभव के वेतनों पर प्रभाव ज्ञात करने हेतु निगमों के 344 कार्यकारी अधिकारियों से आँकड़े संग्रहित किए गए। निम्नलिखित मॉडल आकलित किया गया :

$$\widehat{\ln Sal}_i = 2.3501 + 3.6306\ D_{2i} + 2.6354\ D_{3i} + 0.1527\ WE_i + 0.234\ (D_{3i} * WE_i)$$

, se = (1.86) (1.67) (0.762) (0.02) (0.09)

$R^2 = 0.8968$

ln : प्राकृतिक लघुगणक

Sal : वार्षिक समग्र वेतन लाखों रुपयों में

WE : कार्यानुभव वर्षों में

D_{2i} = 1 यदि किसी के पास MBA डिग्री है

= 0 अन्यथा

D_{3i} = 1 पुरुष अधिकारी हेतु

= 0 महिला अधिकारी हेतु

(*i*) महिला MBA अधिकारियों व पुरुष गैर-MBA अधिकारियों हेतु अलग-अलग समाश्रयण समीकरण लिखिए।

(*ii*) पुरुष व महिला अधिकारियों हेतु कार्यानुभव में एक वर्ष की वृद्धि के कारण वेतन में होने वाले सापेक्ष परिवर्तन ज्ञात कीजिए। क्या ये 5% सार्थकता स्तर पर सांख्यिकीय रूप से भिन्न हैं ?

(*iii*) अब मान लीजिए कि अधिकारियों के इस प्रतिदर्श में से 48 महिला MBA अधिकारी हैं व 156 पुरुष MBA अधिकारी हैं। अर्जित आय व कार्यानुभव के मध्य सम्बन्ध ज्ञात करने हेतु हमने तीन समाश्रयण किए।

समाश्रयण A : 156 पुरुष MBA अधिकारियों हेतु :

$$\widehat{Inc}_i = \hat{a}_i + \hat{b}_i \ WE_i \qquad RSS^A = 3.701$$

समाश्रयण B : 48 महिला MBA अधिकारियों हेतु :

$$\widehat{Inc}_i = \hat{\alpha}_i + \hat{\beta}_i \ WE_i \qquad RSS^B = 4.803$$

समाश्रयण C : 204 (156 पुरुष + 48 महिला) MBA अधिकारियों हेतु :

$$\widehat{Inc}_i = \hat{\gamma}_i + \hat{\delta}_i \ WE_i \qquad RSS^C = 9.7602$$

उपर्युक्त आँकड़ों की सहायता से 10% सार्थकता स्तर पर एक परीक्षण कीजिए, यह देखने के लिए कि क्या समाश्रयण C को करना, दोनों उप-प्रतिदर्शों के समाश्रयणों को करने के मुकाबले सांख्यिकीय रूप से बेहतर है।

(*b*) बाहरी सूचना की सहायता से किस प्रकार बहुसंरेखता की समस्या को नियन्त्रित किया जा सकता है, एक उदाहरण की सहायता से समझाइए।

5. (*a*) Prove that the slope coefficient from a regression of Y_i on X_i is the same as the slope coefficient from a regression of y_i on x_i, where y_i and x_i are deviations of Y_i and X_i from their respective means.

(*b*) Consider the following regression output :

$\hat{Y}_i = 0.2033 + 0.6560X_i$

se = (0.0976) (0.1961)

RSS = 0.0544

where

Y = expenditure on food in thousands of Rupees

X = income in thousands of Rupees.

The regression results are obtained from a sample of 19 households. Suppose the income (X_0) is 0.58. On the basis of the regression results given above, what is the mean expenditure on food ? Establish a 95% confidence interval for the predicted mean expenditure on food, given that $\bar{X} = 0.5$.

(*c*) A researcher estimated the following demand function for money

$\widehat{\ln M_t} = 2.6027 - 0.4024 \ln R_t + 0.59 \ln Y_t + 0.524 \ln M_{t-1}$

se = (1.24) (0.36) (0.34) (.0992)

$R^2 = 0.92$ $d = 1.23$ $n = 41$

ln : Natural log

M_t : real cash balances

R_t : long term interest rate

Y_t : aggregate real national income

(*i*) Test at 10% significance the presence of first order autocorrelation (conduct a two-tailed test).

(*ii*) In the light of the above result, is the researcher correct in using the *t*-test to check the impact of long term interest on demand for money ?

