

This question paper contains 22 printed pages.

Your Roll No.

Sl. No. of Ques. Paper : 5968 F
Unique Paper Code : 227304
Name of Paper : Introductory Econometrics
Name of Course : B.A. (Hons.) Economics
Semester : III
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 दिये गये निर्धारित स्थान पर अपना अनुक्रमांक लिखिये।)

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

टिप्पणी:— इस प्रश्नपत्र का उत्तर अंग्रेज़ी या हिन्दी किसी एक भाषा में दीजिए; लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए।

The question paper consists of 7 questions.

Attempt any five questions.

All questions carry equal marks.

Use of simple non-programmable calculator is allowed.

Statistical tables are attached for your references.

इस प्रश्न-पत्र में 7 प्रश्न हैं।

किन्हीं पाँच प्रश्नों के उत्तर दीजिए।

सभी प्रश्नों के अंक समान हैं।

साधारण अप्रोग्रामनीय कैल्कुलेटर का उपयोग किया जा सकता है।

आपके संदर्भ हेतु सांख्यिकीय सारिणियाँ संलग्न हैं।

Q1. State whether each of the following statement is true or false. Give reasons or proofs.

- a) Consider the regression involving standardized variables, $Y_i^* = B_1^* + B_2^* X_i^* + u_i^*$, where $Y_i^* = \frac{Y_i - \bar{Y}}{S_Y}$, $X_i^* = \frac{X_i - \bar{X}}{S_X}$, \bar{X} and \bar{Y} are sample means, S_x and S_y are the sample standard deviations of X and Y respectively. The intercept term, in this regression, is always zero.
- b) The assumptions of homoscedasticity and absence of autocorrelation are necessary to use the Ordinary Least Squares (OLS) technique to estimate the coefficients of a linear regression.
- c) Keeping as many dummy variables as the categories of the qualitative variable always results in a dummy variable trap.
- d) The mean of the sample values of Y (\bar{Y}) is same as the mean of the estimated Y values ($\bar{\hat{Y}}$) in the following two models:
 Model A: $Y_i = B_1 + B_2 X_i + u_i$
 Model B: $Y_i = A_2 X_i + v_i$
- e) If the disturbance term, u_i , in the regression model, $Y_i = B_1 + B_2 X_i + u_i$, is heteroscedastic, then $\text{var}(\hat{B}_2) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2}$ (where $x_i = X_i - \bar{X}$).

(5×3=15)

बताइये कि निम्नलिखित वक्तव्य सत्य हैं अथवा असत्य। आवश्यकतानुसार कारण या प्रमाण दीजिए।

- a) मानकीकृत (standardized) चरों वाले एक समाश्रयण (regression) पर विचार कीजिए,
 $Y_i^* = B_1^* + B_2^* X_i^* + u_i^*$,
 जहाँ $Y_i^* = \frac{Y_i - \bar{Y}}{S_Y}$, $X_i^* = \frac{X_i - \bar{X}}{S_X}$, \bar{X} व \bar{Y} क्रमशः X व Y के प्रतिदर्श माध्य (sample mean) हैं तथा S_x व S_y क्रमशः इसके प्रतिदर्श मानक विचलन (sample standard deviations) हैं। इस समाश्रयण में अन्तःखण्ड (intercept) हमेशा शून्य होता है।
- b) एक रेखीय (linear) समाश्रयण के गुणांकों के साधारण न्यूनतम वर्ग (ordinary least squares) की सहायता से आकलन हेतु प्रसरण-समता (homoscedasticity) व स्व-सहसम्बन्ध (autocorrelation) की अनुपस्थिति की मान्यताएँ आवश्यक हैं।
- c) मात्रात्मक (quantitative) चर की श्रेणियों की संख्या जितने मूक (dummy) चर रखने से हमेशा मूक चर जाल (dummy variable trap) उत्पन्न होता है।

- d) निम्नलिखित मॉडलों में Y के प्रतिदर्श मानों (sample values) का माध्य (\bar{Y}) व Y के आकलित (estimated) मानों का माध्य ($\bar{\hat{Y}}$) समान होते हैं।

$$\text{Model A: } Y_i = B_1 + B_2 X_i + u_i$$

$$\text{Model B: } Y_i = A_2 X_i + v_i$$

- e) यदि समाश्रयण मॉडल $Y_i = B_1 + B_2 X_i + u_i$ में त्रुटि पद प्रसरण-विषम (heteroscedastic) है तो

$$\text{var}(\hat{B}_2) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2}, \text{ (जहाँ } x_i = X_i - \bar{X}\text{)}$$

