इकाई 7 बहु समाश्रयण मॉडल : आकलन*

इकाई की रूपरेखा

7.0 उद्देश्य

प्रस्तुत इकाई को पढ़ने के बाद आप इस योग्य होंगे कि –

- एक से अधिक व्याख्यात्मक चर वाले बहु समाश्रयण मॉडल निर्दिष्ट कर सकें;
- समाश्रयण मॉडल के प्राचलों के गुणधर्म बताते हुए साधारण न्यूनतम वर्ग (OLS) विधि से उनका आकलन कर सकें;
- किसी आकलित बहु समाश्रयण मॉडल के परिणामों की व्याख्या कर सकें;
- बहु समाश्रयण मॉडलों में आव्यूह अंकन प्रयोग किए जाने के लाभ गिना सकें;
- आकलन की उस अधिकतम संभाव्यता विधि की व्याख्या कर सकें जो यह दर्शाती है कि अधिकतम संभाव्यता आकल (MLE) और साधारण न्यूनतम वर्ग (OLS) आकल उपगामी रूप से एक समान होते हैं;
- किसी दो व्याख्यात्मक चरों वाले सरल बहु समाश्रयण मॉडल के उदाहरण में निर्धारण गुणांक (R^2) हेतु निष्पीड़न अवकलित कर सकें; तथा
- व्यवहार में समायोजित R^2 को क्यों प्राथमिकता दी जाती है, यह निर्दिष्ट करते हुए R^2 और समायोजित R^2 के बीच अंतर अंतर स्पष्ट कर सकें।

7.1 विषय–प्रवेश

अब तक हम उस सरल समाश्रयण मॉडल से भली–भाँति परिचित हो चुके हैं जिसमें एक आश्रित चर और एक स्वतंत्र चर होता है। इसमें आश्रित चर की व्याख्या स्वतंत्र चर द्वारा की जाती है। आइए, अब हम बहु समाश्रयण मॉडल के विषय में चर्चा करते हैं। किसी भी बहु समाश्रयण मॉडल में एक आश्रित चर और एक से अधिक स्वतंत्र चर होते हैं। सरलतम संभव बहु समाश्रयण मॉडल एक तीन–चर समाश्रयण मॉडल होता है, जिसमें एक आश्रित चर और दो व्याख्यात्मक चर होते हैं। ऐसा तीन–चर बहु समाश्रयण समीकरण या मॉडल निम्नानुसार व्यक्त किया जाता है –

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \quad \dots (7.1)$$

---

* डॉ. पूजा शर्मा, सहायक प्राध्यापक, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय।

इस पूरी इकाई के दौरान हम अधिकांशतः किसी बहु समाश्रयण मॉडल पर ही चर्चा करेंगे, जैसा कि ऊपर समीकरण (7.1) में निर्दिष्ट है। यहाँ Y आश्रित चर है और X_2 व X_3 स्वतंत्र चर हैं। साथ ही, u_i एक प्रसंभाव्य त्रुटि पद है। इस त्रुटि पद की व्याख्या साधारण समाश्रयण मॉडल की भाँति ही है। आपको आश्चर्य हो सकता है कि समीकरण (7.1) में कोई X_1 क्यों नहीं है। इसका उत्तर यह है कि यहाँ सभी प्रेक्षणों के लिए X_1 को ही अव्यक्त रूप से 1 के रूप में लिया जाता है।

उपर्युक्त समीकरण में प्राचल β_1 ही अवरोधन पद है। हम पद Y, X_2 और X_3 को आर्थिक सिद्धांत से लिए गए कुछ चर मानकर चल सकते हैं। हम इसे एक माँग फलन के रूप में भी ले सकते हैं, जहाँ Y किसी वस्तु की माँग की मात्रा को दर्शाएगा, जबकि X_2 और X_3 क्रमशः उस वस्तु की कीमत और उपभोक्ता की आय को दर्शाएँगे। एक अन्य उदाहरण के रूप में हम किसी दो निविष्टों वाले उत्पादन/माँग फलन पर विचार कर सकते हैं। यहाँ Y किसी वस्तु की उत्पादित अथवा माँग की मात्रा है, जबकि X_2 और X_3 पूँजी निविष्ट हैं। हम इस प्रकार के अनेक उदाहरणों के विषय में भी सोच सकते हैं।

7.2 बहु रैखिक समाश्रयण मॉडल सम्बन्धी अवधारणाएँ

स्मरण करें कि सरल समाश्रयण मॉडल कुछ अवधारणाओं पर आधारित होता है। ये अवधारणाएँ समाश्रयण मॉडल के लिए मानदण्ड का काम करती हैं। जब इन अवधारणाओं को किसी समाश्रयण मॉडल द्वारा संतुष्ट किया जाता है तो हम इसे पारंपरिक रैखिक समाश्रयण मॉडल (CLRM) कहते हैं। पारंपरिक रैखिक समाश्रयण मॉडल के लिए प्रस्तुत मुख्य अवधारणाएँ साधारण समाश्रयण मॉडल के समान ही होती हैं। हाँ, एक भिन्नता अवश्य है, यथा, यह बहुसंरेखता विषयक एक नई अवधारणा से संबंध रखती है। चूँकि हम यहाँ एक से अधिक स्वतंत्र चर X_i पर विचार कर रहे हैं, अब यह मानकर चलना आवश्यक होगा कि X_i चर पूरी तरह से सहसंबद्ध नहीं होते। आइए, अब हम इस नई अवधारणा के साथ पारंपरिक रैखिक समाश्रयण मॉडल (CLRM) की अवधारणाओं को निम्नवत् दोहराते हैं –

1. समाश्रयण मॉडल अपने प्राचलों में रैखिक होता है। इस अवधारणा का तात्पर्य यह है कि इसमें आश्रित चर इसके प्राचलों βs का एक रैखिक फलन होता है। बहरहाल, समाश्रयण मॉडल अपने व्याख्यात्मक चर में गैर–रैखिक हो सकता है।

2. चर u_i और X_i के बीच कोई सहप्रसरण नहीं होता है। इसका यह तात्पर्य है कि समीकरण (7.1) की भाँति किसी भी बहु समाश्रयण मॉडल में त्रुटि पद और व्याख्यात्मक चर के बीच कोई सहसंबंध नहीं होता, यथा –

$$Cov(u_i, X_{2i}) = Cov(u_i, X_{3i}) = 0 \quad \text{... (7.2)}$$

उक्त समस्या से बचने के लिए हम यह मानकर चलते हैं कि सभी व्याख्यात्मक चर अपनी प्रकृति में गैर–प्रसंभाव्य होते हैं। इसका तात्पर्य यह है कि व्याख्यात्मक चर X द्वारा लिए गए मानों को दोहराए गए प्रतिदर्शों में नियत माना जाता है।

3. त्रुटि पदों का माध्य शून्य होता है। दूसरे शब्दों में, व्याख्यात्मक चर X_{2i} और X_{3i} पर सशर्त त्रुटि पद का अपेक्षित मान शून्य होता है। संकेताक्षरों में इसको निम्नवत् व्यक्त किया जाएगा –

$$E(u_i) = 0 \text{ अथवा } E(u_i|X_{2i}, X_{3i}) = 0 \quad \text{... (7.3)}$$

4. कोई स्वसहसंबंध नहीं : इस अवधारणा का अर्थ है कि व्यक्तिगत प्रेक्षणों के त्रुटि पदों के बीच कोई आनुक्रमिक सहसंबंध या स्वसहसंबंध नहीं होता। इसका तात्पर्य यह है कि i^{th} प्रेक्षण से संबद्ध त्रुटि पद u_i और j^{th} प्रेक्षण से संबद्ध त्रुटि पद u_j के बीच सहप्रसरण शून्य होता है। संकेताक्षरों में इसको निम्नवत् व्यक्त किया जाएगा –

$$\mathrm{cov}(u_i, u_j) = 0 \quad \text{... (7.4)}$$

5. समविसारिता : समविसारिता की अवधारणा का अर्थ है कि त्रुटि विचरण सभी प्रेक्षणों के लिए नियत होता है। संकेताक्षरों में इसको निम्नवत् व्यक्त किया जाएगा –

$$var(u_i^2) = \sigma^2 \qquad \dots (7.5)$$

6. यहाँ X चरों के बीच कोई वास्तविक संरेखता नहीं होती। यह एकाधिक समाश्रयण मॉडल के लिए बनाई गई नई अतिरिक्त अवधारणा है। इसका अर्थ यह है कि X_2 और X_3 के बीच कोई वास्तविक रैखिक संबंध नहीं होता। इसे ही शून्य पूर्ण बहुसंरेखता की अवधारणा के रूप में संदर्भित किया जाता है।

7. आपके प्रेक्षणों की संख्या n आकलित किए जाने वाले कुल प्राचलों की संख्या से अधिक ही होनी चाहिए। दूसरे शब्दों में, प्रेक्षणों की संख्या n व्याख्यात्मक चरों की संख्या k से अधिक होनी चाहिए।

8. शून्य विनिर्देशन अभिनति : यह मानकर चला जाता है कि मॉडल सही ढंग से निर्दिष्ट है। शून्य विनिर्देशन अभिनति की अवधारणा का अर्थ है कि मॉडल को निर्दिष्ट करते समय कोई त्रुटि स्वीकार नहीं की गई। इसका तात्पर्य यह है कि समाश्रयण मॉडल को निर्दिष्ट करते समय किसी अप्रासंगिक चर को शामिल कर लेने और किसी प्रासंगिक चर को छोड़ देने संबंधी त्रुटियों का ध्यान रखा जाता है।

9. यहाँ कोई मापन त्रुटि नहीं होती अर्थात् सभी X और Y चरों को सही–सही मापा जाता है।

7.2.1 बहु रैखिक समाश्रयण मॉडल की व्याख्या

समीकरण (7.1) की भाँति बहु समाश्रयण मॉडल में अवरोधन β_1 आश्रित चर Y के अपेक्षित मान को तब मापता है जब व्याख्यात्मक चर X_2 और X_3 के मान शून्य होते हैं। अन्य दो प्राचल β_2 और β_3 आंशिक समाश्रयण गुणांक हैं। आइए, इन गुणांकों के विषय में कुछ और जानें। समाश्रयण गुणांक β_2 और β_2 को आंशिक प्रवणता गुणांक के रूप में भी जाना जाता है। गुणांक β_2 चर Y के औसत मान में परिवर्तन चर X_3 के मान को स्थिर रखते हुए X_2 के मान में प्रति इकाई परिवर्तन [यथा E(Y)] मापता है।

संकेताक्षरों में इसका अर्थ होगा – $\beta_2 = \frac{\Delta E(Y)}{\Delta X_2}$

उक्त समीकरण चर X_3 का मान स्थिर रखते हुए चर Y के औसत मान पर चर X_2 में एक इकाई परिवर्तन का 'प्रत्यक्ष' अथवा 'निवल' प्रभाव दर्शाता है। इसी प्रकार, चर β_3 चर X_2 का मान स्थिर रखते हुए चर X_3 में प्रति इकाई परिवर्तन Y का औसत मान मापता है, यथा β_2 और β_3 को समीकरण $\beta_3 = \frac{\Delta E(Y)}{\Delta X_3}$ से दर्शाया जाता है। इस प्रकार, बहु समाश्रयण के समाश्रयण गुणांक अन्य चरों के प्रभाव को स्थिर रखते हुए आश्रित चर पर एक व्याख्यात्मक चर के प्रभाव को मापते हैं।

7.3 बहु समाश्रयण मॉडल का आकलन

किसी भी बहु समाश्रयण समीकरण का आकलन समष्टि समाश्रयण फलन $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$ की व्याख्या करने के लिए किया जाता है। इस फलन में दो घटक होते हैं। प्रथम E ($Y_i | X_{2i}, X_{3i}$) से दर्शाया गया नियतात्मक घटक होता है। इसे 'समष्टि समाश्रयण रेखा' के रूप में भी जाना जाता है। दूसरा घटक u_i से दर्शाया गया यादृच्छिक घटक होता है। उक्त फलन (PRF) का आकलन किसी प्रतिदर्श का प्रयोग करके किया जाता है। फिर आकलित फलन (यथा, प्रतिदर्श समाश्रयण फलन) को समीकरण $Y_i = b_1 + b_2 X_{2i} + b_3 X_{3i} + e_i$ से दर्शाया जाता है।

स्मरण करें कि $Y_i = \hat{Y}_i + e_i$ होता है, जहाँ $\hat{Y}_i$ ही पद $E(Y_i|X_{2i}, X_{3i})$ से दर्शाया गया Y का आकलित मान होता है और चर e_i अवशिष्ट पद होता है। उक्त प्रतिदर्श समाश्रयण फलन में चर b_1 समष्टि अवरोधन β_1 का आकलक होता है जबकि b_2 और b_3 क्रमशः समष्टि आंशिक

समाश्रयण गुणांक β_2 और β_3 के आकलक कहलाते हैं। अवशिष्ट पद e_i समष्टि त्रुटि पद u_i का आकलक होता है। हमें ज्ञात है कि कि प्रतिदर्श समाश्रयण रेखा साधारण न्यूनतम वर्ग (OLS) विधि में वर्गों के अवशिष्ट योगफल को न्यूनतम करके प्राप्त की जाती है, यथा –

$$Min \sum e_i^2 = \sum\left(Y_i - \hat{Y}_i\right)^2$$

$$= \sum(Y_i - b_1 - b_2X_{2i} - b_3X_{3i})^2 \text{ [चूँकि } \hat{Y}_i = b_1 + b_2X_{2i} + b_3X_{3i}]$$

अब हम तीन प्रथम–कोटि शर्तों पर विचार करेंगे, यथा, $\frac{\partial \sum e_i^2}{\partial b_1} = 0, \frac{\partial \sum e_i^2}{\partial b_2} = 0$ और $\frac{\partial \sum e_i^2}{\partial b_3} = 0.$

उक्त तीन आंशिक अवकलजों से हम आकलक निम्नवत् ज्ञात करते हैं –

(i) $b_1 = \bar{Y} - b_2\bar{X}_2 - b_3\bar{X}_3$

(ii) $b_2 = \frac{(\sum y_i.x_{2i})(\sum x_{3i}^2)-(\sum y_i x_{3i})(\sum x_{2i}x_{3i})}{(\sum x_{2i}^2)(\sum X_{3i}^2)-(\sum X_{2i}X_{3i})^2}$