(*iii*) Based on the information given above, can you conduct the Breusch Godfrey test to check for first order autocorrelation ? Give reasons. 2,6,5+1+1

(*a*) सिद्ध कीजिए कि Y_i के X_i पर समाश्रयण में ढाल गुणांक वही होगा जो कि y_i के x_i पर समाश्रयण में होगा, जहाँ y_i व x_i, Y_i व X_i के अपने-अपने माध्यों से विचलन हैं।

(*b*) निम्नलिखित समाश्रयण परिणामों पर विचार कीजिए :

$\hat{Y}_i = 0.2033 + 0.6560X_i$

se = (0.0976) (0.1961)

RSS = 0.0544

जहाँ

Y = भोजन पर व्यय, हज़ारों रुपयों में

X = आय, हज़ारों रुपयों में

19 परिवारों के एक प्रतिदर्श से समाश्रयण परिणाम प्राप्त किये जाते हैं। मान लीजिए कि आय (X_0) 0.58 है। उपर्युक्त समाश्रयण परिणामों के आधार पर भोजन पर माध्य व्यय क्या है ? माध्य के पूर्वकथित मान हेतु 95% विश्वास्यता अन्तराल ज्ञात कीजिए, यदि $\bar{X} = 0.5$.

(*c*) एक शोधकर्ता ने मुद्रा के लिए निम्नलिखित मांग फलन आकलित किया :

$$\widehat{\ln M_t} = 2.6027 - 0.4024 \ln R_t + 0.59 \ln Y_t + 0.524 \ln M_{t-1}$$

se = (1.24) (0.36) (0.34) (.0992)

$R^2 = 0.92$ $d = 1.23$ $n = 41$

ln : प्राकृतिक लघुगणक

M_t : वास्तविक नकद शेष

R_t : दीर्घकालीन ब्याज दर

Y_t : समग्र वास्तविक राष्ट्रीय आय

(*i*) 10% सार्थकता स्तर पर प्रथम क्रम के स्वसहसम्बन्ध की उपस्थिति हेतु परीक्षण कीजिए। (द्वि-पुच्छीय (two-tailed) परीक्षण कीजिए)।

(*ii*) उपर्युक्त समाश्रयण परिणामों के परिप्रेक्ष्य में क्या शोधकर्ता द्वारा मुद्रा की मांग पर दीर्घकालीन ब्याज दर के प्रभाव के परीक्षण हेतु *t*-परीक्षण का उपयोग किया जाना सही है ?

(*iii*) उपर्युक्त परिणामों के आधार पर क्या आप प्रथम क्रम से स्वसहसम्बन्ध के परीक्षण हेतु ब्रूश-गॉडफ्री परीक्षण (Breusch-Godfrey test) कर सकते हैं ? कारण दीजिए।

6. (*a*) Let the population regression function be :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

How will you transform the model to obtain homoscedastic errors under each of the following cases, assuming all other CLRM assumptions for u_i hold ? Do the transformed regressions in each of the cases have an intercept term ?

(*i*) $u_i = \varepsilon_i(X_{2i})$

(*ii*) $E(u_i^2) = \sigma^2(X_{3i})^{1/2}$

It is given that $\varepsilon_i \sim N$ (mean = 0, variance = σ^2).

(*b*) From the annual data for the manufacturing sector for 1899-1992 the following regression results were obtained :

$$\widehat{\ln Y} = 2.81 - 0.53 \ln K + 0.61 \ln L + 0.047\, t \quad \text{.....(1)}$$

se = (1.38) (0.44) (0.14) (0.006)

$R^2 = 0.97$ P-value of F test = 0.000

Where

ln : Natural log

Y = index of real output

K = index of real capital input

L = index of labor input

t = time or trend.

Using the same data, he also obtained the following regression result :

$$\widehat{\ln\left(\frac{Y}{L}\right)} = -0.11 + 0.18 \ln\left(\frac{K}{L}\right) + 0.006\, t \quad \text{.....(2)}$$

se = (0.03) (0.05) (0.002)

$R^2 = 0.95$ P-value of F test = 0.000

(*i*) In regression (1), what are the a priori signs of the coefficients log K and log L ?

Do the results conform to this expectation ?

(*ii*) Is there multicollinearity in regression (1) ? How do you know ?

(*iii*) In regression (2), interpret the coefficient of $\ln\left(\frac{K}{L}\right)$.

(*iv*) Do you suspect multicollinearity in second regression ? Why or why not ?

3+3,4+2+3

(*a*) मान लीजिए कि समष्टि समाश्रयण फलन निम्न प्रकार है :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$