(5×3=15)

- Q2. (a) The regression results for 32 countries for the year 2012-13 are reported as follows (standard errors are mentioned in the parentheses):

$$\hat{Y}_i = -2.03 - 0.34 \ln X_i$$

$$(\text{se}) = (0.104) \quad (0.16)$$

$$\text{TSS} = 278.00$$

where Y_i is the corruption index of a country – an indicator of the level of corruption prevailing in the country, and X_i is country's per capita GDP (in thousand dollars).

- Interpret the slope coefficient.
- Compute the t statistic under the null hypothesis that the slope coefficient is zero.
- Using the relationship between the t-statistic and F-statistic for the above simple regression model, compute R^2 for the above regression.
- Construct the ANOVA table for the above regression.

- (b) Consider the following regression results based on annual data of a certain country for the years 1965 to 2014, (the t-ratios are mentioned in the brackets):

$$\widehat{\text{IMPORTS}}_t = -108.2 - 0.045 \text{GNP}_t + 0.931 \text{CPI}_t$$

(t)

(1.032)

(1.844)

$$R^2 = 0.9894$$

where IMPORTS : Imports (in billion dollars)

GNP : Gross National Product (in billion dollars)

CPI : Consumer Price Index

- Do you find any indicators of Multicollinearity in the above results?
- Discuss any two measures to remedy the problem of Multicollinearity.

(1+1+4+4, 2+3)

- (a) 32 देशों हेतु वर्ष 2012-13 हेतु रिपोर्ट किए गए निम्नलिखित समाश्रयण परिणामों पर विचार कीजिए (मानक त्रुटियाँ कोष्ठकों में दी हुई हैं):

$$\hat{Y}_i = -2.03 + 0.34 \ln X_i$$

$$(\text{se}) = (0.104) \quad (0.16)$$

$$\text{TSS} = 278.00$$

जहाँ Y_i एक देश का भ्रष्टाचार सूचकांक है, अर्थात उस देश में व्याप्त भ्रष्टाचार के स्तर का संकेतक है तथा X_i उस देश का प्रति व्यक्ति GDP (हजारों डॉलरों में) है।

- ढाल गुणांक (slope coefficient) की व्याख्या कीजिए।
- ढाल गुणांक शून्य है, इस शून्य परिकल्पना (null hypothesis) के अधीन t प्रतिदर्शज की गणना कीजिए।
- साधारण समाश्रयण मॉडल हेतु t -प्रतिदर्शज व F -प्रतिदर्शज के मध्य सम्बन्ध की सहायता से उपरोक्त समाश्रयण हेतु R^2 की गणना कीजिए।
- उपरोक्त समाश्रयण हेतु ANOVA सारिणी की रचना कीजिए।

(b) वर्षों 1965 से 2014 की अवधि के लिए किसी देश के वार्षिक आँकड़ों पर आधारित निम्नलिखित समाश्रयण परिणामों पर विचार कीजिए (t -अनुपात कोष्ठकों में दिए हुए हैं):

$$\widehat{IMPORTS}_t = -108.2 - 0.045 GNP_t + 0.931 CPI_t \quad R^2 = 0.9894$$

(t) (1.032) (1.844)

जहाँ $IMPORTS$: आयात (अरबों डॉलरों में)

GNP : सकल राष्ट्रीय उत्पाद (अरबों डॉलरों में)

CPI : उपभोक्ता मूल्य सूचकांक

- क्या आप उपरोक्त परिणामों में बहुसंरेखता (multicollinearity) के कोई संकेतक पाते हैं?
- बहुसंरेखता की समस्या के निराकरण हेतु किन्हीं दो उपायों का विवेचन कीजिए।

(1+1+4+4, 2+3)

Q3. (a) Consider the following regression results based on a sample of 30 farms (standard errors are mentioned in the parentheses):

$$\hat{Y}_i = 384.105 + 3.67 X_i$$

(se) = (151.538) (1.00) Residual Sum of Squares (RSS) = 6776

where Y : output of rice per acre (in tons)

X : quantity of organic manure applied per acre (in kilograms)

- What is the predicted mean output of rice per acre if the amount of organic manure applied (X_0) is 8 kilograms?
- Establish a 95% confidence interval for the predicted mean output per acre when 8 kilograms of organic manure is applied, given that the sample average of organic manure applied per acre (\bar{X}) is 5 kilograms.