(iii) $b_3 = \frac{(\sum y_i.x_{3i})(\sum x_{2i}^2)-(\sum y_i x_{2i})(\sum x_{2i}x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2)-(\sum x_{2i}x_{3i})^2}$

यहाँ प्राचलों के संगत विचरण और मानक त्रुटियाँ निम्नानुसार दर्शायी जाती हैं –

$$V(b_1) = \left[\frac{1}{n} + \frac{\bar{X}_2^2\sum x_{3i}^2+\bar{X}_3^2\sum x_{2i}^2-2\bar{X}_2\bar{X}_3\sum x_{2i}x}{\sum x_{2i}^2\sum x_{3i}^2-(\sum x_{2i}x_{3i})^2}\right]\sigma^2$$

$$SE(b_1) = +\sqrt{V(b_1)}$$

$$V(b_2) = \frac{\sum x_{3i}^2}{(\sum x_{2i}^2)(\sum x_{3i}^2)-(\sum x_{2i}x_{3i})^2} \times \sigma^2$$

$$SE(b_2) = +\sqrt{V(b_2)}$$

$$V(b_3) = \frac{\sum x_{2i}^2}{(\sum x_{2i}^2)(\sum x_{3i}^2)-(\sum x_{2i}x_{3i})} \times \sigma^2$$

$$V(b_3) = \frac{\sigma^2}{\sum x_{3i}^2(1-r_{23}^2)}$$

$$SE(b_3) = +\sqrt{V(b_3)}$$

इसके अलावा यह भी ध्यान देने की बात है कि –

(i) $COV(b_2, b_3) = \frac{-r_{23}\sigma^2}{(1-r_{23}^2)\sqrt{x_{2i}^2}\sqrt{x_{3i}^2}}$

साथ ही, त्रुटि विचरण और आंशिक सहसंबंध गुणांक निम्नवत् दर्शाए जाते हैं –

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-k} = \frac{RSS}{n-k} \qquad \text{... (7.6)}$$

तीन व्याख्यात्मक चरों वाले किसी भी समाश्रयण मॉडल (उदाहरणार्थ, समीकरण (7.1)) के लिए हमें ज्ञात है कि –

$$\hat{\sigma}^2 = \frac{RSS}{n-3}$$

$$r_{23} = \frac{(\sum x_{2i}x_{3i})^2}{\sum x_{2i}^2\sum x_{3i}^2} \qquad \text{... (7.7)}$$

आप देखेंगे कि उपर्युक्त निषीड़नों में रोमन लिपि के छोटे अक्षर माध्य से विचलन निरूपित करते हैं। हम जानते हैं कि, चूँकि हम यहाँ 'पारंपरिक' रैखिक बहु समाश्रयण मॉडल पर विचार कर रहे हैं, अवरोधन के OLS अनुमानक और आंशिक समाश्रयण गुणांक निम्नलिखित गुणधर्मों का पालन करते हैं–

a) समाश्रयण रेखा माध्य, $\bar{Y}, \bar{X}_2$ और $\bar{X}_3$ से होकर गुजरती है। किसी भी k–चर रैखिक समाश्रयण मॉडल में एक समाश्रेय Y_i होता है और $(k-1)$ समाश्रयी होते हैं क्योंकि गुणांकों में

से एक ही अवरोधन पद β_1 कहलाता है। अतएव, इस अवरोधन पद का आकल निम्नवत् प्राप्त होता है –

$$b_1 = \bar{Y} - b_2\bar{X}_2 - b_3\bar{X}_3$$

b) आकलित $\widehat{Y\imath}$ का औसत मान वास्तविक औसत मान Y_i, के बराबर ही होता है, यथा – $\bar{\hat{Y}}_i = \bar{Y}$

c) $\frac{1}{n}\sum e_i = \bar{e}_i = 0.$

d) $Cov(e_i, X_{2i}) = Cov(e_i, X_{3i}) = 0$, यथा, अवशिष्ट e_i पद X_{2i} और X_{3i} से सहसंबद्ध होता है। संकेताक्षरों में, $(\sum e_i X_{2i}) = (\sum e_i X_{3i}) = 0.$

e) $Cov\left(e_i, \hat{Y}_i\right) = 0$, यथा, अवशिष्ट e_i पद $\hat{Y}_i$ और $\sum e_i \ \hat{Y} = 0$ से सहसंबद्ध नहीं होता है। जब पद r_{23}, चर X_2 और X_3 के बीच सहसंबंध गुणांक अंक 1 की ओर बढ़ता है तो चर b_2 और b_3 के विचर $\sigma^2, \sum x_{2i}^2$ अथवा $\sum x_{3i}^2$ के ज्ञात मानों की ओर बढ़ते हैं।

f) ऊपर f) को ध्यान में रखते हुए, पद r_{23} और $\sum x_{2i}^2$ अथवा $\sum x_{3i}^2$ के मान ज्ञात हों तो OLS आकलकों के विचरण चर σ^2 के प्रत्यक्ष समानुपातिक होते हैं।

g) CLRM की अवधारणाओं को देखते हुए, आंशिक समाश्रयण गुणांक के OLS आकलक न केवल रैखिक और अनभिनत होते हैं, बल्कि वे सभी अनभिनत आकलकों के वर्ग में न्यूनतम विचरण भी दर्शाते हैं, यथा वे BLUE होते हैं। दूसरे शब्दों में, वे गॉस–मार्कोव प्रमेय को संतुष्ट करते हैं।

7.4 आकलन की अधिकतम संभावना विधि

'अधिकतम संभावना आकलन' की विधि एक प्रायिकता बंटन फलन (pdf) के प्राचलों का आकलन करती है। यह यह इस फलन (pdf) के प्रायिकता फलन को अधिकतम करके किया जाता है। इसीलिए प्रायिकता फलन को अधिकतम करने वाले आकलकों को 'अधिकतम प्रायिकता आकलक' कहा जाता है। इस अवधारणा को बेहतर ढंग से समझने के लिए, आइए, पहले अधिकतम संभावना आकलक $(\tilde{\beta})$ अवकलित करते हैं। यहाँ हम संकेताक्षर $(\tilde{\beta})$ का प्रयोग करते हैं ताकि ML आकलकों को OLS आकलकों $(\hat{\beta})$ से भिन्न दर्शाया जा सके। मान लीजिए कि pdf सामान्य बंटन का अनुसरण करता है। तदनुसार, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

इस pdf के दोनों पक्षों पर संभावना फलन का लघुगणक लेने पर निम्नलिखित समीकरण प्राप्त होता है –

$$lnL = -\frac{n}{2}ln\sigma^2 - \frac{n}{2}ln(2\pi) - \frac{1}{2}\sum\frac{(Y_i - \beta_1 - \beta_2 X_{2i} - \beta_k X_{ki})^2}{\sigma^2}$$

उपर्युक्त फलन का अंशतः चरों $\beta_1, \beta_2, ..., \beta_k$ और σ^2 के संदर्भ में अवकलन करने पर निम्नलिखित $(k+1)$ समीकरण प्राप्त होते हैं –

$$\frac{\partial \ln L}{\partial \beta_1} = \frac{1}{\sigma^2}\sum(Y_i - \beta_1 - \beta_2 X_{2i} - \ldots - \beta_k X_{ki})\,(-1) \quad (1)$$

$$\frac{\partial \ln L}{\partial \beta_2} = \frac{1}{\sigma^2}\sum(Y_i - \beta_1 - \beta_2 X_{2i} - \ldots - \beta_k X_{ki})\,(-X_{2i}) \quad (2)$$

....

$$\frac{\partial \ln L}{\partial \beta_k} = \frac{1}{\sigma^2}\sum(Y_i - \beta_1 - \beta_2 X_{2i} - \ldots - \beta_k X_{ki})\,(-X_{ki}) \quad (k)$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^2}\sum(Y_i - \beta_1 - \beta_2 X_{2i} - \ldots - \beta_k X_{ki})^2 \quad (k+1)$$

इन समीकरणों को शून्य पर निर्धारित करके (यथा, इष्टतमीकरण के लिए प्रथम–कोटि शर्त लागू करके) तथा शर्तों को पुनर्व्यवस्थित करके और साथ ही, $\tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_k$ और $\tilde{\sigma}^2$ को अधिकतम संभावना आकलों (MLEs) के रूप में निरूपित करके हमें निम्नलिखित समीकरण प्राप्त होते हैं–

$$\sum Y_i = n\tilde{\beta}_1 + \tilde{\beta}_2 \sum X_{2i} + \cdots + \tilde{\beta}_k \sum X_{ki}$$

$$\sum Y_i X_{2i} = \tilde{\beta}_1 + \sum X_{2i} + \tilde{\beta}_2 \sum X_{2i}^2 + \ldots + \tilde{\beta}_k \sum X_{2i} X_{ki}$$

……………………………………………………………

$$\sum Y_i X_{ki} = \tilde{\beta}_1 \sum X_{ki} + \tilde{\beta}_2 \sum X_{2i} X_{ki} + \ldots + \tilde{\beta}_k \sum X_{ki}^2$$

उपर्युक्त समीकरण सटीक रूप से आकलन की OLS विधि के सामान्य समीकरण हैं। अतएव, पद $\tilde{\beta}$ के अधिकतम संभावना आकल पद $\tilde{\beta}$ के OLS आकलों की भाँति ही हैं। तदनुसार, अधिकतम संभावना आकलों (अथवा OLS आकलकों) को ऊपर दिए गए $(k+1)^{वें}$ समीकरण में प्रतिस्थापित कर सरलीकृत करते हुए हमें पद σ^2 के अधिकतम संभावना आकल निम्नवत् प्राप्त होते हैं –

$$\tilde{\sigma}^2 = \frac{1}{n}\sum\left(Y_i - \tilde{\beta}_1 - \tilde{\beta}_2 X_{2i} - \ldots - \tilde{\beta}_k X_{ki}\right)^2$$

$$= \frac{1}{n}\sum \hat{u}_i^2$$

यह ध्यान देने की बात है कि यह आकलक $\hat{\sigma}^2 = \sum u_i^2/(n-k)$ के OLS आकलक से भिन्न होता है। चूँकि परवर्ती σ^2 का कोई अनभिनत आकलक होता है, पद का $\tilde{\sigma}^2$ अधिकतम संभावना आकल एक अभिनत आकलक होता है। हालाँकि आप देखेंगे कि उपगामितः $\tilde{\sigma}^2$ भी अनभिनत ही होता है। इसका अर्थ है कि स्पर्शोन्मुख रूप से अधिकतम संभावना आकल (MLE) और साधारण न्यूनतम वर्ग (OLS) के आकल एक समान ही होते हैं। इसके अलावा, अधिकतम संभावना आकल अभिनत होता है परन्तु वह समनुरूप होता है।

बहु समाश्रयण मॉडलों के लिए उपर्युक्त बीजीय निष्पीड़न दुर्वह हो जाते हैं। अतः हम किसी भी बहु समाश्रयण मॉडल को दर्शाने के लिए हम आव्यूह बीजगणित (जिस पर हमने अपने पिछले पाठ्यक्रम BECC 104 में चर्चा की है) का सहारा ले सकते हैं। इसके लिए, मान लीजिए कि

$$X_0 = \begin{bmatrix} 1 \\ X_{02} \\ X_{03} \\ \cdot \\ \cdot \\ \cdot \\ X_{0k} \end{bmatrix} \qquad \ldots (7.8)$$

सभी *X* चरों के मानों का वह सदिश रूप है जिसके लिए हम $\hat{Y}_0$ का पूर्वानुमान करना चाहते हैं, जो कि चर *Y* का औसत पूर्वानुमान होगा। अब अदिश रूप में आकलित बहु समाश्रयण समीकरण निम्नवत् होगा –

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \ldots + \hat{\beta}_k X_{ki} + u_i \qquad \ldots (7.9)$$

आव्यूह संकेतन में समीकरण (7.9) को संक्षिप्ततः निम्नवत् लिखा जा सकता है –

$$\hat{Y}_i = x'_i \hat{\beta} \qquad \ldots (7.10)$$

जहाँ $x'_i = [1 \quad X_{2i} \quad X_{3i} \ldots \quad X_{ki}]$

और

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\beta}_k \end{bmatrix}$$

समीकरण (7.9) अथवा (7.10) ही पद Y_i का औसत पूर्वानुमान है, जो कि ज्ञात x'_i के अनुरूप है। अतएव, यदि x'_i समीकरण (7.8) में दिए गए मान के अनुरूप ही हो तो समीकरण (7.10) निम्नवत् दिखाई देने लगता है –

$$\left(\hat{Y}_I | x'_0\right) = x'_0 \hat{\beta} \qquad \text{... (7.11)}$$

जहाँ, x_0 के मान निर्दिष्ट हैं। आप देखेंगे कि समीकरण (7.11) हमें $E(Y_i|x'_0)$ का एक अनभिनत पूर्वानुमान देता है क्योंकि $E\left(x'_0 \hat{\beta}\right) = x'_0 \hat{\beta}$ होता है। पद $\left(\hat{Y}_0 | x'_0\right)$ के विचरण का आकल निम्नवत् दर्शाया जाता है –

$$Var\left(\hat{Y}_0 | x'_0\right) = \sigma^2 x'_0 (X'X)^{-1} x_0 \qquad \text{... (7.12)}$$

जहाँ σ^2 पद u_i का विचरण है, जबकि u_i, x'_0 उन X चरों के ज्ञात मान हैं जिनके लिए हम भावी मानों का पूर्वानुमान करना चाहते हैं, और $(X'X)$ आव्यूह है। व्यवहारतः, हम σ^2 के स्थान पर उसका अनभिनत आकलक σ^2 ही लिखते हैं।

बोध प्रश्न 1

[दिए गए स्थान में प्रश्नों के उत्तर 50–100 शब्दों में दें।]

1) किसी बहु समाश्रयण मॉडल का सरलतम रूप सोदाहरण निर्दिष्ट करें। यह सबसे सरल रूप क्यों है, स्पष्ट करें।

....

....

....

....

2) पारंपरिक रैखिक समाश्रयण मॉडल (CLRM) के लिए किए गए आकलनों को वृहद पदों में एक–एक कर बताएँ। बहु समाश्रयण मॉडल के लिए अतिरिक्त अवधारणा क्या है?

