

Lib

03/12/18

Morning

This question paper contains 16+4 printed pages + 7 attached chart]

Roll No.

--	--	--	--	--	--	--	--	--	--

S. No. of Question Paper : 7945

Unique Paper Code : 12277502

IC

Name of the Paper : Applied Econometrics

Name of the Course : B.A. (Honours) Economics CBCS DSE-I

Semester : V

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 अंग्रेजी या हिंदी किसी एक भाषा में दीजिए परन्तु सभी उत्तरों का माध्यम एक ही होना चाहिए।

The question paper consists of seven questions.

Answer any *five* questions.

Marks allotted to each question are indicated against it.

Use of simple non-programmable calculator is allowed.

Statistical tables are attached for your reference.

इस प्रश्न-पत्र में सात प्रश्न हैं।

किन्हीं पाँच प्रश्नों के उत्तर दीजिए।

प्रत्येक प्रश्न को आवंटित अंक उसके सामने दिए गए हैं।

साधारण अप्रोग्रामनीय कैलकुलेटर का उपयोग किया जा सकता है।

आपके सन्दर्भ हेतु सांख्यिकीय सारणियाँ संलग्न हैं।

P.T.O.

1. State whether the following statements are true or false. Give reasons for your answer :

(a) For the three-variable model :

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

if the coefficient of determination for the regression of X_2 on X_3 is equal to 0.25, tolerance (TOL) will be 0.5 and it indicates presence of multicollinearity.

(b) In the simple regression model :

$$Y_i = \alpha + \beta X_i + u_i,$$

suppose X and u are correlated and Z is an instrumental variable (IV) for X . Then, higher is the value of R_{xz}^2 , larger is the standard error of the IV estimator of β and wider is the confidence interval for β .

(c) If there are five regions in a country—South, North, West, East and Central, then only four dummy variables must be included in the regression model to account for regional variations.

(d) Suppose there are $k = 6$ candidate regressors. If value of Mallows's C_p criterion is obtained as 6.456 for Model A with four regressors and 5.105 for Model B with five regressors, then Model B is preferred over Model A.

(e) Durbin Watson d statistic can also be used to test model specification errors. $5 \times 3 = 15$

बताइए कि निम्नलिखित कथन सत्य हैं अथवा असत्य। अपने उत्तर हेतु कारण भी दीजिए :

(a) तीन चरों वाले मॉडल :

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

में यदि X_2 के X_3 पर समाश्रयण में निर्धारण गुणांक (coefficient of determination) 0.25 है, तो सह्यता (tolerance, TOL) 0.5 होगी तथा यह बहुसंरेखता (multicollinearity) की उपस्थिति का संकेत देती है।

(b) सरल समाश्रयण मॉडल :

$$Y_i = \alpha + \beta X_i + u_i$$

में मान लीजिए कि X व u सहसम्बन्धित हैं तथा Z , X हेतु एक यांत्रिक चर (instrumental variable, IV) है। इस स्थिति में R_{xz}^2 का मान जितना अधिक होगा, β के IV आकलन का मानक विचलन उतना ही अधिक होगा तथा β हेतु विश्वास्यता अन्तराल (confidence interval) भी उतना ही चौड़ा होगा।

- (c) यदि एक देश में पाँच क्षेत्र हैं—दक्षिण, उत्तर, पश्चिम, पूर्व तथा केन्द्रीय, तो मॉडल में क्षेत्रीय विचरण (variation) को शामिल करने हेतु केवल चार मूक (dummy) चरों को शामिल किया जाना चाहिए।
- (d) मान लीजिए कि आपके पास $k = 6$ प्रत्याशी समाश्रयक (regressors) हैं। यदि मैलो के C_p मापदण्ड (Mallows's C_p criterion) का मान चार चरों वाले मॉडल A हेतु 6.456 प्राप्त होता है तथा पाँच चरों वाले मॉडल B हेतु 5.105 प्राप्त होता है तो मॉडल A की अपेक्षा मॉडल B बेहतर है।
- (e) डर्बिन-वॉट्सन के d प्रतिदर्शज का उपयोग मॉडल विनिर्देशन त्रुटियाँ (model specification errors) का परीक्षण करने हेतु भी किया जा सकता है।

2. (a) In the three-variable regression model,

$$y = X\beta + u, i = 1, 2, \dots, n,$$

given that

$$\hat{\beta} = (X'X)^{-1}X'y$$

(i) Show that $\hat{\beta}$ is an unbiased estimator of β .