(b) The relationship between gross domestic product (GDP) in rupees lakhs and the expenditure on infrastructure ($INFRA$) in rupees lakhs for 44 countries for the fiscal year 2014-15 is postulated as follows:

$$\text{Model A: } GDP_i = \alpha_1 + \alpha_2 INFRA_i + u_i$$

Another study uses an alternative model,

$$\text{Model B: } GDP_i = \beta_1 + \beta_2 INFRA_i + \beta_3 INFRA_i^2 + \beta_4 INV_i + v_i$$

where INV refers to investment in manufacturing (in rupees lakhs)

The regression results are as follows (t-ratios mentioned in the brackets):

Model	A	B
Constant	526.341	543.895
$INFRA$ (t)	4.453 (3.572)	2.602 (2.891)
$INFRA^2$ (t)	—	0.258 (4.281)
INV (t)	—	5.485 (6.238)
R^2	0.2514	0.693

- What is the impact of infrastructure on GDP in Model B.
- Test for overall significance of Model B at 1% level of significance.
- Find Adjusted R^2 for Model A.
- Which model will you choose – Model A or Model B? State the null and alternative hypotheses clearly. Use 1% significance level.

(1+5, 1+3+2+3)

(a) 30 खेतों के एक प्रतिदर्श पर आधारित निम्नलिखित समाश्रयण परिणामों पर विचार कीजिए (मानक त्रुटियाँ कोष्ठको में दी हुई हैं):

$$\hat{Y}_i = 384.105 + 3.67 X_i$$

$$(se) = (151.538) \quad (1.00)$$

$$\text{अवशिष्ट वर्ग योग (RSS)} = 6776$$

जहाँ Y : चावल की प्रति एकड़ उपज (टनों में)

X : प्रति एकड़ प्रयुक्त जैविक खाद की मात्रा (किलोग्राम में)

- यदि जैविक खाद की प्रयुक्त मात्रा (X_0) 8 किलोग्राम है तो चावल की पूर्वकथित (predicted) प्रति एकड़ माध्य उपज क्या है?
- यदि जैविक खाद की प्रयुक्त मात्रा (X_0) 8 किलोग्राम है व इसका प्रतिदर्श माध्य (\bar{X}) 5 किलोग्राम है तो चावल की पूर्वकथित प्रति एकड़ माध्य उपज हेतु 95% विश्वास्यता अन्तराल (confidence interval) स्थापित कीजिए।

(b) 44 देशों हेतु वित्तीय वर्ष 2014-15 हेतु सकल घरेलू उत्पाद (GDP , लाखों रुपयों में) व आधारभूत संरचनाओं (infrastructure) पर व्यय ($INFRA$, लाखों रुपयों में) के मध्य सम्बन्ध निम्न प्रकार होने का दावा किया गया है:

$$\text{मॉडल A: } GDP_i = \alpha_1 + \alpha_2 INFRA_i + u_i$$

एक अन्य अध्ययन में निम्नलिखित मॉडल का उपयोग किया गया है

$$\text{मॉडल B: } GDP_i = \beta_1 + \beta_2 INFRA_i + \beta_3 INFRA_i^2 + \beta_4 INV_i + v_i$$

जहाँ INV विनिर्माण (manufacturing) में निवेश है (लाखों रुपयों में)।

समाश्रयण परिणाम निम्न प्रकार है: (t-अनुपात कोष्ठकों में दिए हुए हैं)

मॉडल	A	B
स्थिरांक	526.341	543.895
$INFRA$ (t)	4.453 (3.572)	2.602 (2.891)
$INFRA^2$ (t)	—	0.258 (4.281)
INV (t)	—	5.485 (6.238)
R^2	0.2514	0.693

- मॉडल B में आधारभूत संरचना का GDP पर क्या प्रभाव है
- मॉडल B की सम्पूर्ण (overall) सार्थकता हेतु 1% सार्थकता स्तर पर परीक्षण कीजिए।
- मॉडल A हेतु समायोजित R^2 ज्ञात कीजिए।
- आप कौनसा मॉडल चुनेंगे – मॉडल A या मॉडल B? शून्य तथा वैकल्पिक (alternative) परिकल्पनाएँ स्पष्टतः लिखिए। 1% सार्थकता स्तर का उपयोग कीजिए।

(1+5, 1+3+2+3)

Q4. (a) Let the true model of wage determination be :

$$\log (WAGE_i) = B_1 + B_2 IQ_i + B_3 EDU_i + u_i$$

where IQ refers to Intelligent Quotient of a person and EDU represents years of education.