....

....

....

....

3) किसी बहु समाश्रयण मॉडल के आकलित प्राचलों की व्याख्या कैसे की जाती है?

....

....

....

....

4) उस गुणधर्म का प्रतिफल निर्दिष्ट करें जिससे साधारण न्यूनतम वर्ग (OLS) के आकलक गॉस–मार्कोव प्रमेय का पालन करते हैं?

....

....

....

....

7.5 निर्धारण गुणांक : R^2

बहु समाश्रयण में सुमानकता अथवा समंजन सुष्ठुता का मापदंड R^2 द्वारा दर्शाया जाता है। इसे ही 'निर्धारण का गुणांक' भी कहा जाता है। यह 'वर्गों के व्याख्यायित योगफल' और 'वर्गों के कुल योग' का अनुपात होता है। पद R^2 अवकलित करने के लिए हम प्रतिदर्श समाश्रयण फलन या समीकरण को निम्नानुसार मानकर चलते हैं–

$$Y_i = b_1 + b_2X_{2i} + b_3X_{3i} + e_i \quad \text{... (7.13)}$$

जहाँ $b_1 = \bar{Y} - b_2\bar{X}_2 - b_3\bar{X}_3$ होता है। पद b_1 को समीकरण (7.13) में प्रतिस्थापित कर X_{2i} और X_{3i} में उनके माध्य लेकर हमें निम्नलिखित समीकरण प्राप्त होता है –

$$Y_i = \bar{Y} - b_2\bar{X}_2 - b_3\bar{X}_3 + b_2X_{2i} + b_3X_{3i} + e_i$$

अतएव, $Y_i - \bar{Y} = b_2(X_{2i} - \bar{X}_2) + b_3(X_{3i} - \bar{X}_3) + e_i$

उपर्युक्त को रोमन लिपि के छोटे अक्षरों में दोबारा लिखकर, यथा, माध्य से विचलन के रूप में विचार कर, हमें निम्नलिखित समीकरण प्राप्त होता है –

$$y_i = b_2x_{2i} + b_3x_{3i} + e_i \quad \text{... (7.14)}$$

हमें ज्ञात है कि $\hat{Y}_i - \bar{Y} = \hat{Y}_i$ जहाँ –

$$\hat{Y}_i = b_1 + b_2X_{2i} + b_3X_{3i}$$

$$\bar{Y} = b_1 + b_2\bar{X}_2 + b_3\bar{X}_3$$

$$\therefore\ \hat{Y}_i - \bar{Y} = (b_1 + b_2X_{2i} + b_3X_{3i}) - (b_1 + b_2\bar{X}_2 + b_3\bar{X}_3)$$

$$\hat{Y}_i - \bar{Y} = b_2(X_{2i} - \bar{X}_2) + b_3(X_{3i} - \bar{X}_3)$$

$$\hat{Y}_i - \bar{Y} = b_2x_{2i} + b_3x_{3i} \quad \text{... (7.15)}$$

अब निम्नलिखित समीकरण पर विचार करें –

$$y_i = \hat{y}_i + e_i$$

दोनों पक्षों का वर्ग कर उनका योगफल लेने पर हमें निम्नलिखित समीकरण प्राप्त होंगे –

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 + 2\sum \hat{y}_i e_i$$

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 + 0 \text{ चूँकि } Cov(\hat{y}_i, e_0) = 0$$

$$\therefore \sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 \qquad \text{... (7.16)}$$

इसका अर्थ है कि TSS = ESS + RSS अब निम्नलिखित समीकरण पर विचार करें –

$R^2 = \dfrac{ESS}{TSS}$ जहाँ ESS = $\sum \hat{y}_i^2$

चूँकि $e_i = y_i - \hat{y}_i$ जहाँ $\hat{y}_i = b_2 x_{2i} + b_3 x_{3i}$ हमें प्राप्त होता है : $e_i = y_i - (b_2 x_{2i} + b_3 x_{3i})$ अब,

$$\sum e_i^2 = \sum (e_i e_i)$$

$$= \sum [e_i (y_i - b_2 x_{2i} - b_3 x_{3i})]$$

$$= \sum e_i y_i - b_2 \sum e_i x_{2i} - b_3 \sum e_i x_{3i}$$

$\therefore \sum e_i^2 = \sum e_i y_i$ [चूँकि $\sum e_i x_{2i} = \sum e_i x_{3i} = 0$].

$$\sum e_i^2 = \sum y_i e_i = \sum y_i (y_i - b_2 x_{2i} - b_3 x_{3i})$$

$$\Rightarrow \sum e_i^2 = \sum y_i^2 - b_2 \sum y_i x_{2i} - b_3 \sum y_i x_{3i} \qquad \text{... (7.17)}$$

समीकरण (7.17) को समीकरण (7.16) में प्रयोग करने पर हमें निम्नलिखित समीकरण प्राप्त होगा –

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum y_i^2 - b_2 \sum y_i x_{2i} - b_3 \sum y_i x_{3i}$$

$\Rightarrow \sum \hat{y}_i^2 = b_2 \sum y_i x_{2i} + b_3 \sum y_i x_{3i}$ = ESS

अतएव, $R^2 = \dfrac{ESS}{TSS} = \dfrac{b_2 \sum y_i x_{2i} + b_3 \sum y_i x_{3i}}{\sum y_i^2}$ ... (7.18)

किसी k–चर बहु समाश्रयण मॉडल में पद R^2 और आंशिक समाश्रयण गुणांक (b_i) के विचरण के बीच संबंध निम्नलिखित समीकरण द्वारा दर्शाया जाता है–

$$V(b_i) = \frac{\sigma^2}{\sum x_y^2} - \left(\frac{1}{1 - R_i^2}\right)$$

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum e_i^2}{\sum y_i^2}$$

$$= 1 - \frac{(n-k)\hat{\sigma}^2}{(n-1)Sy^2}$$

$\because \sum e_i^2$ or $\hat{\sigma}^2 = \dfrac{\sum e_i^2}{n-k} \Rightarrow \sum e_i^2 = (n-k)\hat{\sigma}^2$

$$Sy^2 = \frac{\sum y_i^2}{n-1} \Rightarrow \sum y_i^2 = (n-1)Sy^2$$

7.6 समायोजित R^2

एक ही आश्रित चर के साथ दो समाश्रयण मॉडलों की तुलना करते हुए, जबकि X चरों की संख्या भिन्न हो, उच्चतम R^2 वाले मॉडल को चुनने में सावधानी बरतनी चाहिए। यह समझने के लिए कि यह क्यों महत्वपूर्ण है, आइए, इस पर विचार करें – $R^2 = \frac{ESS}{TSS} = \frac{1-RSS}{TSS} = \frac{1-\sum e_i^2}{\sum y_i^2}$

यह ध्यान देने की बात है कि जैसे–जैसे व्याख्यात्मक चरों की संख्या बढ़ती है, अंश ESS भी बढ़ता है। दूसरे शब्दों में, जब स्वतंत्र चर की संख्या, k, में वृद्धि होती है तो चर R^2 का मान भी बढ़ता है। चर R^2 के उपर्युक्त निष्पीड़न का अर्थ है कि पद R^2 मॉडल में स्वतंत्र चरों की संख्या को कोई महत्व नहीं देता है। इसी कारण व्याख्यात्मक चरों की भिन्न संख्या के साथ दो समाश्रयणों की तुलना करने के लिए हमें R^2 का प्रयोग नहीं करना चाहिए। अब हमें निर्धारण के एक वैकल्पिक गुणांक की आवश्यकता होगी, जो कि आकलित प्राचलों की संख्या, यथा k, को ध्यान में रखेगा। इसके लिए हम एक ऐसे मापदंड को लेकर चलते हैं जिसे समायोजित R^2 कहा जाता है और निम्नानुसार परिभाषित किया जाता है–

$$\bar{R}^2 = 1 - \frac{RSS/n-k}{TSS/n-1} = 1 - \frac{\sum e_i^2/(n-k)}{\sum y_i^2/(n-1)}$$

जहाँ k अवरोधन पद के साथ मॉडल में प्राचलों की संख्या है। उपर्युक्त को सरल रूप में निम्नवत् भी लिखा जा सकता है –

$$\bar{R}^2 = \frac{1-\hat{\sigma}^2}{S_y^2}$$

जहाँ $\hat{\sigma}^2$ एक ऐसा अवशिष्ट विचरण है जो वास्तविक σ^2 का एक अनभिनत आकलक होता है, जबकि S_y^2 चर Y का प्रतिदर्श विचरण है। अब $\bar{R}^2$ और R^2 के बीच संबंध निम्नलिखित समीकरण से दर्शाया जाता है –

$$\bar{R}^2 = 1 - (1-R^2)\frac{n-1}{n-k} \qquad \text{... (7.19)}$$

अब यह तय करने के लिए कि दो समाश्रयणों की तुलना करने के लिए हमें चर R^2 प्रयोग करना चाहिए अथवा चर $\bar{R}^2$, हमें निम्नलिखित बातों पर ध्यान देना चाहिए–

i) यदि $k > 1, \bar{R}^2 < R^2$ हो तो इसका अर्थ होगा है कि जब व्याख्यात्मक चरों की संख्या का X बढ़ती है तो समायोजित R^2 सामान्य R^2 की अपेक्षा कम वृद्धि दर्शाता है।

ii) पद $\bar{R}^2$ ऋणात्मक हो सकता है परंतु पद R^2 अनिवार्यतः कोई ऋणेतर चर ही होता है। ऐसा इसलिए है कि समीकरण (7.18) में –

यदि $R^2 = 1, \bar{R}^2 = 1.$

यदि $R^2 = 0$ हो तो $\bar{R}^2 = \frac{1-k}{n-k}$ होता है। अतः, यदि $k > 1$ हो तो $\bar{R}^2 < 0$ होगा।

इस प्रकार, समायोजित R^2 ऋणात्मक हो सकता है। ऐसे उदाहरणों में $\bar{R}^2$ के मान को शून्य मानकर चलने की प्रथा है। तदनुसार, किसी समाश्रयण की सुमानकता दर्शाने के लिए उक्त दो में से कौन–सा श्रेष्ठ है, इस आशय का निर्णायक मत देना संभव नहीं होता। बहरहाल, व्यवहारतः अनेक समाश्रयण मॉडलों में समायोजित R^2 का प्रयोग मॉडल की सुमानकता निर्धारित करने के

लिए किया जाता है क्योंकि यह समाश्रयियों की संख्या को ध्यान में रखता है और इसके चलते आकलित प्राचलों की संख्या को भी ध्यान में रखता है।

बोध प्रश्न 2

[दिए गए स्थान में प्रश्नों के उत्तर 50–100 शब्दों में दें।]

1) साधारण न्यूनतम वर्ग (OLS) आकल और अधिकतम संभावना आकल (MLE) के बीच अंतर स्पष्ट करें।

....

....

....

....

2) पद R^2 को कैसे परिभाषित किया जाता है? उपयुक्त निष्पीड़नों के साथ स्पष्ट करें।

....

....

....

....

3) सामान्य R^2 की तुलना में समायोजित R^2 के महत्व पर प्रकाश डालें।

....

....

....

....

4) सामान्य R^2 और समायोजित R^2 परस्पर कैसे संबद्ध हैं? इन दोनों चरों के बीच अंतर स्पष्ट करें।

....

....

....

....

5) समायोजित R^2 के ऋणात्मक होने की स्थिति से व्यवहारतः कैसे निपटा जाता है?

....

....

....

....

7.7 सार–संक्षेप

इस इकाई में बहु समाश्रयण मॉडल और उसके अनुमानों का वर्णन किया गया है। पारंपरिक रैखिक समाश्रयण मॉडल (CLRM) की अवधारणाओं को दोहराते हुए यह इकाई दर्शाती करती है कि बहु समाश्रयण मॉडलों में बहुसंरेखता पर एक अतिरिक्त अवधारणा कैसे आवश्यक है। इसमें प्राचलों की व्याख्या, यथा अवरोधन, और आंशिक समाश्रयण गुणांक को स्पष्ट किया गया है।

इस इकाई में सबसे पहले OLS (साधारण न्यूनतम वर्ग) विधि द्वारा बहु समाश्रयण मॉडल के प्राचलों के आकलन पर चर्चा की गई है। एक वैकल्पिक विधि अर्थात् अधिकतम संभावना अनुमान की विधि (MLE) को इसी इकाई में आगे प्रस्तुत किया गया है।

यह दर्शाया गया है कि स्पर्शोन्मुख रूप से OLS और MLE संपाती होते हैं। 'निर्धारण के गुणांक' या सुमानकता की अवधारणा का भी यहाँ वर्णन किया गया है। अंत में, समायोजित R^2 की आवश्यकता और प्रयोग के विषय में बताया गया है।

7.8 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1) बहु समाश्रयण मॉडल एक ऐसा निदर्श प्रस्तुत करता है जिसमें एक से अधिक स्वतंत्र या व्याख्यात्मक चर होते हैं। तदनुसार, सबसे सरल बहु समाश्रयण मॉडल एक आश्रित चर और दो स्वतंत्र चरों के साथ सामने आता है। इसी प्रकार का एक मॉडल $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$ के रूप में निर्दिष्ट किया जाता है। उदाहरणस्वरूप एक ऐसे उत्पादन फलन को देखें जिसमें आश्रित चर उत्पादन दर्शाता हो, जबकि स्वतंत्र चर दो निविष्ट, यथा श्रम और पूँजी, दर्शाते हों। व्यष्टि अर्थशास्त्र में यह आश्रित चर के रूप में किसी वस्तु के उपभोग और दो स्वतंत्र चरों के रूप में मूल्य एवं आय के बीच संबंध दर्शा सकता है।