- (ii) Show that $\text{var} - \text{cov}(\hat{\beta}) = \sigma^2 (X' X)^{-1}$. What is the dimension of the var-cov matrix of $\hat{\beta}$?
- (iii) Write down the correlation matrix R. How many zero order correlation coefficients are there ?
- (b) (i) What a priori restrictions are imposed on the β coefficients in the Koyck model ?
- (ii) Suppose consumption (C) depends on current and past values of GDP (X). Using annual data for a country for the period 1987-2013, the Koyck estimation results are given below :

$$\hat{C}_t = 405.1740 + 0.4004 X_t + 0.5000 C_{t-1}$$

$$p\text{-values} = (0.0001) \quad (0.00001) \quad (0.00000)$$

$$R^2 = 0.9911$$

What is the estimated long-run consumption function ? Write down the short-run and long-run estimated MPC ? Interpret them. Compute mean and median lag for the above consumption function.

- (a) तीन चरों वाले समाश्रयण मॉडल

$$y = X\beta + u, \quad i = 1, 2, \dots, n,$$

में यह दिया हुआ है कि

$$\hat{\beta} = (X'X)^{-1}X'y$$

- (i) दर्शाइए कि $\hat{\beta}$, β का एक अनभिनत (unbiased) आकलक है।
- (ii) दर्शाइए कि $\text{var} - \text{cov}(\hat{\beta}) = \sigma^2 (X' X)^{-1}$ । इस मैट्रिक्स का क्या आकार (dimension) है ?
- (iii) सहसम्बन्ध मैट्रिक्स R लिखिए। इसमें शून्य क्रम (zero order) के कितने सहसम्बन्ध गुणांक हैं ?

(b) (i) कॉएक मॉडल में β गुणांकों पर क्या पूर्व-प्रतिबन्ध (a priori restrictions) लागू होते हैं ?

(ii) मान लीजिए कि उपभोग (C), वर्तमान व पिछली अवधियों के GDP (X) पर निर्भर करता है। एक देश हेतु 1987-2013 के वार्षिक आँकड़ों की सहायता से कॉएक आकलन (Koyck estimation) से निम्नलिखित परिणाम प्राप्त हुए :

$$\hat{C}_t = 405.1740 + 0.4004 X_t + 0.5000 C_{t-1}$$

$$p\text{-values} = (0.0001) \quad (0.00001) \quad (0.00000)$$

$$R^2 = 0.9911$$

आकलित दीर्घकालीन (long-run) उपभोग फलन क्या है ? आकलित लघुकालीन व दीर्घकालीन MPC लिखिए। इनको समझाइए। उपर्युक्त उपभोग फलन हेतु माध्य (mean) व माध्यिका (median) विलम्बन (lag) ज्ञात कीजिए।

3. (a) (i) Suppose the true model is :

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

and the model fitted is :

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + v_i$$

What are $E(\hat{\alpha}_1)$, $E(\hat{\alpha}_2)$ and $E(\hat{\alpha}_3)$ equal to ? Are the OLS estimators of the fitted model unbiased ? Also compare the variances of $\hat{\alpha}_2$ and $\hat{\beta}_2$. Is $\hat{\alpha}_2$ efficient ? Explain.

(ii) When is a data point said to exert high 'leverage' ? If a data point exerts high leverage, is it always an 'influential point' ? Justify.

7

P.T.O.

- (b) A researcher studied quarterly total cost (TC) function for eight mobile-phone manufacturing companies for 10 years. The results are reported below :

Dependent variable: TC			
Method : EGLS			
Used sample 2004-2013			
Included observations : 40			
Number of cross-sections : 8			
Total panel observations : 320			
Variable	Coefficient	Std. Error	t-Statistic
C	47.30772	1.340279	35.29692
RM	0.523554	0.012030	43.52132
ADV	2.220745	0.149031	14.90118
Random Effects			
Company 1		0.258081	
Company 2		-2.415602	
Company 3		0.848119	
Company 4		-1.775884	
Company 5		1.190163	
Company 6		-1.573142	
Company 7		0.472518	
Company 8		2.995747	

where RM = cost incurred on raw material and ADV = cost incurred on advertisement.