- A researcher omits EDU variable and runs the regression:

$$\log (WAGE_i) = A_1 + A_2 IQ_i + v_i$$

Will a_2 (the OLS estimator of A_2) be unbiased? Explain your answer, using the information that EDU has a positive impact on $WAGES$ and the slope coefficient in the regression of EDU on IQ is positive.

- If we add an additional (irrelevant) regressor, X , to the true model and run the model

$$\log (WAGE_i) = A_1 + A_2 IQ_i + A_3 EDU_i + A_4 X_i + v_i,$$

will a_1 and a_2 (the OLS estimators of A_1 and A_2 respectively) be unbiased? What is the expected value of a_4 (the OLS estimator of A_4)?

(b) The regression model,

$$\ln(\text{Crime}_i) = B_1 + B_2 \ln(\text{PCGNP}_i) + u_i$$

is run for a group of developing and developed countries for the year 2005-06 to study the impact of per capita GNP (PCGNP) on number of crimes per lakh of population occurring in the country. The regression was also run for the two groups of countries separately. The following regression results are reported:

Model	Sample size	Residual Sum of Squares
Developed countries	14	3.701
Developing countries	18	4.803
All countries	32	9.766

Using this information, test whether the combined regression is preferred compared to the two sub-sample regressions. Use 10% level of significance.

(c) Consider the regression model, $Y_i = B_1 + B_2 X_{2i} + B_3 X_{3i} + u_i$. Write the formula for $\text{var}(b_2)$, where b_2 is the OLS estimator of B_2 and define the terms used. What is Variance Inflation Factor (VIF)? How does it affect $\text{var}(b_2)$, given that the regressors X_2 and X_3 are found to be highly but not perfectly correlated in the sample data collected for estimating the above model.

(3+3, 5, 4)

(a) मान लीजिए कि मजदूरी निर्धारण का सही मॉडल निम्न प्रकार है:

$$\log(WAGE_i) = B_1 + B_2 IQ_i + B_3 EDU_i + u_i$$

जहाँ IQ बुद्धिमत्ता (intelligence quotient) को तथा EDU शिक्षा के वर्षों को व्यक्त करता है।

- i. एक शोधकर्ता EDU को हटा देता है तथा समाश्रयण $\log(WAGE_i) = A_1 + A_2 IQ_i + v_i$ को आकलित करता है तो क्या a_2 (A_2 का OLS आकलक) अनभिनत (*unbiased*) होगा? अपने उत्तर की व्याख्या कीजिए, यह ध्यान में रखते हुए कि EDU का WAGES पर धनात्मक प्रभाव होता है तथा व EDU के IQ पर समाश्रयण में ढाल गुणांक धनात्मक होता है।
- ii. यदि हम सही मॉडल में एक अतिरिक्त (अप्रसंगिक) चर X जोड़ दें व मॉडल $\log(WAGE_i) = A_1 + A_2 IQ_i + A_3 EDU_i + A_4 X_i + v_i$, को आकलित करें, तो क्या a_1 व a_2 (क्रमशः A_1 व A_2 के OLS आकलक) अनभिनत होंगे? a_4 (A_4 के OLS आकलक) का प्रत्याशित मान क्या है?