2) (i) यह मॉडल अपने प्राचलों में रेखीय होता है; (ii) चर u_i और X_i सहसंबद्ध नहीं होते, यथा, $cov(u_i, X_{2i}) = Cov(u_i, X_{3i}) = 0$; (iii) त्रुटि पद की सशर्त प्रत्याशा शून्य होती है, यथा, $E(u_i | X_{2i}, X_{3i}) = 0$; (iv) त्रुटिपद सहसंबद्ध नहीं होते अथवा यहाँ कोई स्वसहसंबंध देखने में नहीं आता, यथा, $\mathrm{cov}(u_i, u_j) = 0$; (v) यहाँ समविसारिता दिखाई देती है अथवा त्रुटि विचरण भिन्नता नहीं दर्शाते, यथा, $var(u_i^2) = \sigma^2$; (vi) शून्य बहुसंरेखता अथवा पूर्ण संरेखता, यथा, *Corr* $(X_i, X_j) \neq 1$; (vii) प्रेक्षणों की संख्या (n) आकलित प्राचलों की संख्या (k) से अधिक होती है; (viii) यहाँ कोई विनिर्देशन अभिनति नहीं देखी जाती, यथा, मॉडल में न तो कोई प्रासंगिक चर छूटा है और न ही उसमें कोई अप्रासंगिक चर शामिल किया गया है; तथा (ix) चर $\boldsymbol{X}$ और Y में कोई मापन त्रुटि नहीं है।

   उपर्युक्त अवधारणाओं में अवधारणा (vi) ही बहु समाश्रयण मॉडलों में अपेक्षित अतिरिक्त अवधारणा है।

3) अवरोधन β_1 आश्रित चर Y का प्रत्याशित मान मापता है, बशर्त व्याख्यात्मक चरों X_2 और X_3 के मान ज्ञात हों। दूसरी ओर, β_2 चर Y के औसत मान में परिवर्तन [यथा, E(Y)] मापता है, जो कि X_3 के मान को नियत रखते हुए चर X_2 में प्रति इकाई परिवर्तन होता है। इसका अर्थ है कि $\beta_2 = \frac{\Delta E(Y)}{\Delta X_2}$ होता है। इसी भाँति β_3 को भी परिभाषित किया जा सकता है।

4) CLRM की अवधारणाओं के तहत आंशिक समाश्रयण गुणांक के OLS आकलक न केवल रैखिक और अनभिनत होते हैं, बल्कि सभी अनभिनत आकलकों के वर्ग में न्यूनतम विचरण भी दिखाई देता है, यथा, वे BLUE (सर्वोत्तम रैखिक अनभिनत आकलक) होते हैं। यही वह गुणधर्म है जो OLS आकलों से गॉस–मार्कोव प्रमेय को संतुष्ट कराता है।

बोध प्रश्न 2

1) OLS आकलक वर्गों के अवशिष्ट योगफल को कम करके प्राप्त किए जाते हैं, यथा Min $\sum e_i^2 = \sum(Y_i - \hat{Y}_i)^2$ दूसरी ओर, MLE आकलक संबंधित pdf के 'संभावना फलन' को अधिकतम करके प्राप्त किए जाते हैं। इस प्रकार, इन दो विधियों के उपागम में यह एक बुनियादी अंतर है। हालाँकि, एक बार जब प्रथम–कोटि शर्तें लागू हो जाती हैं और सरलीकृत कर दी जाती हैं तो हमें MLE उपागम में जो समीकरण प्राप्त होते हैं वे उन सामान्य समीकरणों के समान ही होते हैं जो हमें OLS विधि में मिलते हैं। इसलिए, उन समीकरणों को हल करके प्राप्त किए गए प्राचलों के आकल एक समान ही होते हैं। बहरहाल, चर σ^2 के अनभिनत आकलन से संबंधित एक अनिवार्य अंतर भी होता है। OLS विधि में इस अनभिनत आकलन के लिए निष्पीड़न का भाजक *'n-k'* होता है जबकि MLE विधि में यह *'n'* होता है। यह महत्वपूर्ण अंतर छोटे प्रतिदर्शों के अभिनत MLE उपागम में σ^2 के आकल प्रदान करता है। बड़े प्रतिदर्शों के लिए यह अंतर अनभिनत होता है। इसीलिए MLE और OLS के आकल संपाती होते हैं और स्पर्शोन्मुख रूप से, OLS आकल और MLE आकल एक दूसरे से मेल खाते हैं।

2) किसी भी ऐसे दो स्वतंत्र चरों वाले बहु समाश्रयण मॉडल में जिसका प्रतिदर्श सममाश्रण फलन $Y_i = b_1 + b_2X_{2i} + b_3X_{3i} + e_i$ के रूप में दिया गया हो

   तो R^2 को $R^2 = \frac{ESS}{TSS} = \frac{b_2\sum y_i x_{2i} + b_3 \sum y_i x_{3i}}{\sum y_i^2}$ के रूप में दर्शाया जाता है।

3) व्याख्यात्मक चरों की भिन्न संख्या के साथ दो बहु समाश्रयणों की तुलना करने के लिए R^2 पर निर्भर होना भ्रामक हो सकता है। ऐसा इसलिए है कि R^2 व्याख्यात्मक चरों की संख्या को हिसाब में नहीं लेता है।

4) वे इस प्रकार संबंधित हैं : $\overline{R}^2 = 1-(1-R^2)\frac{n-1}{n-k}$ एक महत्वपूर्ण अंतर यह है कि जबकि R^2 ऋणात्मक नहीं हो सकता है, समायोजित R^2 ऋणात्मक हो सकता है।

5) जब यह ऋणात्मक होता है तो इसे परंपरागत रूप से शून्य के रूप में लिया जाता है।

इकाई 8 बहु समाश्रयण मॉडल : अनुमिति

इकाई की रूपरेखा

8.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप निम्न के लिए सक्षम होंगे:

- बहु प्रतिगमन के मामले में सामान्यता की धारणा की आवश्यकता की व्याख्या
- करना;
- व्यक्तिगत अनुमानकों पर परिकल्पना के परीक्षण की प्रक्रिया का वर्णन करना;
- प्रतिगमन प्रतिमान के समग्र महत्व का परीक्षण करना;
- दो प्रतिगमन गुणांकों की समानता के लिए परीक्षण;
- चाउ परीक्षण को उपयोग करने की प्रक्रिया की व्याख्या करना;
- बहु प्रतिगमन प्रतिमान के आधार पर भविष्यवाणी करना; तथा
- व्यक्तिगत और संयुक्त दोनों, परिकल्पना के परीक्षण से प्राप्त परिणामों की व्याख्या करना ।

* डॉ. पूजा शर्मा, सहायक प्राध्यापक, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय।

8.1 विषय–प्रवेश

पिछली इकाई में हमने बहु प्रतिगमन प्रतिमानों की व्याख्या और आकलन के बारे में चर्चा की थी। हमने उन मान्यताओं को देखा जो सामान्य न्यूनतम वर्ग (OLS) और अधिकतम संभावना (ML) अनुमान के लिए आवश्यक हैं। वर्तमान इकाई में हम बहु प्रतिगमन प्रतिमान में परिकल्पना परीक्षण की विधियों को देखेंगे।

याद कीजिए कि इस पाठ्यक्रम की इकाई 3 में हमने परिकल्पना परीक्षण की प्रक्रिया का उल्लेख किया था। इसके अलावा, इकाई 5 में हमने दो परिवर्तनीय प्रतिगमन प्रतिमान के मामले में परिकल्पना परीक्षण की प्रक्रिया की व्याख्या की। आइए अब हम परिकल्पना परीक्षण की प्रक्रिया को बहु प्रतिगमन प्रतिमानों तक विस्तारित करें। जहां तक परिकल्पना परीक्षण का संबंध है, कई प्रतिगमन प्रतिमान में दो परिदृश्य हो सकते हैं: (i) व्यक्तिगत गुणांक का परीक्षण, और (ii) कुछ मापदंडों का संयुक्त परीक्षण। हम चाउ परीक्षण को उपयोग करके प्रतिगमन प्रतिमान की संरचनात्मक स्थिरता के परीक्षण की विधि पर चर्चा करते हैं। इसके अलावा हम तीन महत्वपूर्ण परीक्षणों पर चर्चा करते हैं जैसे, संभावना अनुपात परीक्षण, वाल्ड परीक्षण, और लैग्रेंज गुणक परीक्षण। अंत में हम बहु प्रतिगमन समीकरण के आधार पर भविष्यवाणी करते है ।

परिकल्पना परीक्षण में एक धारणा यह है कि त्रुटि पद u_i सामान्य वितरण का अनुसरण करता है। क्या एक पद की सामान्यता के परीक्षण के लिए कोई तरीका है? हम इस मुद्दे पर भी चर्चा करेंगे। हालाँकि, आइए हम कई प्रतिगमन प्रतिमान की बुनियादी मान्यताओं के अवलोकन के साथ शुरू करें |

8.2 बहु प्रतिगमन प्रतिमान की मान्यता

इकाई 7 में हमने दो व्याख्यात्मक चरों X_2 और X_3 वाले बहु प्रतिगमन प्रतिमान पर विचार किया। प्रसंभाव्य त्रुटि पद u_i है

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \qquad \dots (8.1)$$

8.2.1 पारंपरिक मान्यताएं

एकाधिक प्रतिगमन प्रतिमान के संबंध में सात धारणाएं हैं। इनमें से अधिकांश धारणाएँ त्रुटि पद के संबंध में हैं। इन मान्यताओं के बारे में हमने पिछली इकाई में चर्चा की थी। आइए संक्षेप में उन धारणाओं का फिर से उल्लेख करें |

a) प्रतिगमन प्रतिमान मापदंडों और पद में रैखिक है |

b) त्रुटि पदों का माध्य निराकरणीय है। दूसरे शब्दों में, व्याख्यात्मक पद X_{2i} और X_{3i} पर सशर्त त्रुटि पद का अपेक्षित मान निराकरणीय है

$$E(u_i) = 0 \text{ or } E(u_i | X_{2i}, X_{3i}) = 0$$

c) त्रुटि पदों के बीच कोई क्रमिक सहसंबंध (या स्वत: सहसंबंध) नहीं है। त्रुटि पदों सहसंबद्ध नहीं हैं। इसका तात्पर्य है कि i^{th} अवलोकन u_i से संबद्ध त्रुटि पद और j^{th} अवलोकन से संबद्ध त्रुटि पद के बीच सहप्रसरण, u_j निराकरणीय है

$$\text{cov}(u_i, u_j) = 0$$

d) समविसारिता: समविसारिता की धारणा बताती है कि त्रुटि विपदण पूरी आबादी में स्थिर है। प्रत्येक अवलोकन से जुड़े त्रुटि पद के प्रसरण का एक ही प्रसरण होता है

$$var(u_i) = \sigma^2$$

e) व्याख्यात्मक पद की बहिर्जातता: व्याख्यात्मक पद और त्रुटि पद के बीच कोई संबंध नहीं है। इस धारणा को बहिर्जातता भी कहा जाता है, क्योंकि व्याख्यात्मक पद को बहिर्जात माना जाता है (बाहर से दिया गया है; एक्स प्रतिमान के अंदर निर्धारित नहीं है)। इसके विपरीत, Y प्रतिमान के भीतर निर्धारित होता है। जब व्याख्यात्मक पद को त्रुटि पद के साथ सहसंबद्ध किया जाता है, तो इसे अंतर्जात समस्या कहा जाता है। इस समस्या से बचने के लिए, हम मानते हैं कि व्याख्यात्मक पद नमूनों में स्थिर रखे जाते हैं।

f) स्वतंत्र पद एक दूसरे के रैखिक संयोजन नहीं हैं। यदि स्वतंत्र चरों के बीच पूर्ण रैखिक संबंध है, तो व्याख्यात्मक पद सामंजस्य में चलते हैं और मापदंडों का अनुमान लगाना संभव नहीं है। इसे बहुसंरेखण समस्या भी कहते हैं।

g) त्रुटि पद सामान्य रूप से वितरित किया जाता है। मापदंडों के आकलन के लिए ओएलएस पद्धति में यह धारणा आवश्यक नहीं है। यह विश्वास अंतराल और परिकल्पना परीक्षण के निर्माण के लिए आवश्यक है। पिछली इकाई में चर्चा की गई अधिकतम संभावना विधि में, मापदंडों का अनुमान लगाने के लिए हमने मान लिया था कि त्रुटि पद सामान्य वितरण का अनुसरण करता है.

8.2.2 त्रुटि-पद (पद) की सामान्यता के लिए परीक्षण

जैसा कि पहले बताया गया है हम त्रुटि पद की सामान्यता की धारणा को देखते हैं। त्रुटि पद की सामान्यता का परीक्षण करने के लिए हम जार्के–बेरा परीक्षण (जिसे अक्सर JB परीक्षण कहा जाता है) उपयोग करते हैं। यह एक स्पर्शोन्मुख या बड़ा नमूना परीक्षण है। हम प्रतिगमन प्रतिमान में त्रुटि पदों को नहीं जानते हैं; हम अवशेषों को जानते हैं। इसलिए, जेबी परीक्षण OLS अवशेषों पर आधारित है। आँकड़ों से दो अवधारणाओं को याद करें: विषमता और वक्रता-मात्रा (कुर्टोसिस)। एक तिरछा वक्र (अर्थात असममित) एक सामान्य वक्र से भिन्न होता है। एक लेप्टोकुर्टिक या प्लेटीकुर्टिक वक्र (यानी, लंबा या छोटा ऊंचाई में) सामान्य वक्र से अलग होता है। जार्के–बेरा परीक्षण विषमता और वक्रता-मात्रा (कुर्टोसिस) के उपायों का उपयोग करता है।

हम जानते हैं कि एक सामान्य वितरण के लिए S = 0 और K = 3 होना चाहिए । इन दो मानों से एक संकेत विचलन पुष्टि करेगा कि पद सामान्य रूप से वितरित नहीं है ।

जार्के और बेरा ने *JB*-सांख्यिकी का निर्माण किया

$$JB = \frac{n}{6}\left[S^2 + \frac{(K-3)^2}{4}\right] \qquad \ldots (8.2)$$

जहां

n = नमूने का आकार

S = तिरछापन का माप $(\frac{\mu_3}{\sigma^3})$

K = कुर्टोसिस का माप $(\frac{\mu_4}{\mu_2^2})$

विषमता और वक्रता-मात्रा (कुर्टोसिस) को एक पद के क्षणों के संदर्भ में मापा जाता है। जैसा कि आप BECC 107, इकाई 4 से जानते हैं, पद X_i के rth परिघात (moment) की गणना करने का सूत्र है

$$\mu_r = \frac{1}{n}\sum_{i-1}^{n} f_i(X_i - \bar{X})^r \qquad \ldots (8.4)$$

प्रसरण (variance) दूसरा परिघात है μ_2 |

समीकरण (8.2) में JB आँकड़ा स्वतंत्रता के 2 डिग्री के साथ काई-वर्ग वितरण का अनुसरण करता है $\sim\chi^2_{(2)}$ |

यदि कोई पद सामान्य वितरण का अनुसरण करता है, तो आइए हम JB आँकड़ों का मान ज्ञात करें। सामान्य वितरण के लिए, जैसा कि ऊपर बताया गया है S = 0 और K = 3। इन मानों को समीकरण (8.2) में प्रतिस्थापित करने पर हम निम्नलिखित समीकरण प्राप्त करते हैं

$$JB = \frac{n}{6}[0 + 0] = \frac{n}{6} \times 0 = 0 \qquad \ldots (8.3)$$

एक पद के लिए सामान्य रूप से वितरित नहीं किया गया जेबी आँकड़े तेजी से बड़े मूल्यों को ग्रहण करेंगे। निराकरणीय परिकल्पना है

H0: यादृच्छिक पद सामान्य वितरण का अनुसरण करता है।

हम JB आँकड़ों से इस प्रकार निष्कर्ष निकालते हैं:

a) यदि JB सांख्यिकी का परिकलित मान स्वतंत्रता के 2 डिग्री के लिए χ^2 के सारणीबद्ध मान से अधिक है, तो हम निराकरणीय परिकल्पना को अस्वीकार करते हैं। हम अनुमान लगाते हैं कि यादृच्छिक पद सामान्य रूप से वितरित नहीं होता है

b) यदि JB आंकड़े का परिकलित मान स्वतंत्रता के 2 डिग्री के लिए χ^2 के सारणीबद्ध मान से कम है, तो हम निराकरणीय परिकल्पना को अस्वीकार नहीं करते हैं। हम अनुमान लगाते हैं कि यादृच्छिक पद सामान्य रूप से वितरित किया जाता है।

बोध प्रश्न 1

1) बहु प्रतिगमन प्रतिमान की मान्यताओं की सूची बनाएं।

....

