[This question paper contains 12 printed pages.]

Your Roll No...............

Sr. No. of Question Paper : 7550 **J**

Unique Paper Code : 2273100027

Name of the Paper : Introduction to Causal Inference

Name of the Course : **Common Pool of DSE**

Semester : VI

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. There are **two** Sections, **A & B** with a total of **6** questions.
3. Question **1** from **Section A** is compulsory.
4. Answer any **4** out of the **5** questions from **Section B.**
5. Marks are given at the end of each question.
6. Use of simple calculators is allowed.
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. इसमें दो खंड अ और ब हैं, जिनमें कुल 6 प्रश्न हैं।
3. खंड अ से प्रश्न 1 अनिवार्य है।
4. खंड ब से 5 प्रश्नों में से किन्हीं 4 के उत्तर दीजिए।
5. प्रत्येक प्रश्न के अंत में अंक दिए गए हैं।
6. साधारण कैलकुलेटर का उपयोग करने की अनुमति है।
7. इस प्रश्न-पत्र का उत्तर अंग्रेजी या हिंदी किसी एक भाषा में दीजिए, लेकिन सभी उत्तरों का माध्यम एक ही होना चाहिए।

Section A

1. (a) Express the Simple Difference in Outcomes (SDO) in terms of the Average Treatment Effect (ATE), selection bias and the heterogeneous treatment effect bias.

   (b) With regard to part (b), how does unit randomisation solve the selection bias and the heterogeneous treatment effect bias?

   (c) What if selection into treatment is independent only of Y^0 and not of Y^1?

   (d) Explain the assumptions involved in SUTVA

   (e) A researcher evaluates the effect of a job training program on participants' employment status. However, some participants later mentor or refer their friends (in the control group) to informal opportunities. Does this situation violate SUTVA? If so, which part? What are the implications for causal inference? (6+3+2+4+3)

Section B

2. A government introduced a cash transfer program in 2022 for low-income households with an annual income below ₹300,000. The benefit was a monthly transfer of ₹1,500. A researcher used a sharp regression discontinuity design (RDD) to estimate the impact of the program on annual household consumption (in ₹) using data from 2023.

   The running variable is household income, and the cutoff for eligibility is ₹300,000. The outcome variable is annual household consumption. The researcher estimates the following local linear regression model on either side of the cutoff:

Variable	Coefficient	Std. Error	p-value
Intercept	240,000	3000	0.000
Treatment (D)	18,000	5,000	0.001
Income -300,000	0.15	0.02	0.000

(a) What is the estimated causal effect of the program? Interpret the coefficient on the treatment variable clearly.

(b) How does the RDD estimate differ conceptually from an OLS regression of consumption on treatment status and income?

(c) What is the key identifying assumption of the RDD used here? Why is it important?

(d) Describe one test the researcher could use to assess whether the running variable (income) was manipulated around the cutoff.

(e) Why is bandwidth choice important in RDD? What happens if it's too wide or too narrow?

(f) Suppose a placebo test was performed at ₹350,000 (where no policy exists), and the estimated effect was statistically significant. What would this imply?

(3+4+4+3+2+2)

3. A development economist is studying the effect of household income on children's educational attainment in rural farming regions. Since income may be endogenous, they use rainfall deviation as an instrumental variable. The following regression results are presented:

Dependent Variable	Independent Variables	Coefficient	Std. Error	t-stat
First Stage (Income)	Rainfall deviation (in mm)	1.5	0.4	3.75
	Constant	5.0	1.0	
	$R^2 = 0.18$			
Reduced Form (Years of Schooling)	Rainfall deviation (in mm)	0.6	0.25	2.4
	Constant	6.2	0.5	
	$R^2 = 0.12$			
IV Estimate (2SLS)	Income (in \$1000s)	0.40	0.17	2.35
	Constant	5.7	0.6	

(a) Interpret the coefficients from the first-stage and reduced-form regressions.

(b) Use the Wald estimator to manually compute the IV estimate.

(c) Discuss assumptions for rainfall as a valid instrument. Comment on the strength of the instrument in this case using the information that is provided.