(b) किसी देश की प्रति व्यक्ति GNP (PCGNP) के उस देश में प्रति लाख जनसंख्या होने वाले अपराधों की संख्या (CRIME) पर प्रभाव का अध्ययन करने हेतु वर्ष 2005-06 हेतु विकसित व विकासशील देशों के एक समूह हेतु एक समाश्रयण मॉडल,

$$\ln(\text{Arrests}_i) = B_1 + B_2 \ln(\text{PCGNP}_i) + u_i$$

का आकलन किया गया। इसका आकलन देशों के दोनों समूहों हेतु अलग-अलग भी किया गया तथा निम्नलिखित परिणाम प्राप्त हुए:

मॉडल	प्रतिदर्श का आकार	अवशिष्ट वर्ग योग (RSS)
विकसित देश	14	3.701
विकासशील देश	18	4.803
सभी देश	32	9.766

उपरोक्त सूचना की सहायता से परीक्षण कीजिए कि क्या उप-प्रतिदर्शवार समाश्रयणों की अपेक्षा संयुक्त समाश्रयण अधिक उपयुक्त है। 10% सार्थकता स्तर का उपयोग कीजिए।

(c) समाश्रयण मॉडल $Y_i = B_1 + B_2 X_{2i} + B_3 X_{3i} + u_i$ पर विचार कीजिए। $\text{var}(b_2)$ हेतु सूत्र लिखिए जहाँ b_2, B_2 का OLS आकलक है, तथा इसमें आने वाले पदों को परिभाषित कीजिए। प्रसरणवर्द्धन गुणांक (VIF) क्या है? यह $\text{var}(b_2)$ को किस प्रकार प्रभावित करता है, यदि उपरोक्त मॉडल को आकलित करने हेतु संग्रहित प्रतिदर्श के आँकड़ों में यह पाया गया है कि X_2 व X_3 के मध्य सहसम्बन्ध उच्च है परन्तु पूर्ण (perfect) नहीं।

(3+3, 5, 4)

Q5. (a) Consider the regression model, $BILL_i = B_1 + B_2 \left(\frac{1}{FRIENDS_i} \right) + B_3 D_i + u_i$,

where *BILL*: bill paid by a student for a post-paid mobile connection for the month of July, 2015 (in rupees)

FRIENDS: number of close friends of the student

$D_i = 1$ if the student's parents live in the same city

= 0 otherwise

The regression results are reported as follows:

$$\widehat{BILL}_i = 539 - 45 \left(\frac{1}{FRIENDS_i} \right) - 134 D_i$$

$$(p\text{-value}) = (0.025) \quad (0.003) \quad (0.012)$$

n=45

- Is the data time series or cross section?
- Write the regression equations for the students whose parents stay in the same city and for students whose parents do not stay in the same city
- What are the upper limits to the bill paid for the two categories of the students?

iv. Test the statistical significance of B_3 at 5% level of significance. State the null and alternative hypothesis.

(b) Consider the following formulations of the two-variable population regression functions (PRF):

$$\text{Model I: } Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$\text{Model II: } Y_i = \alpha_1 + \alpha_2 (X_i - \bar{X}) + v_i$$

- Prove that b_2 is an unbiased estimator of β_2 , where b_2 is the usual OLS estimator of β_2 in Model I.
- What will be the OLS estimator of α_2 in Model II? Are the OLS estimators of β_2 and α_2 identical?
- What will be the OLS estimator of α_1 in Model II?

(1+2+2+3, 2+3+2)

(a) समाश्रयण मॉडल

$$BILL_i = B_1 + B_2 \left(\frac{1}{FRIENDS_i} \right) + B_3 D_i + u_i$$

पर विचार कीजिए जहाँ

BILL: एक छात्र द्वारा माह जुलाई 2015 हेतु पोस्ट-पेड मोबाइल कनेक्शन हेतु अदा किया गया बिल है (रुपयों में)

FRIENDS: छात्र के घनिष्ठ मित्रों की संख्या

$D_i = 1$ यदि छात्र के माता-पिता उसी शहर में रहते हैं
 $= 0$ अन्यथा

निम्नलिखित समाश्रयण परिणाम प्राप्त हुए:

$$\widehat{BILL}_i = 539 - 45 \left(\frac{1}{FRIENDS_i} \right) - 134 D_i$$

$$(p \text{ value}) = (0.025) \quad (0.003) \quad (0.012) \quad n=45$$

- आँकड़े काल-श्रेणी (time series) हैं या अनुप्रस्थ (cross-section) ?
- जिन छात्रों के माता-पिता उसी शहर में रहते हैं उनके लिए तथा जिनके माता-पिता उसी शहर में नहीं रहते हैं उनके लिए समाश्रयण समीकरण लिखिए।
- इन दो श्रेणियों के छात्रों द्वारा अदा किए गए बिलों की ऊपरी सीमाएँ क्या हैं?
- B_3 की सांख्यिकीय सार्थकता का 5% सार्थकता स्तर पर परीक्षण कीजिए। शून्य व वैकल्पिक परिकल्पनाओं का उल्लेख कीजिए।