- (i) What is the mean value of all the cross-sectional intercepts ?
- (ii) Interpret the random effect values of Company 2 and Company 3.
- (iii) Explain the rationale for the sum of random effects of all eight cross-section units being zero.
- (iv) Suppose the Hausman test statistic reported in the study is χ^2 (2 d.f) = 4.868021. Clearly mention the null and alternative hypotheses. Choose 5% level of significance. Which model would you choose—Fixed effect or Random Effects ?

8

- (a) (i) मान लीजिए कि सही मॉडल है :

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

तथा फिट किया गया मॉडल है :

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + v_i$$

$E(\hat{\alpha}_1)$, $E(\hat{\alpha}_2)$ और $E(\hat{\alpha}_3)$ क्या हैं ? क्या फिट मॉडल के ओएलएस अनुमानक निष्पक्ष हैं ? $\hat{\alpha}_2$ व $\hat{\beta}_2$ के प्रसरणों (variances) की भी तुलना कीजिए। क्या $\hat{\alpha}_2$ कुशल (efficient) है ? समझाइए।

- (ii) यह कब कहा जाता है कि एक आँकड़ा-बिन्दु (data point) ऊँचा उत्तोलन (leverage) डालता है ? यदि कोई डेटा बिंदु उच्च उत्तोलन करता है, तो क्या यह 'प्रभावशाली बिंदु' भी है ? अपने उत्तर के पक्ष में तर्क दीजिए।

- (b) एक शोधकर्ता ने 8 मोबाइल-फोन विनिर्मात्री कम्पनियों हेतु 10 वर्षों हेतु त्रैमासिक (quarterly) कुल लागत (TC) फलन का अध्ययन किया। परिणाम निम्न प्रकार है :

Dependent variable : TC

Method: EGLS

Used sample 2004-2013

Included observations : 40

Number of cross-sections : 8

Total panel observations : 320

Variable	Coefficient	Std. Error	t-Statistic
C	47.30772	1.340279	35.29692
RM	0.523554	0.012030	43.52132
ADV	2.220745	0.149031	14.90118

Random Effects

Company 1	0.258081
Company 2	-2.415602
Company 3	0.848119
Company 4	-1.775884
Company 5	1.190163
Company 6	-1.573142
Company 7	0.472518
Company 8	2.995747

जहाँ RM = कच्चे माल पर व्यय तथा ADV = विज्ञापन पर व्यय

- (i) सभी अनुप्रस्थीय अन्तःखण्डों (cross-section intercepts) के माध्य (mean) का मान क्या है ?
- (ii) कम्पनी 2 व कम्पनी 3 के यादृच्छिक प्रभावों (random effects) के मानों को समझाइए।
- (iii) सभी 8 अनुप्रस्थ इकाइयों के यादृच्छिक प्रभावों का योग शून्य होने के औचित्य (rationale) को समझाइए।
- (iv) मान लीजिए कि अध्ययन परिणामों में दिया गया हॉसमैन परीक्षण प्रतिदर्शज का मान χ^2 (2 d.f) = 4.868021 है। शून्य (zero) व वैकल्पिक (alternative) परिकल्पनाओं को स्पष्टतः लिखिए। 5% सार्थकता स्तर का चयन कीजिए। आप कौनसा मॉडल चुनेंगे—स्थिर प्रभाव या यादृच्छिक प्रभाव ?

4. (a) Suppose the desired level of inventories for a firm (Y_t^*) depends on the level of sales of the firm, X_t . The model is as follows :

$$Y_t^* = \beta_0 + \beta_1 X_t + u_t$$

- (i) Show how the Partial Adjustment hypothesis regarding Y_t^* can be used to transform the model into an autoregressive model. Write down the transformed model.
- (ii) Show that the stochastic explanatory variable in the model, Y_{t-1} is distributed independently of the disturbance term.

(iii) If u_i are serially independent, can the same be said of the disturbance terms in the transformed model ?

(iv) How can we get the long-run estimated level of inventories from the short-run estimated function ?

7

(b) In the two-variable regression model :

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

suppose Y and X take values given by the vectors :

$$Y = (-4, -2, 1, 5) \text{ and } X = (-2, -1, 1, 2).$$

(i) Calculate $\hat{\beta}$ vector using matrix algebra.

(ii) Write down the Explained Sum of Squares (ESS), Total Sum of Squares (TSS) and Residual Sum of Squares (RSS) using matrix notation and find their values.

(iii) What is the estimated value of R^2 ?

(iv) Use matrix notation to write down the standard error of the regression and obtain the estimated value.