(d) What additional information would you require to use a thumb rule check for weak instruments in this context? Discuss the implications of having a weak instrument.

(e) What explains a larger IV estimate compared to OLS in the case of a valid IV estimation? (4+3+5+3+3)

4. A state government implemented a new minimum wage law in 2022 that increased the hourly minimum wage from \$10 to \$15. Another similar state did not implement this policy. You are given data from both states for the years 2020, 2021 (pre-treatment) and 2023 (post-treatment). The variable of interest is Average Weekly Hours Worked (AWHW) by low-wage workers.

You run the following DiD regression:

$$AWHW_{it} = \beta_0 + \beta_1.Post_t + \beta_2.Treated_i + \beta_3.Treated_i.Post_t + \varepsilon_{it}$$

Where :

– $Treated_i$ = 1 if the individual is in the state that raised the minimum wage, 0 otherwise.

– $Post_t$ = 1 for the year 2023, 0 for 2021 and 2020.

Variable	**Coefficient**	**Std. Error**	**p-value**
Intercept	32.0	0.5	0.000
Post	-1.2	0.7	0.092
Treated	0.4	0.6	0.500
Post. Treated	-2.5	0.8	0.002

(a) What is the estimated effect of the minimum wage policy on average weekly hours worked? Interpret the DiD coefficient in plain language. Based on the regression table, is the effect statistically significant at the 5% level? Justify your answer using the appropriate values.

(b) What is the predicted average weekly hours worked in 2023 for workers in the treatment state?

(c) What is the identifying assumption of the DiD method? Why is it important?

(d) Suppose the treatment state was experiencing a downward trend in hours worked even before the policy change. How would this violate the identifying assumption? What impact could this have on the DiD estimate?

(e) Explain one robustness check or alternative approach that could help validate the DiD assumption in this context. (4+2+4+4+4)

5. A nonprofit organization offers a job-readiness training program to unemployed youth. Participation in the program is voluntary. The researcher suspects that those who choose to participate differ systematically from those who don't (e.g., in education and motivation). To estimate the causal effect of the program on monthly earnings (measured in Rs.), the researcher uses propensity score matching (PSM). The propensity score is estimated using a logistic regression model based on pre-treatment covariates: education, age, and baseline earnings.

After matching treated individuals with similar untreated individuals (One-to-One nearest neighbor matching with replacement), the researcher estimates the following regression on the matched sample:

$$Y_i = \alpha + \tau . D_i + \epsilon_i$$

where :

Y_i : Monthly earnings for individual i

D_t : Indicator for participation in the program

Variable	Coefficient	Std. Error	p-value
Intercept	8,500	400	0.000
Treated (D)	2,200	600	0.001

(a) Interpret the coefficient on the "Treated" variable. What does this estimate represent?

(b) Why is a logit model preferred over a Linear Probability Model in constructing the propensity score?

(c) What is the Conditional Independence Assumption (CIA) in the context of PSM, and why is it critical?

(d) Describe a method to assess whether the common support assumption is met.

(e) Suppose some treated units have no comparable matches in the control group. What issue arises? How might you address it?

(f) What are the potential limitations of using nearest-neighbor matching with replacement? (3+3+3+3+3+3)

6. In 2021, the government introduced a performance-based grant program for public schools that scored below 60 on a standardized quality index. The goal was to improve student learning outcomes by providing additional resources and training to these lower-performing schools. The policy was implemented nationwide. Schools with an index score of 59 or lower received the intervention, while those scoring 60 or above did not. A researcher wants to estimate the impact of the program on average student test scores using two methods:

Regression Discontinuity Design (RDD): Schools just above and below the cutoff (scores between 50-70) are used in a local linear regression.

Difference-in-Differences (DiD): The researcher compares schools below and above the cutoff before (2020) and after (2022) the policy.

