(i) इंस्ट्रूमेंटल वैरिएबल (IV) रिग्रेशन में 'कमजोर इंस्ट्रूमेंट' ('weak instrument') शब्द से आप क्या समझते हैं। कमजोर इंस्ट्रूमेंट एक समस्या क्यों हैं? एकल एंडोजेनस रिग्रेसर वाले मॉडल में कमजोर इंस्ट्रूमेंट की जाँच के लिए रुल ऑफ थम्ब (rule of thumb) क्या है?

(ii) लॉगिट, प्रोबिट और लीनियर प्रोबेबिलिटी मॉडल (Logit, Probit and Linear Probability Model) के उपयोग को कौन से कारक नियंत्रित करते हैं? चर्चा कीजिए।

[This question paper contains 16 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 4581 **I**

Unique Paper Code : 2273100008

Name of the Paper : Advanced Econometrics (NEP)

Name of the Course : **Pool of Discipline Specific Elective (DSE)**

Semester : V

Duration : 3 Hours Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. The question paper consists of **six** questions.
3. **Q1** is compulsory. It carries **10** marks.
4. Do any **four** out of **five** questions from **Q2** to **Q6**. Each question carries **20** marks.
5. Use of simple non programmable calculator is allowed.
6. Statistical tables are attached for your reference.

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

छात्रों के लिए निर्देश

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

2. प्रश्न पत्र में छह प्रश्न हैं ।

3. प्रश्न 1 अनिवार्य है । इसके 10 अंक हैं ।

4. प्रश्न 2 से प्रश्न 6 तक **पाँच** में से कोई **चार** प्रश्न हल करें । प्रत्येक प्रश्न के 20 अंक हैं ।

5. सामान्य गैर-प्रोग्रामीय कैलकुलेटर का उपयोग करने की अनुमति है ।

6. आपके संदर्भ के लिए सांख्यिकीय तालिकाएँ संलग्न हैं ।

7. इस प्रश्न-पत्र का उत्तर अंग्रेजी या हिंदी किसी एक भाषा में दीजिए, लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए ।

1. (a) Consider the following multiple regression model containing three independent variables, satisfying assumptions of classical linear regression model:

(iii) इस मॉडल में, आप गुणांकों (coefficients) की व्याख्या किस प्रकार करेंगे?

(iv) निश्चित प्रभाव अनुमान के अपने ज्ञान के आधार पर, यदि विशिष्ट त्रुटियों (idiosyncratic errors) u_{it} में क्रमिक सहसंबंध हो, तो क्या संभावित समस्याएँ उत्पन्न हो सकती हैं?

(v) आप इसका परीक्षण किस प्रकार करेंगे? निश्चित प्रभाव अनुमानक (fixed effects estimator) के लिए स्वतंत्रता की डिग्रियां (degrees of freedom) क्या होगी?

6. (i) What do you understand by the term 'weak instrument' in an instrumental variables (IV) regression. Why are weak instruments a problem? What is the rule of thumb for checking for weak instruments in a model with a single endogenous regressor?

(ii) What factors govern the use of Logit, Probit and Linear Probability Model? Discuss.

(10×2=20)

एक पैनल डेटा सेट पर विचार कीजिए जहाँ आश्रित चर Y_{it} वर्ष t में फर्म i के लिए फर्म उत्पादकता के लॉग का प्रतिनिधित्व करता है। यह डेटा 54 फर्मों के संबंध में, जिन्होंने प्रत्येक वर्ष स्क्रैप दरों की सूचना दी, तीन वर्षों, 1987, 1988 और 1989 के लिए हैं। आपकी रुचि फर्म उत्पादकता के संबंध में नौकरी प्रशिक्षण कार्यक्रम (बाइनरी चर $grant_{it}$ द्वारा इंगित है) के प्रभाव का आकलन करने में है। मॉडल में समय-नियत प्रभावों को नियंत्रित करने के लिए डमी वर्ष (d88, d89) शामिल हैं। अप्रत्याशित फर्म-विशिष्ट प्रभाव (unobserved firm-specific effects) (a_i) व्याख्यात्मक चर (explanatory variables) के साथ सहसंबंधित हैं।