8

(a) मान लीजिए कि एक फर्म हेतु स्टॉक (inventories) का वांछित (desired) स्तर (Y_i^*) फर्म की बिक्री के स्तर X_i पर निर्भर करता है। मॉडल निम्न प्रकार है :

$$Y_i^* = \beta_0 + \beta_1 X_i + u_i$$

- (i) दर्शाइए कि किस प्रकार Y_t^* हेतु आंशिक समायोजन परिकल्पना (Partial Adjustment hypothesis) की सहायता से इस मॉडल को एक स्व-समाश्रयी मॉडल (autoregressive model) में रूपान्तरित किया जा सकता है। रूपान्तरित मॉडल को लिखिए।
- (ii) दर्शाइए कि मॉडल में यादृच्छिक (stochastic) व्याख्याकारी चर, Y_{t-1} त्रुटि पद (disturbance term) से स्वतन्त्रतः वितरित है।
- (iii) यदि u_t स्व-सहसम्बन्धित (serially independent) नहीं है, तो क्या रूपान्तरित मॉडल के बारे में भी यही बात कही जा सकती है ?
- (iv) हम आकलित लघुकालीन (short-run) फलन से स्टॉक का आकलित दीर्घकालीन (long-run) स्तर किस प्रकार प्राप्त कर सकते हैं ?
- (b) द्विचर समाश्रयण मॉडल :

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

में मान लीजिए कि Y व X क्रमशः सदिशों :

$$Y = (-4, -2, 1, 5) \text{ व } X = (-2, -1, 1, 2).$$

में दिए गए मान लेते हैं।

- (i) मैट्रिक्स बीजगणित की सहायता से $\hat{\beta}$ सदिश की गणना कीजिए।
- (ii) मैट्रिक्स संकेतन में व्याख्याकृत वर्गयोग (Explained Sum of Squares, ESS), कुल वर्गयोग (Total Sum of Squares, TSS) व अवशिष्ट वर्गयोग (Residual Sum of Squares, RSS) लिखिए तथा इनके मान ज्ञात कीजिए।

(iii) R^2 का आकलित मान क्या है ?

(iv) मैट्रिक्स संकेतन में समाश्रयण की मानक त्रुटि (standard error) लिखिए तथा इसका आकलित मान ज्ञात कीजिए।

5. (a) Consider the Keynesian model of income determination :

$$\text{Consumption function : } C_t = \beta_0 + \beta_1 Y_t + u_t \quad 0 < \beta_1 < 1$$

$$\text{Income identity : } Y_t = C_t + I_t (= S_t)$$

where C = consumption expenditure; Y = Income; I = investment (assumed exogenous);

S = savings; t = time and u = stochastic disturbance term.

(i) Identify the endogenous variables in the model.

(ii) If we estimate the consumption function in isolation, that is, as a single equation model, show that $\hat{\beta}_1$ is a biased estimator.

(iii) Given $\sum y_t u_t = 7.78$ and $\sum y_t^2 = 110$, what is the estimated bias in $\hat{\beta}_1$? 7

(b) (i) A researcher runs the following regression using least squares method :

$$\text{Model A : } Y_t = \beta_1 + \beta_2 X_{2t} + u_t$$

In order to test whether another variable, X_3 , should be added to the model, the residuals obtained from Model A, \hat{u}_t , are regressed on all explanatory variables in Model A and X_3 as well.

$$\hat{u}_t = \alpha_1 + \alpha_2 X_{2t} + \alpha_3 X_{3t} + v_t$$

If the auxiliary regression given above has $R^2 = 0.4832$ and $n = 60$, apply the Lagrange Multiplier (LM) test for adding variables. Use 5% level of significance. State null and alternate hypotheses clearly. Should X_3 be added to the model ?

(ii) Let the true population regression be :

$$Y_i = \alpha + \beta X_i^* + u_i$$

Suppose all assumptions of CLRM are satisfied and that actual values of X^* are not observable. The observed values of X^* differ from the actual values as follows :

$$X_i = X_i^* + v_i$$

where v_i denotes the measurement error in X_i .

Write the estimable model. Show that the composite error term is no longer independent of the explanatory variable.