RDD Results (Local Linear Regression, Bandwidth = 10)

$$Y_i = \alpha + \tau.D_i = f(Score_i - 60) + \varepsilon_i$$

Variable	Coefficient	Std. Error	p-value
Intercept	65.0	1.5	0.000
Treated (Score < 60)	3.2	1.1	0.004
Score (centered)	0.4	0.1	0.000

DiD Regression Results

$$Y_{it} = \alpha + \delta.Post_t + \gamma.Treated_i + \tau.(Post_t.Treated_i) + \varepsilon_it$$

Variable	Coefficient	Std. Error	p-value
Intercept	66.0	0.9	0.000
Post	1.1	0.5	0.030
Treated	0.6	0.7	0.400
Post.Treated	2.0	0.8	0.010

(a) Assuming that the RDD is valid, why might the RDD estimate be larger than the DiD estimate?

(b) Explain how the comparison groups differ between the two methods.

(c) Discuss a situation where the RDD estimate would be preferred over DiD.

(d) What are the limitations of using only schools near the cutoff in RDD?

(e) Under what conditions would DiD provide biased estimates, even if the RDD appears valid?

(f) Are there any advantages if we combine RDD with DiD (having a double difference taken locally around the cutoff) to strengthen causal claims?

(3+3+3+3+3+3)

खण्ड अ

1. (क) एवरेज ट्रीटमेंट इफेक्ट (एटीई), चयन पक्षपात और विषम ट्रीटमेंट इफेक्ट पक्षपात के संदर्भ में प्रतिफल में साधारण अंतर (एसडीओ) को व्यक्त करें।

(ख) भाग (ख) के संबंध में, इकाई यादृच्छिकरण चयन पक्षपात और विषम ट्रीटमेंट इफेक्ट पक्षपात को कैसे दूर कर सकता है?

(ग) क्या होगा यदि ट्रीटमेंट में चयन केवल Y^0 से स्वतंत्र है और Y^1 से नहीं?

(घ) SUTVA में शामिल धारणाओं की व्याख्या कीजिए।

(ङ) एक शोधकर्ता प्रतिभागियों की रोजगार स्थिति पर नौकरी प्रशिक्षण कार्यक्रम के प्रभाव का मूल्यांकन करता है। हालाँकि, कुछ प्रतिभागी बाद में अपने दोस्तों (नियंत्रण समूह में) को अनौपचारिक अवसरों के लिए सलाह देते हैं या संदर्भित करते हैं। क्या यह स्थिति SUTVA का उल्लंघन करती है? यदि हाँ, तो कौन सा भाग? इसका कारणात्मक निष्कर्ष पर क्या प्रभाव पड़ेगा?

(6 + 3 + 2 + 4 + 3)

खण्ड ब

2. सरकार ने 2022 में ₹300,000 से कम वार्षिक आय वाले कम आय वाले परिवारों के लिए नकद हस्तांतरण कार्यक्रम शुरू किया। इसका लाभ ₹1,500 का मासिक हस्तांतरण था। एक शोधकर्ता ने 2023 के डेटा का उपयोग करके वार्षिक घरेलू खपत (₹ में) पर कार्यक्रम के प्रभाव का अनुमान लगाने के लिए एक तीव्र प्रतीपगमन असंततता डिजाइन (आरडीडी) का उपयोग किया।

रनिंग चर घरेलू आय है, और पात्रता के लिए कटऑफ ₹300,000 है। प्रतिफल चर वार्षिक घरेलू उपभोग है। शोधकर्ता कटऑफ के दोनों ओर निम्नलिखित स्थानीय रैखिक प्रतीपगमन मॉडल का अनुमान लगाता है :

चर	गुणांक	प्रमाप त्रुटि	p-मान
अंतःखंड	240,000	3000	0.000
ट्रीटमेंट (D)	18,000	5,000	0.001
आय -300,000	0.15	0.02	0.000

(क) कार्यक्रम का अनुमानित कारणात्मक प्रभाव क्या है? ट्रीटमेंट चर पर गुणांक की स्पष्ट व्याख्या कीजिए।

(ख) ट्रीटमेंट स्टेटस और आय पर उपभोग के ओएलएस प्रतीपगमन से आरडीडी अनुमान वैचारिक रूप से कैसे भिन्न है?

(ग) यहां इस्तेमाल किए गए आरडीडी की मुख्य पहचान धारणा क्या है? यह महत्वपूर्ण क्यों है?