....

....

....

2) सामान्यता के लिए जार्क-बेरा परीक्षण बताएं।

....

....

....

....

8.3 एकल मापदंड का परीक्षण

जनसंख्या प्रतिगमन कार्य हमें ज्ञात नहीं है। हम नमूना आँकड़ा के आधार पर मापदंडों का अनुमान लगाते हैं। चूँकि हम त्रुटि प्रसरण σ^2 को नहीं जानते हैं, इसलिए हमें z-परीक्षण (सामान्य वितरण के आधार पर) के बजाय t-परीक्षण उपयोग करना चाहिए।

आइए समीकरण (8.1) पर दी गई समष्टि प्रतिगमन रेखा पर विचार करें।

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

साधारण न्यूनतम वर्ग (OLS) विधि द्वारा अनुमानित नमूना प्रतिगमन रेखा है

$$\hat{Y}_i = b_1 + b_2 X_{2i} + b_3 X_{3i} \qquad \ldots (8.4)$$

जहां b_1, b_2 और b_3 क्रमश. β_1, β_2 और β_3 के अनुमानक हैं। त्रुटि पद का अनुमानक σ^2 द्वारा दिया गया है $\hat{\sigma}^2 = \frac{RSS}{n-k}$।

परिकल्पना परीक्षण के दो दृष्टिकोण हैं: (i) महत्व दृष्टिकोण का परीक्षण, और (ii) विश्वास अंतराल दृष्टिकोण। हम नीचे दोनों दृष्टिकोणों पर चर्चा करते हैं |

8.3.1 महत्व दृष्टिकोण का परीक्षण

इस दृष्टिकोण में हम निम्नानुसार आगे बढ़ते हैं:

(i) उस मापदंड का बिंदु अनुमान लें जिसे हम परीक्षण करना चाहते हैं, अर्थात, b_1, या b_2 or b_3 |

(ii) निराकरणीय परिकल्पना करें। मान लीजिए कि हम उम्मीद करते हैं कि पद X_2 का Y पर कोई प्रभाव नहीं है। इसका तात्पर्य है कि β_2 का मान निराकरणीय होना चाहिए। इस प्रकार, निराकरणीय परिकल्पना $H_0: \beta_2 = 0$ है। इस मामले में वैकल्पिक परिकल्पना क्या होनी चाहिए? वैकल्पिक परिकल्पना है $H_A: \beta_2 \neq 0$.

(iii) अगर $\beta_2 \neq 0$, फिर β_2 या तो सकारात्मक या नकारात्मक हो सकता है। इस प्रकार हमें टू-टेल परीक्षण का उपयोग करना चाहिए। तदनुसार, t-अनुपात का महत्वपूर्ण मूल्य तय किया जाना है |

(iv) आइए एक और परिदृश्य पर विचार करें। मान लीजिए हम उम्मीद करते हैं कि β_3 सकारात्मक होना चाहिए। इसका तात्पर्य है कि हमारी निराकरणीय परिकल्पना $H_0: \beta_3 > 0$ है। वैकल्पिक परिकल्पना है $H_A: \beta_3 \leq 0$ |

(v) यदि $\beta_3 > 0$, तो β_3 या तो निराकरणीय या ऋणात्मक हो सकता है। इस प्रकार महत्वपूर्ण क्षेत्र या अस्वीकृति क्षेत्र t-प्रायिकता वक्र के एक तरफ स्थित है। इसलिए, हमें वन-टेल परीक्षण उपयोग करना होगा। तदनुसार टी-अनुपात का महत्वपूर्ण मूल्य तय किया जाना है |

(vi) याद रखें कि निराकरणीय परिकल्पना आर्थिक सिद्धांत या तर्क पर निर्भर करती है। इसलिए, आपको कुछ तर्क के अनुसार निराकरणीय परिकल्पना निर्धारित करनी होगी। यदि आप उम्मीद करते हैं कि व्याख्यात्मक पद का आश्रित पद पर कोई प्रभाव नहीं होना चाहिए, तो निराकरणीय परिकल्पना में मापदंड को निराकरणीय के रूप में निर्धारित करें.

(vii) महत्व के स्तर पर निर्णय लें। यह उस त्रुटि की सीमा का प्रतिनिधित्व करता है जिसे आप होने देना चाहते हैं। यदि महत्व का स्तर 5 प्रतिशत (α = ०.०५) है, तो निराकरणीय परिकल्पना पर आपका निर्णय ५ प्रतिशत बार गलत होगा। यदि आप महत्व का 1 प्रतिशत स्तर ($\alpha = 0.01$) लेते हैं, तो निराकरणीय परिकल्पना पर आपका निर्णय 1 प्रतिशत बार गलत होगा (अर्थात, यह 99 प्रतिशत बार सही होगा) |

(viii) t-अनुपात की गणना करें। यहाँ मानक त्रुटि अनुमानक के प्रसरण का धनात्मक वर्गमूल है। बहु प्रतीपगमन प्रतिमान में OLS अनुमानकों के प्रसरण का सूत्र इकाई 7 में दिया गया है.

$$t = \frac{b_2 - \beta_2}{se(b_2)} \qquad \ldots (8.5)$$

(ix) t-अनुपात के परिकलित मान की t-अनुपात के सारणीबद्ध मान से तुलना करें। टी-टेबल पढ़ते समय दो मुद्दों के बारे में सावधान रहें: (i) महत्व का स्तर, और (ii) स्वतंत्रता की डिग्री। महत्व के स्तर का हमने ऊपर उल्लेख किया है। स्वतंत्रता की कोटि (n–k) है, जैसा कि आप पिछली इकाई से जानते हैं.

(x) यदि t-अनुपात का परिकलित मान t-अनुपात के सारणीबद्ध मान से अधिक है, तो निराकरणीय परिकल्पना को अस्वीकार करें।

यदि t-अनुपात का परिकलित मान t-अनुपात के सारणीबद्ध मान से कम है, तो निराकरणीय परिकल्पना को अस्वीकार न करें और वैकल्पिक निराकरणीय परिकल्पना को स्वीकार करें।.

8.3.2 विश्वास अंतराल दृष्टिकोण

हमने इकाई 3 और इकाई 5 में अंतराल अनुमान के बारे में चर्चा की है। इस प्रकार, यहां हम केवल आवश्यक बिंदुओं को ही शामिल कर रहे हैं।

(i) याद रखें कि विश्वास अंतराल (CI) प्रत्येक मापदंड के लिए व्यक्तिगत रूप से बनाया गया है। मापदंडों के समूह के लिए एक विश्वास अंतराल नहीं हो सकता।

(ii) महत्व के परीक्षण में ऊपर वर्णित तर्क के आधार पर विश्वास अंतराल का निर्माण होता है |

(iii) मान लीजिए कि हमारे पास निराकरणीय परिकल्पना है $H_0: \beta_2 = 0$ और वैकल्पिक परिकल्पना $H_A: \beta_2 \neq 0$ है। β_2 का अनुमानक b_2 है। हम b_2 की मानक त्रुटि जानते हैं |

(iv) यहाँ भी हम महत्व के स्तर (α) पर निर्णय लेते हैं। हम t-टेबल का संदर्भ लेते हैं और वांछित स्तर के महत्व के लिए t-अनुपात का पता लगाते हैं।

(v) स्वतंत्रता की डिग्री हमें ज्ञात है, अर्थात, (n-k)।

(vi) चूंकि उपरोक्त दो-पुच्छीय परीक्षण का मामला है, हम t प्रायिकता वक्र के प्रत्येक पक्ष पर $\alpha/2$ लेते हैं। इसलिए, हम प्रायिकता $\alpha/2$ और उपयोग स्वतंत्रता की डिग्री के अनुरूप टी-अनुपात लेते हैं।

(vii) याद रखें कि विश्वास अंतराल अनुमानक और उसकी मानक त्रुटि की मदद से बनाया जाता है। हम परीक्षण करते हैं कि मापदंड विश्वास अंतराल के भीतर है या नहीं।

(viii) विश्वास अंतराल की रचना इस प्रकार करें:

$$\left[b_2 - t_{\alpha/2} SE(b_2) \leq \beta_2 \leq b_2 + t_{\alpha/2} SE(b_2)\right]. \qquad \dots (8.6)$$

(ix) विश्वास अंतराल में मापदंड के बचे रहने की प्रायिकता (1-α) है। यदि हमने विश्वास अंतराल को 5 प्रतिशत के रूप में लिया है, तो β_2 विश्वास अंतराल में रहने की संभावना 95 प्रतिशत है

$$P_r\left[b_2 - t_{\alpha/2} SE(b_2) \leq \beta_2 \leq b_2 + t_{\alpha/2} SE(b_2)\right] = (1 - \alpha) \qquad \dots (8.7)$$

(x) यदि मापदंड (इस मामले में, β_2) विश्वास अंतराल में रहता है, तो निराकरणीय परिकल्पना को अस्वीकार न करें ।

(xi) यदि मापदंड विश्वास अंतराल के भीतर नहीं रहता है, तो निराकरणीय परिकल्पना को अस्वीकार करें और वैकल्पिक निराकरणीय परिकल्पना को स्वीकार करें.

बोध प्रश्न 2

1) उन चरणों का वर्णन करें जिनका आप इस परिकल्पना के परीक्षण में अनुसरण करेंगे कि $\beta_2 < 0$.

....

....

....

....

....

2) आंशिक ढलान गुणांक के जनसंख्या मापदंड के लिए एक विश्वास अंतराल बनाएं |

....

....

....

....

....

8.4 समग्र महत्व की परीक्षण

बहु प्रतिगमन प्रतिमान के महत्व का समग्र परीक्षण F-परीक्षण उपयोग करके किया जाता है। हमने इस पाठ्यक्रम की इकाई 5 में दो परिवर्ती प्रतिमान के संदर्भ में F-परीक्षण के बारे में चर्चा की है। बहु प्रतिगमन प्रतिमान के समग्र महत्व के परीक्षण के लिए हम निम्नानुसार आगे बढ़ते हैं:

(i) निराकरणीय परिकल्पना निर्धारित करें। एकाधिक प्रतिगमन प्रतिमान के समग्र महत्व के परीक्षण के लिए निराकरणीय परिकल्पना निम्नानुसार दी गई है:

$$H_0: \beta_2 = \beta_3 = \ldots \beta_k = 0 \qquad \ldots (8.8)$$

(ii) संगत वैकल्पिक परिकल्पना निर्धारित करें

$$H_A: \beta_2 = \ldots = \beta_k \neq 0 \qquad \ldots (8.9)$$

(iii) महत्व के स्तर पर निर्णय लें। इसका वही अर्थ है जो ऊपर वर्णित t-परीक्षण के मामले में है.

(iv) बहु प्रतिगमन प्रतिमान के लिए F-आँकड़ा किसके द्वारा दिया जाता है

$$F = \frac{ESS/(k-1)}{RSS(n-k)} \qquad \ldots (8.10)$$

(v) स्वतंत्रता की मात्रा ज्ञात कीजिए। समीकरण (8.10) में वर्णित F-सांख्यिकी स्वतंत्रता की डिग्री (k-1, n-k) के साथ एफ वितरण का अनुसरण करती है।.

(vi) समीकरण (8.10) के आधार पर F का परिकलित मान ज्ञात कीजिए। इसकी तुलना F के सारणीबद्ध मान (पुस्तक के अंत में दिए गए) से करें। वांछित स्तर के महत्व और स्वतंत्रता की उपयोग डिग्री के लिए सारणीबद्ध एफ मान पढ़ें।

(vii) यदि F का परिकलित मान सारणीबद्ध मान से अधिक है, तो निराकरणीय परिकल्पना को अस्वीकार कर दें।

(viii) यदि परिकलित मान सारणीबद्ध मान से कम है तो निराकरणीय परिकल्पना को अस्वीकार न करें।

8.5 दो मापदंडों के बीच समानता का परीक्षण

हम एक बहु प्रतिगमन प्रतिमान के मापदंडों या प्राचलों के बीच तुलना कर सकते हैं। विशेष रूप से, हम परीक्षण कर सकते हैं कि क्या प्रतिगमन प्रतिमान में दो मापदंड बराबर हैं। इस उद्देश्य के लिए हम उसी प्रक्रिया को उपयोग करते हैं जैसा हमने BECC 107 पाठ्यक्रम में सीखा है.