8

(a) आय निर्धारण के कीनेसियन मॉडल पर विचार कीजिए :

$$\text{उपभोग फलन : } C_t = \beta_0 + \beta_1 Y_t + u_t \quad 0 < \beta_1 < 1$$

$$\text{आय समिका : } Y_t = C_t + I_t (= S_t)$$

जहाँ C = उपभोग व्यय; Y = आय; I = निवेश (बाह्यगत (exogenous) माना गया है);

S = बचत; t = समय, तथा u = यादृच्छिक त्रुटि पद (stochastic error term)

- (i) इस मॉडल के अन्तःजात (endogenous) चरों की पहचान कीजिए।
- (ii) यदि हम केवल उपभोग फलन को आकलित करें, अर्थात् एक समीकरण वाले मॉडल के तौर पर, तो दर्शाइए कि $\hat{\beta}_1$ एक अभिनत (biased) आकलक है।
- (iii) प्रतिदर्श परिणामों : $\sum y_i u_i = 7.78$ व $\sum y_i^2 = 110$, से $\hat{\beta}_1$ में अनुमानित पूर्वाग्रह दर्शाइए।
- (b) (i) एक शोधकर्ता न्यूनतम वर्ग विधि (method of least squares) की सहायता से निम्नलिखित समाश्रयण करता है :

$$\text{मॉडल A : } Y_i = \beta_1 + \beta_2 X_{2i} + u_i$$

एक अन्य चर, X_3 , को इस मॉडल में जोड़ा जाना चाहिए या नहीं, इस बात का परीक्षण करने हेतु मॉडल A से प्राप्त अवशिष्टों, (residuals) \hat{u}_i , को मॉडल A के सभी व्याख्याकारी चरों व X_3 पर समाश्रयित किया जाता है :

$$\hat{u}_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + v_i$$

यदि उपर्युक्त सहायक समाश्रयण में प्राप्त $R^2 = 0.4832$ व $n = 60$ है, तो चर जोड़ने हेतु लैग्रान्ज गुणक (Lagrange Multiplier, LM) परीक्षण कीजिए। 5% सार्थकता स्तर का उपयोग कीजिए। शून्य (null) व वैकल्पिक (alternate) परिकल्पनाएँ स्पष्टतः लिखिए। क्या X_3 को मॉडल में जोड़ा जाना चाहिए ?

- (ii) मान लीजिए कि सही समष्टि समाश्रयण (population regression) निम्न प्रकार है :

$$Y_i = \alpha + \beta X_i^* + u_i$$

मान लीजिए कि CLRM की सभी मान्यताएँ सन्तुष्ट होती हैं तथा X^* के वास्तविक मान प्रेक्षणीय नहीं हैं। X^* के प्रेक्षणीय मान वास्तविक मानों से निम्नानुसार अलग हैं :

$$X_i = X_i^* + v_i$$

जहाँ v_i , X_i में मापन त्रुटि (measurement error) है।

आकलनीय (estimable) मॉडल को लिखिए। दर्शाइए संयुक्त त्रुटि पद (composite error term) अब व्याख्याकारी चर (explanatory variable) से स्वतन्त्र नहीं है।

6. (a) A researcher studied the total cost function (TC) of seven tyre-manufacturing companies (including both private and public) over 30 years.

$$\ln TC_{it} = \beta_1 + \beta_2 \ln X_{2it} + \beta_3 \ln X_{3it} + \beta_4 \ln X_{4it} + u_{it}$$

where X_2 = price of rubber, X_3 = wages and X_4 = rate of interest.

- (i) The double-log cost function was estimated using Pooled OLS method. Suppose the cost structure of the firm depends on the nature of ownership, α_i , which is not directly observable and which is time invariant. What are the consequences of using Pooled OLS method for the estimation of parameters ?
- (ii) Rewrite the model to show how heterogeneity can be accounted for by using the Fixed Effects Least-Squares Dummy Variable (LSDV) model to obtain consistent estimates of the parameters.
- (iii) Outline the test procedure to choose between pooled and fixed effects LSDV model.

- (b) (i) A researcher conducts a study where output of rice (Q) is regressed on number of labour employed on the farm (L) and number of units of machine used (K) for 65 large farms and 61 small farms for the year 2013-14.

$$Q_i = \alpha_1 + \alpha_2 L_i + \alpha_3 K_i + u_i$$

The researcher runs a pooled regression and also runs regressions separately for the two subsamples of large farms and small farms. The residual sum of squares (RSS) so obtained are reported as follows :

Pooled data ($N = 126$)	:	9.1875×10^8
Subsample of large farms ($N = 65$)	:	3.2578×10^8
Subsample of small farms ($N = 61$)	:	2.2684×10^8

Perform the Chow test. Use 1% level of significance. Would you choose the pooled regression ?