(घ) एक परीक्षण का वर्णन कीजिए जिसका उपयोग शोधकर्ता यह जांचने के लिए कर सकता है कि क्या कटऑफ के आसपास रनिंग वेरिएबल (आय) में हेरफेर किया गया था।

(ङ) आरडीडी में बैंडविड्थ का चयन क्यों महत्त्वपूर्ण होता है? यदि बैंडविड्थ बहुत अधिक चौड़ी या बहुत संकरी हो, तो क्या प्रभाव पड़ता है?

(च) मान लीजिए कि ₹3,50,000 पर (जहाँ कोई नीति लागू नहीं है) एक प्लेसबो परीक्षण किया गया, और अनुमानित प्रभाव सांख्यिकीय रूप से महत्त्वपूर्ण पाया गया। इसका क्या तात्पर्य होगा?

(3+4+4+3+2+2)

3. एक विकास अर्थशास्त्री ग्रामीण कृषि क्षेत्रों में पारिवारिक आय का बच्चों की शैक्षिक उपलब्धि पर प्रभाव का अध्ययन कर रहा है। चूंकि आय एंडोजेनस (आंतरिक रूप से प्रभावित) हो सकती है, इसलिए वह वर्षा में होने वाले विचलन को एक उपकरण चर के रूप में उपयोग करता है। नीचे दिए गए प्रतीपगमन परिणाम प्रस्तुत किए गए हैं:

आश्रित चर	स्वतंत्र चर	गुणांक	प्रमाप त्रुटि	t-सांख्यिकीय मान
प्रथम चरण (आय)	वर्षा विचलन (मिमी में)	1.5	0.4	3.75
	स्थिरांक	5.0	1.0	
	$R^2 = 0.18$			
लघु रूप (स्कूली शिक्षा के वर्ष)	वर्षा विचलन (मिमी में)	0.6	0.25	2.4
	स्थिरांक	6.2	0.5	
	$R^2 = 0.12$			
IV अनुमान (2एसएलएस)	आय ($1000s में)	0.40	0.17	2.35
	स्थिरांक	5.7	0.6	

(क) प्रथम चरण और लघु-रूप प्रतीपगमन से गुणांकों की व्याख्या कीजिए।

(ख) IV अनुमान की मैन्युअल गणना करने के लिए वाल्ड अनुमानक का उपयोग कीजिए।

(ग) वर्षा को एक वैध उपकरण चर के रूप में उपयोग करने के लिए किन धारणाओं की आवश्यकता होती है, चर्चा कीजिए। दिए गए जानकारी का उपयोग करते हुए इस मामले में उपकरण की शक्ति पर टिप्पणी कीजिए।

P.T.O.

(घ) इस संदर्भ में कमजोर उपकरणों के लिए थम्ब रुल की जांच का उपयोग करने के लिए आपको कौन सी अतिरिक्त जानकारी की आवश्यकता होगी? कमजोर उपकरण होने के निहितार्थों पर चर्चा कीजिए।

(ङ) एक वैध IV अनुमान के मामले में ओएलएस की तुलना में बड़े IV अनुमान की क्या व्याख्या है? (4+3+5+3+3)

4. एक राज्य सरकार ने 2022 में एक नया न्यूनतम वेतन कानून लागू किया, जिसके तहत प्रति घंटे न्यूनतम वेतन $10 से बढ़ाकर $15 कर दिया गया। इसी तरह के दूसरे राज्य ने इस नीति को लागू नहीं किया। आपको दोनों राज्यों से वर्ष 2020, 2021 (ट्रीटमेंट से पहले) और 2023 (ट्रीटमेंट के बाद) के लिए डेटा दिया गया है। इंटरेस्ट का चर कम वेतन वाले श्रमिकों द्वारा औसत साप्ताहिक कार्य घंटे (AWHW) है। आप निम्नलिखित डीआईडी प्रतीपगमन चलाते हैं :

$$AWHW_{it} = \beta_0 + \beta_1.Post_t + \beta_2.Treated_i + \beta_3.Treated_i.Post_t + \varepsilon_{it}$$

जहां :

- $Treated_i = 1$ यदि व्यक्ति उस राज्य में है जिसने न्यूनतम वेतन बढ़ाया है, अन्यथा 0
- $Post_t$ = 2023 के लिए 1, और 2021 तथा 2020 के लिए 0.