(b) एक द्वि-चर समष्टि समाश्रयण फलन के निम्नलिखित दो सूत्रों (formulations) पर विचार कीजिए:

$$\text{मॉडल I: } Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$\text{मॉडल II: } Y_i = \alpha_1 + \alpha_2(X_i - \bar{X}) + v_i$$

- सिद्ध कीजिए कि b_2, β_2 का एक अनभिन्न आकलक है, जहाँ b_2 मॉडल I में β_2 का OLS आकलक है।
- मॉडल II में α_2 का अनभिन्न आकलक क्या होगा? क्या β_2 व α_2 के OLS आकलक समान होंगे?
- मॉडल II में α_1 का OLS आकलक क्या होगा?

(1+2+2+3, 2+3+2)

Q6. (a) A researcher estimates the regression model, $\ln Y_i = B_1 + B_2 X_i + u_i$, for a cross-section of 50 entities. She obtains OLS residuals (e_i) and suspecting heteroscedasticity, she runs the regression, $\ln(e_i^2) = A_1 + A_2 \ln X_i + v_i$, and reports the regression results as follows: (where \ln indicates natural log)

$$\begin{aligned} \ln(e_i^2) &= 264 + 3.478 \ln X_i \\ (\text{p-value}) &= (0.002) \quad (0.007) \end{aligned}$$

- Identify the above equation - which test of heteroscedasticity does it correspond to? State the null and alternative hypothesis clearly. Do you find evidence of heteroscedasticity? Use 5% level of significance.
- Will the OLS estimators from the original regression be BLUE? Why / why not?
- In order to remedy the problem of heteroscedasticity, another study ran the following transformed model: $\frac{Y_i}{X_i^3} = \frac{B_1}{X_i^3} + \frac{B_2 X_i}{X_i^3} + \frac{u_i}{X_i^3}$.
What is the underlying assumption about the unknown error variance behind the transformation?

(b) The regression model, $Y_i = B_1 + B_2 X_i + u_i$ is estimated for 50 observations. The Durbin-Watson d statistic is reported as 1.985. Give the detailed decision rule for the d-statistic. Do you find an evidence of Autocorrelation at 5% level of significance?

(c) The disturbance term in the regression model, $Y_i = B_1 + B_2 X_{2i} + u_i$, has a non-zero mean, that is, $E(u_i) = \mu_u$. Show that the above regression model can be rewritten as a new model with a different intercept, same slope and a disturbance term with zero mean.

(4+3+1, 4, 3)

(a) एक शोधकर्ता 50 इकाइयों के एक अनुप्रस्थ हेतु समाश्रयण मॉडल $\ln Y_i = B_1 + B_2 X_i + \mu_i$ का आकलन करता है। वह OLS अवशिष्ट (e_i) प्राप्त करती है तथा प्रसरण-विषमता (heteroscedasticity) के संदेह में समाश्रयण $\ln(e_i^2) = A_1 + A_2 \ln X_i + v_i$ करती जिसके परिणामस्वरूप उसे निम्नलिखित समाश्रयण परिणाम प्राप्त होते हैं: (जहाँ \ln प्राकृतिक लघुगणक को व्यक्त करता है।)

$$\begin{aligned} \ln(e_i^2) &= 264 + 3.478 \ln X_i \\ (\text{p-मान}) &= (0.002) \quad (0.007) \end{aligned}$$

- i. उपरोक्त समीकरण की पहचान कीजिए – यह प्रसरण-विषमता के किस परीक्षण के अनुरूप है? शून्य व वैकल्पिक परिकल्पनाओं का स्पष्टतः उल्लेख कीजिए। क्या आपको प्रसरण-विषमता का प्रमाण मिलता है? 5% सार्थकता स्तर का उपयोग कीजिए।
- ii. क्या मूल समाश्रयण के OLS आकलक BLUE होंगे? क्यों / क्यों नहीं?
- iii. प्रसरण-विषमता की समस्या का निदान करने हेतु एक पूर्वकालिक अध्ययन में निम्नलिखित रूपान्तरित मॉडल को आकलित किया गया था: $\frac{Y_i}{X_i^3} = \frac{B_1}{X_i^3} + \frac{B_2 X_i}{X_i^3} + \frac{u_i}{X_i^3}$. इस रूपान्तरण के पीछे त्रुटि पद के अज्ञात प्रसरण के बारे में क्या मान्यता निहित है?