- (ii) Suppose sales of automobiles ($Sales$) in a quarter are related to GDP : $Sales_t = \beta_0 + \beta_1 GDP_t + u_t$. Show how dummy variables can be used to check for seasonality in sales of automobiles across four quarters. 8

- (a) एक शोधकर्ता ने 7 टायर-विनिर्मात्री कम्पनियों (निजी व सार्वजनिक दोनों को शामिल करते हुए) के 30 वर्षों हेतु कुल लागत फलों (TC) का अध्ययन किया :

$$\ln TC_{it} = \beta_1 + \beta_2 \ln X_{2it} + \beta_3 \ln X_{3it} + \beta_4 \ln X_{4it} + u_{it}$$

जहाँ X_2 = रबड़ की कीमत, X_3 = मजदूरी व X_4 = ब्याज दर है।

- (i) द्विगुण लघुगणक (double-log) लागत फलन को सामूहिक (Pooled) OLS की सहायता से आकलित किया गया। मान लीजिए कि फर्म की लागत संरचना स्वामित्व की प्रकृति (α_i) पर निर्भर करती है जो कि सीधे प्रेक्षणीय नहीं है तथा समय के साथ स्थिर है। इस मॉडल में प्राचलों (parameters) के आकलन हेतु सामूहिक (pooled) OLS विधि के उपयोग के क्या परिणाम होंगे ?
- (ii) प्राचलों के सुसंगत (consistent) आकलक प्राप्त करने हेतु स्थिर प्रभाव (fixed effects) न्यूनतम वर्ग (least squares) मूक चर (dummy variable) समाश्रयण की सहायता से किस प्रकार विविधता (heterogeneity) को शामिल किया जा सकता है, यह दर्शाने हेतु मॉडल को पुनर्लिखित कीजिए।
- (iii) सामूहिक व स्थिर LSDV मॉडल में से चयन करने हेतु परीक्षण प्रक्रिया की रूपरेखा दीजिए।
- (b) (i) एक शोधकर्ता एक अध्ययन करता है जिसमें 65 बड़े व 61 छोटे खेतों हेतु वर्ष 2013-14 हेतु चावल के उत्पादन (Q) को खेत में लगाए गए श्रम (L) व प्रयुक्त मशीनों की संख्या (K) पर समाश्रयित करता है :

$$Q_i = \alpha_1 + \alpha_2 L_i + \alpha_3 K_i + u_i$$

शोधकर्ता एक सामूहिक समाश्रयण करता है तथा बड़े व छोटे खेतों हेतु अलग-अलग समाश्रयण भी करता है। इनसे प्राप्त अवशिष्ट वर्गयोग (RSS) निम्न प्रकार हैं :

पूर्ण प्रतिदर्श (N = 126)	:	9.1875×10^8
बड़े खेतों का उप-प्रतिदर्श (N = 65)	:	3.2578×10^8
छोटे खेतों का उप-प्रतिदर्श (N = 61)	:	2.2684×10^8

चाऊ का परीक्षण (Chow test) कीजिए। 1% सार्थकता स्तर का उपयोग कीजिए। क्या आप सामूहिक समाश्रयण को चुनेंगे ?

- (ii) मान लीजिए कि किसी तिमाही में वाहनों की बिक्री (*Sales*) उस तिमाही में GDP से सम्बद्ध है :

$$\text{Sales}_t = \beta_0 + \beta_1 \text{GDP}_t + u_t$$

दर्शाइए कि वाहनों की बिक्री में चारों तिमाहियों (quarters) में मौसमी व्यवहार (seasonality) की जाँच करने हेतु मूक चरों (dummy variables) का किस प्रकार उपयोग किया जा सकता है ?

7. (a) (i) Using least squares method, cost of production of a firm (COST in rupees lakhs) is regressed on total output (Q, in thousand tonnes) for 60 firms for the year 2010-11. In order to test for heteroscedasticity, the squared OLS residuals are regressed on the explanatory variable :

$$\hat{u}_i^2 = \delta_1 + \delta_2 Q_i + v_i$$

If the estimated value of R^2 is 0.2354 for this auxiliary regression, conduct Breusch-Pagan (BP) test. Use 5% level of significance. Do you find evidence of heteroscedasticity ? State null and alternate hypotheses clearly.