चर	गुणांक	प्रमाप त्रुटि	p-मान
अंतःखंड	32.0	0.5	0.000
पोस्ट	-1.2	0.7	0.092
ट्रीटेड	0.4	0.6	0.500
पोस्ट. ट्रीटेड	-2.5	0.8	0.002

(क) न्यूनतम वेतन नीति का औसत साप्ताहिक काम के घंटों पर अनुमानित प्रभाव क्या है? सरल भाषा में डीआईडी गुणांक की व्याख्या कीजिए। प्रतीपगमन तालिका के आधार पर, क्या प्रभाव 5% स्तर पर सांख्यिकीय रूप से महत्वपूर्ण है? उचित मानों का उपयोग करके अपने उत्तर की पुष्टि कीजिए।

(ख) ट्रीटमेंट राज्य में 2023 में काम करने वाले श्रमिकों के लिए अनुमानित औसत साप्ताहिक कार्य घंटे क्या होंगे?

(ग) डीआईडी पद्धति की पहचान करने वाली धारणा क्या है? यह महत्वपूर्ण क्यों है?

(घ) मान लीजिए कि नीति परिवर्तन से पहले ही ट्रीटमेंट राज्य में काम के घंटों में गिरावट का रुझान था। यह पहचान करने वाली धारणा का उल्लंघन कैसे करेगा? इसका डीआईडी अनुमान पर क्या प्रभाव हो सकता है?

(ङ) एक स्थिरता जाँच या वैकल्पिक दृष्टिकोण की व्याख्या कीजिए जो इस संदर्भ में डीआईडी धारणा को मान्य करने में मदद कर सकता है। (4+2+4+4+4)

5. एक गैर-लाभकारी संगठन बेरोजगार युवाओं को नौकरी-तत्परता प्रशिक्षण कार्यक्रम प्रदान करता है। कार्यक्रम में भागीदारी स्वैच्छिक है। शोधकर्ता को संदेह है कि जो लोग भाग लेते हैं वे उन लोगों से व्यवस्थित रूप से भिन्न होते हैं जो भाग नहीं लेते हैं (उदाहरण के लिए, शिक्षा और अभिप्रेरणा में)। मासिक आय (रुपयों में मापी गई) पर कार्यक्रम के कारणात्मक प्रभाव का आकलन करने के लिए शोधकर्ता प्रोपेन्सिटी स्कोर मैचिंग (पीएसएम) का उपयोग करता है। प्रोपेन्सिटी स्कोर को प्री ट्रीटमेंट कोवैरिएट्स: शिक्षा, आयु और प्रारंभिक आय के आधार पर एक लॉजिस्टिक रिग्रेशन मॉडल का उपयोग करके अनुमानित किया जाता है।

ट्रीटेड व्यक्तियों को समान ट्रीटमेंट न किए गए व्यक्तियों के साथ मिलान करने के बाद (प्रतिस्थापन के साथ एक-से-एक निकटतम पड़ोसी मिलान), शोधकर्ता मिलान किए गए नमूने पर निम्नलिखित प्रतीपगमन का अनुमान लगाता है :

$$Y_i = \alpha + \tau . D_i + \epsilon_i$$

जहां :

Y_i : व्यक्ति i के लिए मासिक आय

D_t : कार्यक्रम में भागीदारी के लिए संकेतक

चर	गुणांक	प्रमाप त्रुटि	p-मान
अंतःखंड	8,500	400	0.000
ट्रीटेड(D)	2,200	600	0.001

(क) "ट्रीटेड" चर पर गुणांक की व्याख्या कीजिए। यह अनुमान क्या दर्शाता है?

(ख) प्रोपेन्सिटी स्कोर के निर्माण में रैखिक संभाव्यता मॉडल की तुलना में लॉगिट मॉडल को क्यों प्राथमिकता दी जाती है?

(ग) पीएसएम के संदर्भ में सशर्त स्वतंत्रता धारणा (सीआईए) क्या है, और यह महत्वपूर्ण क्यों है?