आइए हम निम्नलिखित प्रतिगमन प्रतिमान लें:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i \qquad \ldots (8.11)$$

याद रखें कि हम मापदंडों के प्रसरण को नहीं जानते हैं। इस प्रकार, मापदंडों की तुलना के लिए हम t-परीक्षण उपयोग करते हैं। दूसरे, हम मापदंड नहीं करते हैं। इसलिए, हम तुलनात्मक उद्देश्यों के लिए उनके OLS अनुमानक लेते हैं

हमारी निराकरणीय परिकल्पना और वैकल्पिक परिकल्पना इस प्रकार है:

$$H_0: \beta_3 = \beta_4 \quad \text{or} \quad (\beta_3 - \beta_4) = 0 \qquad \ldots (8.12)$$

$$H_1: \beta_3 \neq \beta_4 \quad \text{or} \quad (\beta_3 - \beta_4) \neq 0 \qquad \ldots (8.13)$$

उपरोक्त परिकल्पना के परीक्षण के लिए t-सांख्यिकी निम्नानुसार दी गई है:

$$t = \frac{(b_3 - b_4) - (\beta_3 - \beta_4)}{SE(b_3 - b_4)} \qquad \ldots (8.14)$$

उपरोक्त (n - k) स्वतंत्रता की डिग्री के साथ t-वितरण का अनुसरण करता है|

चूंकि $\beta_3 = \beta_4$ निराकरणीय परिकल्पना के तहत, हम समीकरण (8.14) को निम्नानुसार पुन: व्यवस्थित कर सकते हैं:

$$t = \frac{b_3 - b_4}{\sqrt{V(b_3) + V(b_4) - 2\text{cov}(b_3, b_4)}} \qquad \ldots (8.15)$$

t-सांख्यिकी का परिकलित मान समीकरण (8.15) द्वारा प्राप्त किया जाता है। हम t-अनुपात के परिकलित मान की t-अनुपात के सारणीबद्ध मान से तुलना करते हैं। हम वांछित स्तर के महत्व और स्वतंत्रता की उपयोग डिग्री के लिए टी-टेबल पढ़ते हैं।

यदि t-अनुपात का परिकलित मान सारणीबद्ध मान से अधिक है, तो हम निराकरणीय परिकल्पना को अस्वीकार करते हैं। यदि t-अनुपात का परिकलित मान सारणीबद्ध मान से कम है, तो हम निराकरणीय परिकल्पना को अस्वीकार नहीं करते हैं और वैकल्पिक परिकल्पना को स्वीकार करते हैं।

हमें अपने परिणामों की व्याख्या करने की आवश्यकता है। यदि हम निराकरणीय परिकल्पना को अस्वीकार करते हैं तो हम निष्कर्ष निकालते हैं कि आंशिक ढलान गुणांक β_3 और β_4 सांख्यिकीय रूप से काफी भिन्न हैं। यदि हम निराकरणीय परिकल्पना को अस्वीकार नहीं करते हैं, तो हम यह निष्कर्ष निकालते हैं कि ढलान गुणांक के बीच कोई सांख्यिकीय महत्वपूर्ण अंतर नहीं है β_3 और β_4|

बोध प्रश्न 3

1) बहु प्रतिगमन प्रतिमान के समग्र महत्व का परीक्षण करने के चरणों का उल्लेख करें|

....

....

....

....

2) बताएं कि दो प्राचलों के बीच समानता का परीक्षण कैसे किया जा सकता है|

....

....

....

....

8.6 मापदंडों पर रैखिक प्रतिबंधों का परीक्षण

कई बार हमारे सामने ऐसी स्थितियां आती हैं जहां हमें मापदंडों या प्राचलों पर रैखिक प्रतिबंधों के लिए परीक्षण करना पड़ता है। उदाहरण के लिए, आइए हम कोब-डगलस उत्पादन फलन पर विचार करें.

$$Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} e^{u_i} \qquad \ldots (8.16)$$

जहां Y_i उत्पादन है, X_{2i} पूंजी है और X_{3i} श्रम है। मापदंड β_2 और β_3 हैं। प्रसंभाव्य त्रुटि पद u_i है। अधोलिखित 'i' ith अवलोकन को इंगित करता है। कोब-डगलस उत्पादन फलन पैमाने पर स्थिर प्रतिफल प्रदर्शित करता है यदि मापदंड निम्नलिखित शर्त को पूरा करते हैं:

$$\beta_2 + \beta_3 = 1 \qquad \ldots (8.17)$$

जैसा कि हमने इकाई 6 में चर्चा की है, प्राकृतिक लघुगणक, कॉब-डगलस उत्पादन फलन को रैखिक रूप में व्यक्त किया जा सकता है:

$$ln\, Y_i = ln\, \beta_1 + \beta_2 lnX_{2i} + \beta_3 lnX_{3i} + u_i \qquad \ldots (8.18)$$

मान लीजिए कि फर्मों के नमूने पर आँकड़ा एकत्र किया है; हमारा नमूना आकार n है। उत्पादन फलन कोब-डगलस है जैसा कि ऊपर दिया गया है। हम परीक्षण करना चाहते हैं कि क्या उत्पादन फलन पैमाने पर स्थिर प्रतिफल प्रदर्शित करता है। इस प्रयोजन के लिए हमें F-परीक्षण उपयोग करने की आवश्यकता है। जैसा कि नीचे चर्चा की गई है, हम दो दृष्टिकोणों का पालन कर सकते हैं|

8.6.1 t-परीक्षण दृष्टिकोण

हम परिकल्पना के परीक्षण के लिए दो प्रक्रियाओं पर चर्चा करेंगे |

(a) इस मामले में हमारी निराकरणीय परिकल्पना और वैकल्पिक परिकल्पना इस प्रकार है:

$$H_0: \beta_2 + \beta_3 = 1 \qquad \ldots (8.19)$$

$$H_A: \beta_2 + \beta_3 \neq 1 \qquad \ldots (8.20)$$

उपरोक्त परिकल्पना के परीक्षण के लिए t-सांख्यिकी निम्नानुसार दी गई है:

$$t = \frac{(b_2+b_3)-(\beta_2+\beta_3)}{SE(b_2+b_3)} \qquad \ldots (8.21)$$

उपरोक्त (n - k) स्वतंत्रता की मात्रा के साथ t-वितरण का अनुसरण करता है |

हम समीकरण (8.21) को निम्न प्रकार से पुनर्व्यवस्थित कर सकते हैं:

$$\text{t} = \frac{b_2+b_3-1}{\sqrt{V(b_2)+V(b_3)+2\text{cov}(b_2,b_3)}} \qquad \ldots (8.22)$$

t-सांख्यिकी का परिकलित मान समीकरण (8.22) द्वारा प्राप्त किया जाता है। हम t-अनुपात के परिकलित मान की t-अनुपात के सारणीबद्ध मान से तुलना करते हैं। हम वांछित स्तर के महत्व और स्वतंत्रता की उपयोग डिग्री के लिए टी-टेबल पढ़ते हैं।

यदि t-अनुपात का परिकलित मान सारणीबद्ध मान से अधिक है, तो हम निराकरणीय परिकल्पना को अस्वीकार करते हैं। यदि t-अनुपात का परिकलित मान सारणीबद्ध मान से कम है, तो हम निराकरणीय परिकल्पना को अस्वीकार नहीं करते हैं और वैकल्पिक परिकल्पना को स्वीकार करते हैं।

हमें अपने परिणामों की व्याख्या करने की आवश्यकता है। यदि हम निराकरणीय परिकल्पना को अस्वीकार करते हैं तो हम यह निष्कर्ष निकालते हैं कि फर्म पैमाने पर निरंतर प्रतिफल

प्रदर्शित नहीं करती हैं। यदि हम निराकरणीय परिकल्पना को अस्वीकार नहीं करते हैं, तो हम यह निष्कर्ष निकालते हैं कि फर्म पैमाने पर निरंतर रिटर्न प्रदर्शित करती हैं.

(b) आइए हम फिर से दी गई निराकरणीय परिकल्पना को देखें (8.19).

$H_0: \beta_2 + \beta_3 = 1$

यदि उपरोक्त प्रतिबंध उपयोग होता है, तो हमारे पास होना चाहिए

$\beta_2 = (1 - \beta_3)$

आइए हम उपरोक्त संबंध को कोब-डगलस उत्पादन फलन में प्रतिस्थापित करें

$$ln\,Y_i = ln\,\beta_1 + (1 - \beta_3) lnX_{2i} + \beta_3 lnX_{3i} + u_i \qquad \ldots (8.23)$$

हम निम्न को प्राप्त करने के लिए समीकरण (8.23) में पदों को फिर से व्यवस्थित कर सकते हैं

$$ln\,Y_i - lnX_{2i} = ln\,\beta_1 - \beta_3 lnX_{2i} + \beta_3 lnX_{3i} + u_i$$

या,

$$ln(Y_i/X_{2i}) = \beta_0 + \beta_3 ln(X_{3i}/X_{2i}) + u_i \qquad \ldots (8.24)$$

ध्यान दें कि उपरोक्त प्रतिगमन प्रतिमान में आश्रित पद उत्पादन-श्रम अनुपात है और व्याख्यात्मक पद पूंजी-श्रम अनुपात है। हम समीकरण (8.24) में दिए गए प्रतिगमन प्रतिमान का अनुमान लगा सकते हैं और β_3 के OLS अनुमानक का पता लगा सकते हैं। अगर $\beta_3 = 1$, तो कोब-डगलस उत्पादन पैमाने पर स्थिर प्रतिफल प्रदर्शित करेगा

इसलिए, हम निराकरणीय परिकल्पना और वैकल्पिक परिकल्पना को इस रूप में निर्धारित करते हैं

$H_0: \beta_3 = 1$ और $H_A: \beta_3 \neq 1$

हम उप-खंड 8.3.1 में उल्लिखित व्यक्तिगत मापदंडों के लिए टी-परीक्षण उपयोग करते हैं। यदि निराकरणीय परिकल्पना को खारिज कर दिया जाता है तो हम यह निष्कर्ष निकालते हैं कि फर्म पैमाने पर निरंतर रिटर्न प्रदर्शित नहीं करती हैं.

8.6.2 प्रतिबंधित न्यूनतम वर्ग

ऊपर वर्णित t-परीक्षण दृष्टिकोण सभी मामलों में उपयुक्त नहीं हो सकता है। ऐसी स्थितियां हो सकती हैं जहां हमारे पास परीक्षण के लिए दो से अधिक मापदंड हों। ऐसी परिस्थितियों में हम F-परीक्षण उपयोग करते हैं। इस दृष्टिकोण को प्रतिबंधित न्यूनतम वर्ग कहा जाता है।

आइए समीकरण में दिए गए बहु प्रतिगमन प्रतिमान पर विचार करें (8.11)|

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i$$

मान लीजिए हमें इस परिकल्पना का परीक्षण करना है कि X_3 और X_4 आश्रित पद Y को प्रभावित नहीं करते हैं। ऐसे मामले में, मापदंड β_3 और β_4 शून्य होना चाहिए |

याद रखें कि यदि हम प्रतिगमन प्रतिमान में व्याख्यात्मक पद की संख्या बढ़ाते हैं, तो R^2 वृद्धि होती है, याद रखे की $R^2 = \frac{ESS}{TSS}$ | इस प्रकार, यदि समीकरण (8.11) में दो व्याख्यात्मक पद हटा दिए जाते हैं (अर्थात, उनके गुणांक शून्य हैं), तो R^2 के मान में कमी होगी | यदि चर X_3 और X_4 प्रासंगिक हैं, तो R^2 मूल्य में एक महत्वपूर्ण गिरावट होगी दूसरी ओर, यदि चर X_3 और X_4 प्रतिगमन प्रतिमान के लिए प्रासंगिक नहीं हैं, तो R^2 के मूल्य में गिरावट महत्वहीन होगा | हम प्रतिगमन प्रतिमान की इस गुण का उपयोग मापदंडों के एक समूह पर परिकल्पना का परीक्षण

करने के लिए करते हैं। इसलिए, प्रतिबंधित न्यूनतम वर्गों में F-परीक्षण उपयोग करते समय हमने दो बार प्रतिगमन प्रतिमान का अनुमान लगाया: (i) अप्रतिबंधित प्रतिमान, और (ii) प्रतिबंधित प्रतिमान |

हम निम्नानुसार आगे बढ़ते हैं:

(i) मान लीजिए कि प्रतिगमन प्रतिमान में k व्याख्यात्मक पद हैं |

(ii) इन k व्याख्यात्मक चरों में से, मान लीजिए कि पहले m व्याख्यात्मक पद प्रासंगिक नहीं हैं

(iii) इस प्रकार हमारी निराकरणीय परिकल्पना इस प्रकार होगी:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_m = 0 \qquad \ldots (8.25)$$

(iv) संगत वैकल्पिक परिकल्पना यह होगी कि βs शून्य नहीं हैं।

(v) (8.11) पर दिए गए अप्रतिबंधित प्रतिगमन प्रतिमान का अनुमान लगाएं। अनुमानित प्रतिगमन समीकरण के आधार पर वर्गों का अवशिष्ट योग (RRS) प्राप्त करें। इसे RSS_{UR} के रूप निरूपित करें |

(vi) व्याख्यात्मक चरों को छोड़कर जिसके लिए मापदंड शून्य हैं, प्रतिबंधित प्रतिगमन प्रतिमान का अनुमान लगाएं। इस प्रतिबंधित प्रतिमान से अवशिष्ट योग (RSS) प्राप्त करें। इसे RSS_R के रूप निरूपित करें |

(vii) हमारा F- सांख्यकीय निम्नलिखित है

$$F = \frac{RSS_R - RSS_{UR}/m}{RS\ _{UR}/(n-k)} \qquad \ldots (8.26)$$

(8.26) पर F-सांख्यकीय स्वतंत्रता की मात्रा के साथ (*m*, *n–k*) F-वितरण का अनुसरण करता है |

(ix) समीकरण (8.10) के आधार पर F का परिकलित मान ज्ञात कीजिए। इसकी तुलना F के सारणीबद्ध मान (पुस्तक के अंत में दिए गए) से करें। वांछित स्तर के महत्व और स्वतंत्रता की उपयोग मात्रा के लिए सारणीबद्ध F-मान पढ़ें |

(x) यदि F का परिकलित मान सारणीबद्ध मान से अधिक है, तो निराकरणीय परिकल्पना को अस्वीकार करें.