मॉडल निम्नानुसार विनिर्दिष्ट है :

$$\log(scrap_{it}) = \beta_0 + \beta_1\, grant_{it} + \beta_2\, grant_{it-1} + \beta_3\, \text{d88} + \beta_4\, \text{d89} + a_i + u_{it}$$

(i) इस संदर्भ में निश्चित प्रभाव अनुमानक (fixed effects estimator) क्यों उपयुक्त है, स्पष्ट कीजिए।

(ii) निश्चित प्रभाव परिवर्तन (fixed effects transformation) किस प्रकार फर्म-विशिष्ट अप्रत्याशित विषमता (unobserved heterogeneity) को समाप्त करता है?

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$$

You are interested in estimating the sum of the parameters on x_1 and x_2; call this $\theta_1 = \beta_1 + \beta_2$.

(i) Show that $\hat{\theta}_1 = \hat{\beta}_1 + \hat{\beta}_2$ is an unbiased estimator of θ_1.

(ii) Find $\text{Var}(\hat{\theta}_1)$ in terms of $\text{Var}(\hat{\beta}_1)$, $\text{Var}(\hat{\beta}_2)$ and $\text{Corr}(\hat{\beta}_1, \hat{\beta}_2)$.

(b) Which of the following can cause OLS estimators to be biased? Explain.

(i) Heteroskedasticity

(ii) Omitting an important variable

(iii) A sample correlation coefficient of .95 between two independent variables both included in the model. (5×2=10)

(क) तीन स्वतंत्र चर वाले निम्नलिखित बहु प्रतिगमन मॉडल (multiple regression model) पर विचार कीजिए, जो शास्त्रीय रैखिक प्रतिगमन मॉडल (classical linear regression model) की मान्यताओं को पूरा करता है :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$$

आप x_1 और x_2 पर मापदंड $\theta_1 = \beta_1 + \beta_2$ के योग का आकलन करना चाहते हैं;

(i) दर्शाइए कि $\hat{\theta}_1 = \hat{\beta}_1 + \hat{\beta}_2$, θ_1 का एक निष्पक्ष अनुमानक है।

(ii) प्रसरण ($\hat{\beta}_1$), प्रसरण ($\hat{\beta}_2$) और सह-सबंध ($\hat{\beta}_1, \hat{\beta}_2$) के संदर्भ में प्रसरण ($\hat{\theta}_1$) ज्ञात कीजिए।

(ख) निम्नलिखित में से कौन सा OLS अनुमानकों को पक्षपातपूर्ण बना सकता है? समझाइए।

(i) हेटेरोस्केडैस्टिसिटी।

(ii) किसी महत्वपूर्ण चर को छोड़ देना।

(iii) मॉडल में शामिल दो स्वतंत्र चरों के बीच .95 का प्रतिरूप सहसंबंध गुणांक।

2. Four hundred driver's license applicants were randomly selected and asked whether they passed their driving test ($Pass_i = 1$) or failed their test ($Pass_i = 0$);

The model is specified as :

$$\log(scrap_{it}) = \beta_0 + \beta_1\, grant_{it} + \beta_2\, grant_{it-1} + \beta_3\, d88 + \beta_4\, d89 + a_i + u_{it}$$

(i) Explain why the fixed effects estimator is appropriate in this context.

(ii) How does the fixed effects transformation eliminate the firm-specific unobserved heterogeneity?

(iii) In this model, how would you interpret the coefficients?

(iv) Based on your knowledge of fixed effects estimation, what potential problems could u_{it} arise if there is serial correlation in the idiosyncratic errors?