(घ) यह आकलन करने के लिए एक विधि का वर्णन कीजिए कि क्या सामान्य समर्थन धारणा पूरी हुई है।

(ङ) मान लीजिए कि कुछ ट्रीटेड इकाइयों का नियंत्रण समूह में कोई तुलनीय मिलान नहीं है। क्या समस्या उत्पन्न होती है? आप इसे कैसे संबोधित कर सकते हैं?

(च) प्रतिस्थापन के साथ निकटतम-पड़ोसी मिलान का उपयोग करने की संभावित सीमाएँ क्या हैं? (3+3+3+3+3+3)

6. 2021 में, सरकार ने मानकीकृत गुणवत्ता सूचकांक पर 60 से नीचे स्कोर करने वाले पब्लिक स्कूलों के लिए एक प्रदर्शन-आधारित अनुदान कार्यक्रम शुरू किया। इसका लक्ष्य इन कम प्रदर्शन करने वाले स्कूलों को अतिरिक्त संसाधन और प्रशिक्षण प्रदान करके छात्रों के सीखने के परिणामों में सुधार करना था। नीति को पूरे देश में लागू किया गया। 59 या उससे कम के इंडेक्स स्कोर वाले स्कूलों को हस्तक्षेप प्राप्त हुआ, जबकि 60 या उससे अधिक स्कोर करने वालों को नहीं मिला। एक शोधकर्ता दो तरीकों का उपयोग करके औसत छात्र परीक्षा स्कोर पर कार्यक्रम के प्रभाव का अनुमान लगाना चाहता है :

प्रतीपगमन डिसकंटिन्यूटी डिजाइन (**RDD**) : कटऑफ से ठीक ऊपर और नीचे के स्कूल (50-70 के बीच स्कोर) स्थानीय रैखिक प्रतीपगमन में उपयोग किए जाते हैं।

डिफरेंस-इन-डिफरेंसेज (**DiD**) : शोधकर्ता नीति से पहले (2020) और बाद में (2022) कटऑफ से नीचे और ऊपर के स्कूलों की तुलना करता है।

RDD परिणाम (स्थानीय रैखिक प्रतिगमन, बैंडविड्थ = 10)

$$Y_i = \alpha + \tau.D_i = f(Score_i - 60) + \varepsilon_i$$

चर	गुणांक	प्रमाप त्रुटि	p-मान
अंतःखंड	65.0	1.5	0.000
ट्रीटेड (स्कोर < 60)	3.2	1.1	0.004
स्कोर (सेंटर्ड)	0.4	0.1	0.000

DiD प्रतीपगमन परिणाम

$$Y_{it} = \alpha + \delta.Post_t + \gamma.Treated_i + \tau.(Post_t.Treated_i) + \varepsilon_it$$

चर	गुणांक	प्रमाप त्रुटि	p-मान
अंतःखंड	66.0	0.9	0.000
पोस्ट	1.1	0.5	0.030
ट्रीटेड	0.6	0.7	0.400
पोस्ट. ट्रीटेड	2.0	0.8	0.010

(क) यह मानते हुए कि आरडीडी वैध है, आरडीडी अनुमान डीआईडी अनुमान से बड़ा क्यों हो सकता है?

(ख) बताइए कि दोनों विधियों के बीच तुलना समूह किस प्रकार भिन्न हैं।

(ग) ऐसी स्थिति पर चर्चा कीजिए जहां आरडीडी अनुमान को डीआईडी से अधिक प्राथमिकता दी जाएगी।

(घ) आरडीडी में केवल कटऑफ के निकटवर्ती विद्यालयों का उपयोग करने की क्या सीमाएं हैं?

(ङ) किन परिस्थितियों में डीआईडी पक्षपातपूर्ण अनुमान प्रदान करेगा, भले ही आरडीडी वैध प्रतीत हो?

(च) क्या कोई लाभ है यदि हम आरडीडी को डीआईडी (कटऑफ के आसपास स्थानीय रूप से लिया गया दोहरा अंतर) के साथ जोड़ते हैं ताकि कारण संबंधी दावों को मजबूत किया जा सके?

(3+3+3+3+3+3)

(100)