(b) 50 प्रेक्षणों हेतु समाश्रयण मॉडल $Y_i = B_1 + B_2 X_i + u_i$ आकलित किया गया। इस हेतु डर्बिन-वॉटसन d प्रतिदर्शज 1.985 बताया गया है। d प्रतिदर्शज हेतु विस्तृत निर्णय-नियम दीजिए। क्या आपको 5% सार्थकता स्तर पर स्वसहसम्बन्ध का प्रमाण मिलता है?

(c) समाश्रयण मॉडल $Y_i = B_1 + B_2 X_{2i} + u_i$ में त्रुटि पद का माध्य अशून्य है, अर्थात् $E(u_i) = \mu_u$ दर्शाए कि उपरोक्त मॉडल एक ऐसे नए मॉडल के रूप में पुनर्लिखित किया जा सकता है जिसमें अन्तःखण्ड अलग होगा, ढाल वही होगा तथा त्रुटि पद का माध्य शून्य होगा।

(4+3+1,4, 3)

Q7: (a) The OLS residuals (e_i) obtained for the regression, $Y_i = B_1 + B_2 X_i + u_i$ are found to be autocorrelated.

- i. Discuss how can the problem of Autocorrelation be remedied using Generalized Least Squares (GLS) if the error term is assumed to follow the AR(1) scheme, that is, $u_t = \rho u_{t-1} + \theta_t$, $-1 \leq \rho \leq 1$.
- ii. What is Prais-Winsten Transformation?
- iii. Use the First Difference Method to arrive at the model that seeks to correct for autocorrelation? Does the resulting model have an intercept?

(b) Based on OLS residuals (\hat{u}_i) obtained from the regression, $Y_i = B_1 + B_2 X_i + u_i$, the results of auxiliary regression are reported as follows: (t-ratios are reported in the parentheses)

$$|\hat{u}_i| = -56 + 0.005 X_i$$

$$(t) = (1.210) (1.265)$$

$$R^2 = 0.0274$$

$$n=62$$

Perform Glejser test to test for Heteroscedasticity at 10% level of significance. State the null and alternate hypotheses clearly. Do you find evidence of heteroscedasticity? Can we use Glejser test if sample size is 15? Why / why not?

(c) You are given the following regression results, (t-ratios are reported in the parentheses)

$$\widehat{MONEY}_t = 523.851 + 4.065 CREDIT_t$$

$$(t) = (0.001) \quad (3.587)$$

$$R^2=0.958$$

where MONEY = money supply in the economy (in hundred crores)

CREDIT= credit given by the banks (in hundred crores)

What would be the new intercept and slope coefficients and R^2 if MONEY is expressed in crores and CREDIT in thousand crores.

(3+2+3, 4, 3)

(a) समाश्रयण

$$Y_t = B_1 + B_2 X_t + u_t$$

हनु प्राप्त OLS अवशिष्ट (e_t) स्वसहसम्बन्धित पाए गए हैं। प्रतिदर्श के आँकड़े CLRM की स्वसहसम्बन्ध (autocorrelation) की अनुपस्थिति की मान्यता को सन्तुष्ट नहीं करते हैं।

विवेचन कीजिए कि स्वसहसम्बन्ध की समस्या का सामान्यीकृत वर्ग विधि स त्कस प्रकार निदान किया जा सकता है यदि यह मान लिया जाए कि त्रुटि पद AR(1) प्रणाली का अनुसरण करता है, अर्थात्

$$u_t = \rho u_{t-1} + \theta_t \quad -1 \leq \rho \leq 1$$

- ii. प्राइस-विन्स्टन रूपान्तरण (Prais-Winsten transformation) क्या है?
- iii. प्रथम अन्तर विधि की सहायता से स्वसहसम्बन्ध से मुक्त एक मॉडल व्युत्पन्न काजिए। क्या इस विधि से प्राप्त मॉडल में अन्तःखण्ड होंगे?