(v) How would you test for this? What will be the degrees of freedom for the fixed effects estimator? (4×5=20)

(ii) यदि $\alpha_1 \neq 0$ और $\alpha_2 = 0$ है, तो y_2 के लिए कम किया गया रूप ज्ञात कीजिए।

(iii) यदि $\alpha_1 \neq 0$, $\alpha_2 \neq 0$ और $\alpha_1 \neq \alpha_2$ है, तो y_1 के लिए कम किया गया रूप ज्ञात कीजिए।

(iv) यदि $\alpha_1 \neq 0$, $\alpha_2 \neq 0$ और $\alpha_1 \neq \alpha_2$ हो तो क्या y_2 का कम किया गया रूप है? यदि हाँ, तो बताइए।

(v) क्या आपूर्ति और माँग के उदाहरणों में शर्त $\alpha_1 \neq \alpha_2$ पूरी होने की संभावना है? समझाइए।

5. Consider a panel data set where the dependent variable Y_{it} represents the log of firm productivity for firm in year t. The data is for three years; 1987, 1988 and 1989, on the 54 firms that reported scrap rates in each year. You are interested in estimating the impact of a job training program (indicated by a binary variable $grant_{it}$) on firm productivity. The model includes year dummies (d88, d89) to control for time-fixed effects. The unobserved firm-specific effects (a_i) are correlated with the explanatory variables.

data were also collected on their gender ($Male_i$ = 1 if male and = 0 if female) and their years of driving experience (Experience, in year). The following table summarizes the results from the estimated Probit model :

Dependent Variable: Pass	
Explanatory variable	(1)
Experience	0.041 (0.156)
Male	-0.174 (0.259)
Male x Experience	0.015 (0.019)
Constant	0.806 (0.200)

Note: The figures in column (1) indicate estimated coefficients and the figures below in parenthesis indicate standard errors.

Using results in Column (1),

(i) Akhil is a man with 10 years of driving experience. What is the probability that he will pass the test?

(ii) Julie is a woman with 2 years of driving experience. What is the probability that she will pass the test?

(iii) Does the effect of experience on test performance depend on gender? Explain.

(iv) Can we use a logit regression for the above model? Explain your answer. (5×4=20)

ड्राइविंग लाइसेंस के चार सौ आवेदकों को यादृच्छिक रूप से चुना गया और उनसे पूछा गया कि क्या उन्होंने अपना ड्राइविंग टेस्ट उत्तीर्ण कर लिया है ($Pass_i = 1$) या ड्राइविंग टेस्ट में अनुत्तीर्ण रहे हैं ($Pass_i = 0$); उनके लिंग से संबंधी, यदि पुरुष हों तो ($Male_i = 1$) और यदि महिला हों तो ($Male_i = 0$) और उनके ड्राइविंग का अनुभव वर्षों में (Experience, in years) संबंधी डेटा एकत्र किया गया है। अनुमानित प्रोबिट मॉडल (estimated Probit model) के परिणामों का सार निम्न सारणी में दिया गया है :

(ii) If $\alpha_1 \neq 0$ and $\alpha_2 = 0$, find the reduced form for y_2.

(iii) If $\alpha_1 \neq 0$, $\alpha_2 \neq 0$ and $\alpha_1 \neq \alpha_2$, find the reduced form for y_1.

(iv) Does y_2 have a reduced form in case $\alpha_1 \neq 0$, $\alpha_2 \neq 0$ and $\alpha_1 \neq \alpha_2$? If yes, specify.

(v) Is the condition $\alpha_1 \neq \alpha_2$ likely to be met in supply and demand examples? Explain.

(4×5=20)

बाईं ओर दिखाई दे रहे समान चर y_1 (typically "quantity") के साथ "आपूर्ति और मांग के रूप" में द्वि-समीकरण प्रणाली पर विचार कीजिए :

$$y_1 = \alpha_1 y_2 + \beta_1 z_1 + u_1$$

$$y_1 = \alpha_2 y_2 + \beta_2 z_2 + u_2$$

(i) यदि $\alpha_1 = 0$ या $\alpha_2 = 0$ है, तो समझाइए कि y_1 के लिए कम किया गया रूप (reduced form) क्यों विद्यमान है।