- (ii) Consider the regression model using standardized variables :

$$Y_i^* = \alpha^* + \beta^* X_i^* + u_i$$

where $Y_i^* = \frac{Y_i - \bar{Y}}{S_Y}$ and $X_i^* = \frac{X_i - \bar{X}}{S_X}$, S_Y and S_X are sample standard deviations of Y and X respectively and \bar{Y} and \bar{X} are sample means of Y and X respectively. Show that the intercept term is always zero. How will you interpret $\hat{\beta}^*$?

(b) Consider the regression model :

$$Y_i = \alpha + \beta X_i + u_i$$

Assume that X and u are correlated and let Z be a candidate instrumental variable (IV) for X .

(i) In order for Z to be a valid instrument, what criteria must Z satisfy ?

(ii) Derive the IV estimator for β and show that it is consistent.

(iii) From a sample of 80 observations, the following data are obtained :

$$\sum Y_i^2 = 450, \quad \sum X_i^2 = 500$$

$$\sum Z_i^2 = 400, \quad \sum Y_i = 120, \quad \sum X_i = 80, \quad \sum Z_i = 40,$$

$$\sum X_i Y_i = 300, \quad \sum Z_i Y_i = 300, \quad \sum Z_i X_i = 520$$

Obtain the OLS and IV estimates of β .

8

- (a) (i) न्यूनतम वर्ग विधि की सहायता से 60 फर्मों हेतु वर्ष 2010-11 हेतु, फर्म की उत्पादन लागत (COST, लाख रुपयों में) को कुल उत्पाद (Q , हजार टनों में) पर समाश्रयित किया जाता है। प्रसरण-विषमता (heteroscedasticity) हेतु परीक्षण करने हेतु OLS अवशिष्टों (residuals) के वर्गों को व्याख्याकारी चर (explanatory variable) पर समाश्रयित किया जाता है :

$$\hat{u}_i^2 = \delta_1 + \delta_2 Q_i + v_i$$

यदि इस सहायक समाश्रयण हेतु R^2 का आकलित मान 0.2354 है, तो ब्रूश-पैगन (Breusch-Pagan, BP) परीक्षण कीजिए। 5% सार्थकता स्तर का उपयोग कीजिए। क्या आपको प्रसरण-विषमता का प्रमाण मिलता है ? शून्य व वैकल्पिक परिकल्पनाएँ स्पष्टतः लिखिए।

- (ii) मानकीकृत (standardized) चरों की सहायता से किए जाने वाले निम्नलिखित समाश्रयण पर विचार कीजिए :

$$Y_i^* = \alpha^* + \beta^* X_i^* + u_i$$

जहाँ $Y_i^* = \frac{Y_i - \bar{Y}}{S_Y}$ व $X_i^* = \frac{X_i - \bar{X}}{S_X}$, S_Y व S_X क्रमशः Y व X के प्रतिदर्श मानक विचलन हैं तथा \bar{Y} व \bar{X} क्रमशः Y व X के प्रतिदर्श माध्य हैं। दर्शाइए कि अन्तःखण्ड पद हमेशा शून्य होगा। आप β^* की व्याख्या किस प्रकार करेंगे ?

- (b) समाश्रयण मॉडल :

$$Y_i = \alpha + \beta X_i + u_i$$

पर विचार कीजिए। मान लीजिए कि X व u सहसम्बन्धित हैं तथा मान लीजिए कि Z , X हेतु एक प्रत्याशी यांत्रिक चर (IV) है।

- (i) एक वैध (valid) यांत्रिक चर होने हेतु Z को कौनसे मापदण्ड सन्तुष्ट करने चाहिए ?
(ii) β का IV आकलक व्युत्पन्न कीजिए तथा दर्शाइए कि यह सुसंगत (consistent) है।
(iii) 80 प्रेक्षणों के एक प्रतिदर्श हेतु निम्नलिखित आँकड़े प्राप्त हुए :

$$\sum Y_i^2 = 450, \quad \sum X_i^2 = 500$$

$$\sum Z_i^2 = 400, \quad \sum Y_i = 120, \quad \sum X_i = 80, \quad \sum Z_i = 40,$$

$$\sum X_i Y_i = 300, \quad \sum Z_i Y_i = 300, \quad \sum Z_i X_i = 520$$

β के OLS व IV आकलन ज्ञात कीजिए।