(b) समाश्रयण $Y_t = B_1 + B_2 X_t + u_t$ से प्राप्त OLS अवशिष्टों \hat{u}_t के आधार पर किए गए सहायक समाश्रयण (auxiliary regression) से निम्नलिखित परिणाम प्राप्त हुए: (t-अनुपात कोष्ठकों में दिए हुए हैं)

$$|\hat{u}_t| = -56 + 0.005 X_t$$

$$(t) = (1.210) \quad (1.265)$$

$$R^2 = 0.0274$$

$$n=62$$

10% सार्थकता स्तर पर प्रसरण-विषमता हेतु ग्लेजसर परीक्षण कीजिए। शून्य व वैकल्पिक परिकल्पनाओं का स्पष्टतः उल्लेख कीजिए। क्या आपको प्रसरण-विषमता का प्रमाण मिलता है? यदि प्रतिदर्श का आकार 15 हो तो क्या हम ग्लेजसर परीक्षण का उपयोग कर सकते हैं?

(c) आपको निम्नलिखित समाश्रयण परिणाम दिए गए हैं: (t-अनुपात कोष्ठकों में दिए हुए हैं)

$$\widehat{MONEY}_t = 523.851 + 4.065 CREDIT_t$$

$$(t) = (0.001) \quad (3.587)$$

$$R^2=0.958$$

जहाँ MONEY = अर्थव्यवस्था में मुद्रा की आपूर्ति (सौ करोड़-रुपयों में)

CREDIT= बैंकों द्वारा प्रदत्त साख (credit) (सौ करोड़ रुपयों में)

यदि MONEY को करोड़ रुपयों में व्यक्त किया जाए व CREDIT हजार करोड़ रुपया में व्यक्त किया जाए तो नए अन्तःखण्ड व ढाल गुणांक व R^2 क्या होंगे?

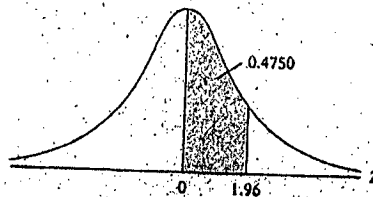
(3+2+3, 4, 3)

TABLE AREAS UNDER THE STANDARDIZED NORMAL DISTRIBUTION

Example

$Pr(0 \leq Z \leq 1.96) = 0.4750$

$Pr(Z \geq 1.96) = 0.5 - 0.4750 = 0.025$



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3685	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4238	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4454	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Note: This table gives the area in the right-hand tail of the distribution (i.e., $Z \geq 0$). But since the normal distribution is symmetrical about $Z = 0$, the area in the left-hand tail is the same as the area in the corresponding right-hand tail. For example, $Pr(-1.96 \leq Z \leq 0) = 0.4750$. Therefore, $Pr(-1.96 \leq Z \leq 1.96) = 2(0.4750) = 0.95$.

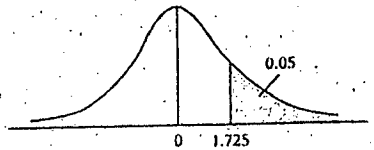
TABLE PERCENTAGE POINTS OF THE t DISTRIBUTION

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.

TABLE 1 UPPER-TAIL AREA VALUES OF THE K^2 DISTRIBUTION

Example
Pr(K^2 > 10.85) = 0.95
Pr(K^2 > 23.85) = 0.25
Pr(K^2 > 31.41) = 0.25

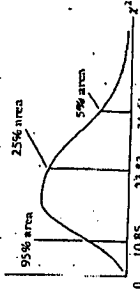


Table with 4 columns for degrees of freedom (1-100) and 10 columns for upper-tail areas (0.95 to 0.05). Values are arranged in a grid format, with a dashed line separating the example values from the rest of the table.

*For df greater than 100 the convention sqrt(2k) - sqrt(2k-2) = z allows the standardized normal distribution, where k represents the degrees of freedom.

2396

TABLE 1. UPPER PERCENTAGE POINTS OF THE F DISTRIBUTION (Continued)

df for numerator N ₁	df for denominator N ₂										df for numerator N ₁										P ₁																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	15	20	24	30	40	50	60		100	150	200	500	∞																		
22	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
24	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
25	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
26	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
28	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
30	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
40	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
50	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
60	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
100	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
200	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01
∞	1.40	1.46	1.47	1.45	1.44	1.42	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28	1.27	1.26	1.24	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13	1.12	1.11	1.10	1.09	1.08	1.07	1.06	1.05	1.04	1.03	1.02	1.01