(ii) मान लीजिए कि आपके पास दो अतिरिक्त चरों, 16 वर्ष की आयु में निकटतम स्कूल की दूरी (dschool) और निकटतम कॉलेज की दूरी (dcollege) से संबंधित डेटा है। चर्चा कीजिए कि क्या ये अच्छे साधन (instruments) हो सकते हैं।

(iii) मान लीजिए कि आप उपर्युक्त मॉडल के आकलन (instruments) में dschool और dcollege का साधन के रूप में उपयोग करते हैं। शिक्षा के लिए छोटा किया गया रूप (reduced form) लिखिए।

(iv) वे स्थितियां बताइए जिनके तहत उपर्युक्त भाग (iii) में मॉडल के मापदंडों की पहचान की जाती है।

4. Consider a two-equation system in "supply and demand form" with the same variable y_1 (typically "quantity") appearing on the left-hand side:

$$y_1 = \alpha_1 y_2 + \beta_1 z_1 + u_1$$

$$y_1 = \alpha_2 y_2 + \beta_2 z_2 + u_2$$

(i) If $\alpha_1 = 0$ or $\alpha_2 = 0$, explain why a reduced form exists for y_1.

आश्रित चर : उत्तीर्ण	
व्याख्यात्मक चर	(1)
अनुभव	0.041 (0.156)
पुरुष	-0.174 (0.259)
पुरुष x अनुभव	0.015 (0.019)
स्थिरांक	0.806 (0.200)

नोट: कॉलम (1) में दिए गए आंकड़े अनुमानित गुणांकों (estimated coefficients) दर्शाते हैं और कोष्ठक (parenthesis) में दिए गए आंकड़े मानक त्रुटियों (standard errors) को दर्शाते हैं।

कॉलम (1) में दिए गए परिणामों का उपयोग करते हुए निम्नलिखित ज्ञात कीजिए :-

(i) अखिल पुरुष है, जिसे 10 वर्ष का ड्राइविंग अनुभव है। उसके परीक्षण उत्तीर्ण करने की संभावना (probability) क्या है?

(ii) जूली महिला है, जिसे 2 वर्ष का ड्राइविंग अनुभव है। उसके परीक्षण उत्तीर्ण करने की संभावना (probability) क्या है?

(iii) क्या परीक्षण में प्रदर्शन पर अनुभव का पड़ने वाला प्रभाव लिंग पर निर्भर करता है? समझाइए।

(iv) क्या हम उपर्युक्त मॉडल के लिए लॉगिट प्रतिगमन (logit regression) का उपयोग कर सकते हैं? अपने उत्तर को स्पष्ट कीजिए।

3. Consider the following model of earnings :

$\ln(\text{earnings}_i) = \beta_0 + \beta_1 \text{ age}_i + \beta_2 \text{ education}_i + u_i$

where earnings is hourly earnings, age is in years, education is highest qualification obtained, u is the unobservable component of the model, i indexes the observation and the β's are the unknown parameters.

(i) What could be the problem in using OLS to estimate the above model?

(ii) Suppose you have data on two additional variables, distance to nearest school (dschool) and distance to nearest college (dcollege) at 16 years of age. Discuss whether these are likely to be good instruments.

(iii) Suppose you use dschool and dcollege as instruments in the estimation of the above model. Write down the reduced form for education.

(iv) State the conditions under which the parameters of the model in part (iii) above are identified.

(5×4=20)

आय के निम्नलिखित मॉडल पर विचार कीजिए :

$\ln(\text{earnings}_i) = \beta_0 + \beta_1 \text{ age}_i + \beta_2 \text{ education}_i + u_i$

जहाँ earnings प्रति घंटे आय है, हम वर्षों में है, education प्राप्त उच्चतम योग्यता है, u मॉडल का अप्राप्य घटक (unobservable component) है, i अवलोकन (observation) को अनुक्रमित (indexes) करता है और β's अज्ञात पैरामीटर हैं।

(i) उपरोक्त मॉडल का आकलन (estimate) के लिए ऐ का उपयोग करने में क्या समस्या हो सकती है?