(xi) यदि परिकलित मान सारणीबद्ध मान से कम है तो निराकरणीय परिकल्पना को अस्वीकार न करें.

जैसा कि पहले उल्लेख किया गया है, वर्गों का अवशिष्ट योग (RSS) और निर्धारण के गुणांक (R^2) संबंधित हैं। इसलिए, R^2 के आधार पर भी F-परीक्षण करना संभव है। यदि हमारे पास अप्रतिबंधित प्रतिमान (R^2_{UR}) के लिए निर्धारण का गुणांक और प्रतिबंधित प्रतिमान (R^2_R) के लिए निर्धारण का गुणांक है, तो हम मापदंडों के निर्धारित के बारे में संयुक्त परिकल्पना का परीक्षण कर सकते हैं।

F-आँकड़ा होगा

$$F = \frac{R^2_{UR} - R^2_R/m}{(1-R^2_{UR})/(n-k)} \qquad \ldots (8.27)$$

जो स्वतंत्रता की डिग्री के साथ एफ-वितरण का अनुसरण करता है (*m*, *n–k*).

निकाले जाने वाले निष्कर्ष और परिणामों की व्याख्या वही होगी जो उपरोक्त बिंदुओं (x) और (xi) में वर्णित है।

8.7 एक प्रतिमान की संरचनात्मक स्थिरता: चाउ परीक्षण

कई बार हमारे सामने ऐसी स्थितियां आती हैं जहां आँकड़ों के पद्धिति में बदलाव होता है। आश्रित और स्वतंत्र पद पूरे नमूने में समान नहीं रह सकते हैं। उदाहरण के लिए, गरीब और अमीर परिवारों का बचत व्यवहार अलग हो सकता है। नीति परिवर्तन के बाद किसी उद्योग का उत्पादन भिन्न हो सकता है। ऐसी स्थितियों में संपूर्ण डेटानिर्धारित के लिए एकल प्रतिगमन चलाना उचित नहीं हो सकता है। अर्थमितीय प्रतिमान की संरचनात्मक स्थिरता की जांच करने की आवश्यकता है।

प्रतिगमन प्रतिमान में संरचनात्मक विराम लाने के लिए विभिन्न प्रक्रियाएं हैं। हम इकाई 9 में कृतृम परिवर्तनीय मामलों के बारे में चर्चा करेंगे। इस इकाई में हम एक बहुत ही सरल और विशिष्ट मामले पर चर्चा करेंगे |

मान लीजिए हमारे पास n प्रेक्षणों पर आँकड़ें है। हमें संदेह है कि पहले n_1 अवलोकन शेष n_2अवलोकन से भिन्न हैं (हमारे पास $n_1 + n_2 = n$ है)। इस मामले में निम्नलिखित तीन प्रतिगमन समीकरण उपयोग करें:

$Y_t = \lambda_1 + \lambda_2 X_t + u_t$ (अवलोकनों की संख्या: n_1) … (8.28)

$Y_t = r_1 + r_2 X_t + v_t$ (अवलोकनों की संख्या: n_2) … (8.29)

$Y_t = \alpha_1 + \alpha_2 X_t + w_t$ अवलोकनों की संख्या: $n = n_1 + n_2$) … (8.30)

यदि दोनों उप-नमूने समान हैं, तो हमारे पास होना चाहिए

$\lambda_1 = r_1 = \alpha_1$ और $\lambda_2 = r_2 = \alpha_2$.

यदि दोनों उप-नमूने भिन्न हैं तो नमूने में संरचनात्मक विराम होगा। इसका तात्पर्य है कि समीकरणों (8.28) और (8.29) के मापदंड अलग-अलग हैं। प्रतिगमन प्रतिमान की संरचनात्मक स्थिरता का परीक्षण करने के लिए हम चाउ (Chow) परीक्षण उपयोग करते हैं |

हम निम्नानुसार प्रक्रिया करते हैं:

(i) प्रतिगमन प्रतिमान (8.28) उपयोग करते। वर्गों का अवशिष्ट योग प्राप्त करें RSS_1.

(ii) प्रतिगमन प्रतिमान (8.29) चलाएँ। वर्गों का अवशिष्ट योग प्राप्त करें RSS_2.

(iii) प्रतिगमन प्रतिमान (8.30) चलाएं। वर्गों का अवशिष्ट योग प्राप्त करें RSS_3.

(iv) प्रतिगमन प्रतिमान (8.30) में हम प्रतिमान को दोनों उप-नमूनों में समान मापदंड रखने के लिए मजबूर कर रहे हैं। इसलिए, आइए हम इस प्रतिमान से प्राप्त वर्गों के अवशिष्ट योग को कहते हैं RSS_R.

(v) चूँकि (8.28) और (8.29) पर दिए गए प्रतिगमन प्रतिमान स्वतंत्र हैं, आइए इसे अप्रतिबंधित प्रतिमान कहते हैं। इसलिए, $RSS_{UR} = RSS_1 + RSS_2$

(vi) मान लीजिए कि दोनों उप-नमूने समान हैं। उस स्थिति में RSS_{UR} और RSS_R में कोई अंतर नहीं होना चाहिए। उस मामले में हमारी निराकरणीय परिकल्पना H_0 है: संरचनात्मक परिवर्तन नहीं है (या, मापदंड स्थिरता है)

(vii) निम्नलिखित परीक्षण आँकड़ों द्वारा उपरोक्त का परीक्षण करें:

$$F = \frac{RS\ \ _R - RSS_{UR})/K}{RS\ \ _{UR}/n_1 + n_2 - 2k} \qquad \ldots (8.31)$$

यह स्वतंत्रता की डिग्री के साथ एफ-वितरण का अनुसरण करता है $k, (n_1 + n_2 - 2k)$, जहां k प्रतिगमन प्रतिमान में व्याख्यात्मक पद की संख्या है

(viii) वांछित स्तर के महत्व और स्वतंत्रता की उपयोग डिग्री के लिए पुस्तक के अंत में दी गई एफ-वितरण तालिका देखें।

(ix) पदण (vii) पर प्राप्त एफ-सांख्यिकी के परिकलित मूल्य के आधार पर निष्कर्ष निकालें.

(x) यदि F का परिकलित मान सारणीबद्ध मान से अधिक है, तो निराकरणीय परिकल्पना को अस्वीकार करें.

(xi) यदि परिकलित मान सारणीबद्ध मान से कम है तो निराकरणीय परिकल्पना को अस्वीकार न करें.

चाउ परीक्षण हमें मापदंड या प्राचल स्थिरता के परीक्षण में मदद करता है। ध्यान दें कि चाउ परीक्षण की तीन सीमाएँ हैं:

(i) हम मानते हैं कि त्रुटि विपदण σ^2 पूरे नमूने में स्थिर है। उप-नमूनों के बीच त्रुटि विपदण में कोई अंतर नहीं है

(ii) संरचनात्मक विराम का बिंदु हमें ज्ञात नहीं है। हम संरचनात्मक परिवर्तन के उस बिंदु को मानते हैं।

(iii) एक से अधिक संरचनात्मक विराम होने पर हम चाउ परीक्षण उपयोग नहीं कर सकते हैं |

8.8 पूर्वानुमान

इकाई 5 में हमने बताया कि कैसे सरल प्रतिगमन प्रतिमान के आधार पर पूर्वानुमान या भविष्यवाणी की जाती है। हम एक ही प्रक्रिया को कई प्रतिगमन प्रतिमान तक बढ़ाते हैं। जैसा कि साधारण प्रतिगमन प्रतिमान के मामले में, कई प्रतिगमन प्रतिमान में दो प्रकार की पूर्वानुमान होती है।

यदि हम व्याख्यात्मक पद के विशेष मूल्यों के अनुरूप आश्रित पद के एक व्यक्तिगत मूल्य की पूर्वानुमान करते हैं, तो हम 'व्यक्तिगत पूर्वानुमान' प्राप्त करते हैं। जब हम व्याख्यात्मक पद के विशेष मूल्यों के अनुरूप Y के अपेक्षित मान की भविष्यवाणी करते हैं,

तो इसे 'माध्य पूर्वानुमान' कहा जाता है। दोनों ही मामलों में Y से अपेक्षित (व्यक्तिगत भविष्यवाणी और माध्य भविष्यवाणी) समान है। माध्य पूर्वानुमान और व्यक्तिगत पूर्वानुमान के बीच का अंतर उनके प्रसरण में निहित है |

8.8.1 माध्य पूर्वानुमान

मान लीजिये

$$X_0 = \begin{bmatrix} 1 \\ X_{02} \\ X_{03} \\ \vdots \\ X_{0k} \end{bmatrix} \qquad (8.32)$$

X पद के मूल्यों के संचालक बनें जिसके लिए हम $\hat{Y}_0$ की भविष्यवाणी करना चाहते हैं

अदिश रूप में अनुमानित बहु प्रतिगमन समीकरण है

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots \hat{\beta}_k X_{ki} + u_i \qquad \dots (8.33)$$

जिसे आव्यूह (मैट्रिक्स) अंकन में सघन रूप से लिखा जा सकता है

$$\hat{Y}_i = X_i'\hat{\beta} \quad \text{... (8.34)}$$

जहां

$$X_i' = [1 \ X_{2i} \ X_{3i} \ldots X_{ki}] \quad \text{... (8.35)}$$

और

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \quad \text{... (8.36)}$$

समीकरण (8.34) दिए गए X_i के अनुरूप Y_i का माध्य अनुमान है।

यदि X_i (8.35) में दिया गया है, तो (8.34) बन जाता है

$$(\hat{Y}_i | X_0') = X_0'\hat{\beta} \quad \text{... (8.37)}$$

जहां x_0 के मान स्थिर हैं। आपको ध्यान देना चाहिए कि (8.36) की निष्पक्ष भविष्यवाणी करता है $E(\hat{Y}_i|X_0')$, चूँकि $E(X_0'\hat{\beta}) = X_0'\hat{\beta}$.

माध्य पूर्वानुमान **का विपदण**

$(\hat{Y}_0|X_0')$ के प्रसरण का अनुमान लगाने का सूत्र इस प्रकार है:

$$\text{var}\,(\hat{Y}_0|X_0') = \sigma^2 X_0'(X'X)^{-1}X_0 \quad \text{... (8.38)}$$

जहाँ σ^2 का प्रसरण u_i है

X_0' वे X पद हैं जिनके लिए हम भविष्यवाणी करना चाहते हैं, और

चूँकि हम त्रुटि प्रसरण (σ^2) को नहीं जानते हैं, हम इसे इसके निष्पक्ष अनुमानक $\hat{\sigma}^2$ से बदल देते हैं

8.8.2 व्यक्तिगत पूर्वानुमान

जैसा कि पहले उल्लेख किया गया है, व्यक्तिगत भविष्यवाणी का अपेक्षित मूल्य व्यक्तिगत पूर्वानुमान के समान है, अर्थात, $\hat{Y}_i$.। व्यक्तिगत पूर्वानुमान का प्रसरण है

$$\text{var}(Y_0|X_0) = \sigma^2[1 + X'_0(X'X)^{-1}X_0] \quad \text{... (8.39)}$$

जहां $\text{var}(Y_0|X_0)$ के लिए है $E[Y_0 - \hat{Y}_0 \mid X]^2$. व्यवहार में हम इसके निष्पक्ष अनुमानक (estimator) द्वारा प्रतिस्थापित करते हैं |

व्यवहार में हम इसके निष्पक्ष अनुमानक $\hat{\sigma}^2$ द्वारा σ^2 प्रतिस्थापित करते हैं |

बोध प्रश्न 4

1) कोब-डगलस उत्पादन पर विचार करें। परिकल्पना के परीक्षण के चरणों को लिखिए कि यह पैमाने पर स्थिर प्रतिफल प्रदर्शित करता है।

....

....

....

....

....

2) चाउ परीक्षण करने के चरणों को लिखिए।

....

....

....

....

....

3) इंगित करें कि व्यक्तिगत भविष्यवाणी में औसत भविष्यवाणी की तुलना में अधिक भिन्नता क्यों है |

....

....

....

....

....

8.9 सार संक्षेप

इस इकाई ने बहु रेखीय प्रतिगमन प्रतिमान की मान्यताओं का वर्णन किया है जो जर्के-बेरा परीक्षण (जे-परीक्षण सामान्यता) द्वारा भी परीक्षण किए गए त्रुटि पद की सामान्यता को मजबूत करता है। व्यक्तिगत गुणांक के बारे में परिकल्पना का परीक्षण इकाई में समग्र महत्व परीक्षण से अलग है। इकाई दो प्रतिगमन गुणांकों की समानता के परीक्षण का भी वर्णन करती है। बाद में चाउ परीक्षण का उपयोग करके संरचनात्मक स्थिरता का परीक्षण किया जाता है। बहु प्रतिगमन का उपयोग स्वतंत्र चरों के दिए गए मानों के लिए आश्रित चरों की भविष्यवाणी के लिए भी किया जाता है। इकाई में व्यक्तिगत और संयुक्त परिकल्पना परीक्षण दोनों का वर्णन किया गया है।

8.10 आपकी प्रगति की जांच करने के लिए उत्तर/सुझाव

बोध प्रश्न 1

1) उप-भाग 8.2.1 देखें और उत्तर दें।

2) जार्के–बेरा परीक्षण आँकड़ा समीकरण (8.2) पर दिया गया है। वर्णन करें कि परीक्षण कैसे किया जाता है |

बोध प्रश्न 2

1) उप-भाग 8.3.1 का संदर्भ लें और उत्तर दें। निराकरणीय और वैकल्पिक परिकल्पनाओं पर निर्णय लें। उन चरणों का वर्णन करें जिनका आप अनुसरण करेंगे।

2) उप-भाग 8.3.2 देखें और उत्तर दें |

बोध प्रश्न 3

1) इसका परीक्षण F-परीक्षण द्वारा किया जा सकता है। विवरण के लिए भाग 8.4 देखें।

2) भाग 8.5 का संदर्भ लें और उत्तर दें |

बोध प्रश्न 4

1. हमने उप-भाग 8.6.1 में व्याख्या की है। इसका संदर्भ लें।

2. भाग 8.7 देखें और उत्तर दें।
3. भाग 8.8 देखें और उत्तर दें। इसका तर्क वही है जो इकाई 5 के भाग 5.7 में चर्चा किए गए द्वि–चरीय समाश्रयण प्रतिमान के मामले में है।

इकाई 9 समाश्रयण मॉडलों का विस्तार : प्रतिरूप चर प्रकरण*

इकाई की रूपरेखा

9.0 उद्देश्य

प्रस्तुत इकाई का अध्ययन करने के बाद आप इस योग्य होंगे कि–

- किसी गुणात्मक अथवा प्रतिरूप चर को परिभाषित कर सकें;
- बहिर्जात चर के रूप में किसी एकल प्रतिरूप वाले एनोवा (ANOVA) मॉडल पर चर्चा कर सकें;
- एक मात्रात्मक और एक प्रतिरूप चर वाले किसी भी एनकोवा (ANCOVA) मॉडल का विस्तृत विवरण दे सकें;
- प्रतिरूप चर समाश्रयण मॉडल के परिणामों की व्याख्या कर सकें;
- 'अवकल अवरोधन गुणांक' और 'अवकल समाश्रयण गुणांक' के बीच अंतर स्पष्ट कर सकें;
- 'अवकल समाश्रयण प्रतिरूपों' पर विचार करते समय सामने आने वाले 'संपाती, समानांतर, समवर्ती और असदृश' समाश्रयण मॉडलों से जुड़ी अवधारणाओं का वर्णन कर सकें; तथा
- स्पष्ट कर सकें कि किस प्रकार दो से अधिक प्रतिरूपों और अन्योन्य प्रतिरूपों को किसी समाश्रयण मॉडल में तैयार किया जा सकता है।

9.1 विषय–प्रवेश

हमारी वास्तविक जीवन–स्थितियों में अनेक चरों का योगदान होता है, जिनमें कुछ चर गुणात्मक होते हैं। इनके उदाहरण हैं– लिंग, विकल्प, राष्ट्रीयता आदि। ऐसे चर द्विभाजित अथवा द्विआधारी हो सकते हैं, यथा वे दो तक सीमित प्रतिक्रियाएँ हो सकते हैं, जैसे कि 'हाँ' या 'नहीं' की स्थिति; या फिर उनके पास दो से अधिक स्पष्ट प्रतिक्रियाएँ हो सकती हैं। हमें ऐसे चरों को समाश्रयण मॉडल में शामिल करने के तरीकों की आवश्यकता पड़ती है। इस इकाई में हम ऐसे ही कुछ उदाहरणों पर विचार करेंगे। बहरहाल, इस इकाई को

* डॉ. पूजा शर्मा, सहायक प्राध्यापक, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय एवं प्रोफेसर बी. एस. प्रकाश, इंदिरा गाँधी राष्ट्रीय मुक्त विश्वविद्यालय।

हम उन समाश्रयणों पर विचार करने तक ही सीमित रखेंगे जिनमें आश्रित चर की मात्रा निर्धारित कर दी जाती है। यह ध्यान देने की बात है कि जब कोई आश्रित चर स्वयं एक प्रतिरूप चर हो तो हमें उसके साथ प्रॉबिट (Probit) अथवा लॉजिट (Logit) जैसे मॉडलों की सहायता से निपटना होता है। ऐसे मॉडलों में आकलन की साधारण न्यूनतम वर्ग (OLS) पद्धति लागू नहीं होती है। बहरहाल, इस इकाई में हम ऐसे उदाहरणों पर विचार नहीं करेंगे। इनके विषय में हम पाठ्यक्रम 'BECE 142: व्यावहारिक अर्थमिति' में पढ़ेंगे।

इस इकाई में हम केवल ऐसे उदाहरणों पर विचार करेंगे जिनमें स्वतंत्र चर एक प्रतिरूप (dummy) चर होता है। गुणात्मक चर चूँकि सीधे परिमाणित नहीं होते हैं, हम उन्हें प्रतिरूप चर के रूप में मानकर परिमाणित (या सुस्पष्ट) बना सकते हैं। उदाहरण के लिए, पुरुष या महिला, नियोजित या बेरोजगार, आदि जैसे चरों पर विचार करें। ये इस अर्थ में परिमाण निर्धारित करने योग्य हैं कि उन्हें 'महिला' के रूप में 1 और 'पुरुष' के रूप में 0 माना जा सकता है। इस प्रकार के अन्य उदाहरण भी हो सकते हैं, यथा 1 यदि हाँ और 0 यदि नहीं; 1 यदि रोजगारप्राप्त हैं और 0 यदि बेरोजगार हैं, इत्यादि।

उपर्युक्त उदाहरण में हमने गुणात्मक प्रतिक्रिया को मात्रात्मक रूप में परिवर्तित कर दिया है। तदनुसार, गुणात्मक चर अब परिमाणित हो गया है। इस प्रकार के समाश्रयण कोई साधारण प्रतिक्रमण दर्शा सकते हैं, यथा यहाँ केवल एक स्वतंत्र चर है जो गुणात्मक है और इसे प्रतिरूप चर के रूप में लिया जाता है। अन्यथा यहाँ दो स्वतंत्र चर हो सकते हैं, जिनमें से एक को प्रतिरूप माना जा सकता है और दूसरा इसका सहपरिवर्त्य है, यथा इसका प्रतिरूप के रूप में माने जाने वाले चर के साथ घनिष्ठ संबंध है। उदाहरण के लिए, व्यक्तियों की कर–पूर्व आय को किसी सीमा स्तर से ऊपर वर्गीकृत किया जा सकता है और फिर उसे प्रतिरूप चर के रूप में माना जा सकता है, यथा 1 अथवा 0 के रूप में ली गई प्रतिक्रिया के साथ सीमा–स्तर आय से ऊपर अथवा नीचे।

अब कर–पश्चात आय, जो कि कर–पूर्व आय का ही एक सहपरिवर्त्य है, पर उसके वास्तविक परिमाणित मान को लेकर विचार किया जा सकता है। इसी प्रकार का विस्तार ऐसी स्थितियों में भी हो सकता है जहाँ आपको अनेक प्रतिरूपों पर विचार करना पड़ता है और ऐसे उदाहरणों में भी जहाँ आपको अन्योन्य प्रतिरूपों पर विचार करना होता है। इस प्रकार के समाश्रयण की प्रकृति, विशेष रूप से उनकी अनुमिति अथवा व्याख्यात्मक सरोकार के लिए, ही इस इकाई में हमारा विचारणीय विषय रहेगा।

9.2 एकल प्रतिरूप प्रकरण : एनोवा (ANOVA) मॉडल

हम सबसे पहले केवल एक स्वतंत्र चर वाले एक साधारण समाश्रयण मॉडल पर विचार करते हैं। इसके अलावा, यह स्वतंत्र चर एक प्रतिरूप चर भी है, यथा –

$$Y_i = \beta_1 + \beta_2 D_i + u_i \qquad \dots (9.1)$$

यहाँ, हम Y को वार्षिक खाद्य व्यय के रूप में लेते हैं और D_i को लिंग पहचान के ऐसे रूप में जहाँ पुरुष होने की स्थिति में वह 0 मान लेगा जबकि महिला होने की स्थिति में वह 1 मान लेगा। ये D_i अंकित चर, तदनुसार, नियत होते हैं और इसलिए अप्रसंभाव्य भी।

अब, यदि हम मान लें कि $u_i \sim N(0, \sigma^2)$ है तो समीकरण (9.1) में प्राचलों का आकलन करने के लिए साधारण न्यूनतम वर्ग (OLS) पद्धति लागू की जा सकती है। यदि हम ऐसा करते हैं तो पुरुषों और महिलाओं के लिए औसत खाद्य व्यय क्रमशः निम्नवत् दर्शाया जाएगा –

$$E(Yi \mid Di = 0) = \beta_1 + \beta_2(0) = \beta_1 \qquad \dots (9.2)$$

$$E(Y_i \mid D_i = 1) = \beta_1 + \beta_2 \qquad \dots (9.3)$$

यहाँ चर β_1 पुरुषों के खाद्य व्यय का औसत या माध्य दर्शाता है। यह वह श्रेणी है जिसके लिए प्रतिरूप चर को मान 0 दिया जाता है। समाश्रयण गुणांक β_2 दर्शाता है कि महिलाओं का औसत खाद्य व्यय पुरुषों के औसत खाद्य व्यय से कितना भिन्न है। तदनुसार, $\beta_1 + \beta_2$ महिलाओं का औसत खाद्य व्यय दर्शाता है। इसे देखते हुए चर β_2 को समाश्रयण गुणांक की संज्ञा दे देना ठीक नहीं होगा क्योंकि यहाँ कोई निरंतर समाश्रयण रेखा नहीं है। अतः चर β_2 'अवकल अवरोधन गुणांक' कहलाएगा। यह दर्शाएगा कि उक्त दो श्रेणियों के बीच 'अवरोधन' पद का मान कितना भिन्न है। अब एक प्रश्न यह उठता है कि यदि हमने '0' के समनुदेशन को पुरुषों और महिलाओं की उक्त दो श्रेणियों के बीच परस्पर बदल दिया होता तो क्या होता (यथा, यदि हमने महिलाओं को '0' मान दिया होता)।

यह ध्यान देने की बात है कि जब तक हमारे पास वर्तमान उदाहरण में केवल दो श्रेणियाँ होती हैं, यथा यह सरल समाश्रयण का ऐसा उदाहरण होता है जिसमें केवल एक स्वतंत्र चर को प्रतिरूप चर D_i के रूप में लिया जाता है और प्रतिक्रियाओं की श्रेणी द्विभाजित अथवा द्विआधारी होती है, मूलतः इससे कोई फर्क नहीं पड़ता कि किस श्रेणी को 1 मान मिलता है और किस श्रेणी को 0 मान मिलता है। हालाँकि, कुछ मामूली अंतर होता है। आइए, देखें कि यह अंतर क्या है। जिस श्रेणी को हम 0 मान देते हैं, उसे 'आधार श्रेणी' कहा जाता है। इसे वैकल्पिक नामों से भी पुकारा जाता है, जैसे 'संदर्भ' या 'तल चिह्न' अथवा 'तुलना श्रेणी'। इस प्रकार के समनुदेशन में अवरोधन मान उस श्रेणी के औसत मूल्य का प्रतिनिधित्व करता है जिसे मान 0 मिलता है (जो कि ऊपर हमारे उदाहरण में 'पुरुष' है)।

समीकरण (9.3) हमें जो बताता है वह इस प्रकार है – इस प्रकार के किसी भी समनुदेशन के आधार पर महिलाओं के खाद्य व्यय का औसत मान 'अवरोधन मान में समाश्रयण गुणांक' जोड़कर प्राप्त किया जाना है। यदि प्रतिरूप का समनुदेशन अन्य प्रकार से किया जाता है, यथा महिलाएँ 0 और पुरुष 1, तो हम अवरोधन पद के संख्यात्मक मान और उसके t–मान में बदलाव देखेंगे। इसे छोड़कर, R^2 मान, आकलित प्रतिरूप चर गुणांक का निरपेक्ष मान और इसकी मानक त्रुटि यथावत रहेंगे। आइए, बेहतर ढंग से समझने के लिए इसे एक उदाहरण की मदद से देखें।

तालिका 9.1 में दिए गए पुरुषों एवं महिलाओं के 'खाद्य व्यय' और आय संबंधी आँकड़ों पर विचार करें। ये आँकड़े विभिन्न आय वर्गों में लोगों की वास्तविक संख्या (जो कि हजारों में है) के आधार पर आकलित मध्य मान हैं। हम सबसे पहले तालिका 9.1 के आँकड़ों से तालिका 9.2 की रचना करते हैं, जैसा कि नीचे दिया गया है।

तालिका 9.1: लिंगानुसार आय और खाद्य व्यय पर आँकड़े

(आँकड़े $ में)

आयु	खाद्य व्यय (महिला)	आय (महिला)	खाद्य व्यय (पुरुष)	आय (पुरुष)
< 25	1983	11557	2230	11589
25-34	2987	29387	3757	33328
35-44	2993	31463	3821	36151
45-54	3156	29554	3291	35448
55-64	2706	25137	3429	32988
> 65	2217	14952	2533	20437

स्रोत : तालिका 6–1, अध्याय 6, गुजराती

तालिका 9.2: आय और लिंग के अनुसार खाद्य व्यय पर आँकड़े

(आँकड़े $ में)

प्रेक्षण	खाद्य व्यय ($)	आय ($)	लिंग
1	1983	11557	1
2	2987	29387	1
3	2993	31463	1
4	3156	29554	1
5	2706	25137	1
6	2217	14952	1
7	2230	11589	0
8	3757	33328	0
9	3821	36151	0
10	3291	35448	0
11	3429	32988	0
12	2533	20437	0

स्रोत : तालिका 6–2, अध्याय 6, गुजराती

लिंग प्रतिरूप चर (इस स्तर पर आय चर को ध्यान में रखे बिना) पर प्रतिक्रमित खाद्य व्यय के परिणाम निम्नलिखित आँकड़े प्रस्तुत करते हैं–

$$\hat{Y}_i = 3176.833 - 503.1667\, D_i$$
$$\text{मानक त्रुटि} = (233.0446) \qquad (329.5749)$$
$$t = (13.6318) \qquad (-1.5267) \qquad R^2 = 0.1890$$

उक्त परिणाम बताते हैं कि पुरुषों का औसत व्यय 3177 डॉलर और महिलाओं का औसत व्यय (3177 – 503 = 2674 डॉलर) है। आकलित D_i सांख्यिकीय रूप से महत्वपूर्ण नहीं है (क्योंकि इसका t–मान मात्र –1.53 है)। इसका अर्थ है कि महिलाओं और पुरुषों के बीच खाद्य व्यय में अंतर सांख्यिकीय रूप से महत्वपूर्ण नहीं है। याद करें कि हमने पुरुषों को '0' मान दिया है। इसलिए अवरोधन मान पुरुषों के लिए माध्य मान निरूपित करता है। इस समनुदेशन में महिलाओं के खाद्य व्यय का माध्य मान प्राप्त करने के लिए हम प्रतिरूप चर के गुणांक के मूल्य को अवरोधन मान में जोड़ते हैं। आइए, अब हम महिलाओं को '0' और पुरुषों को '1' के मान फिर से निर्दिष्ट करते हैं। इससे हमें जो समाश्रयण परिणाम मिलते हैं वे निम्नलिखित हैं–

$$\hat{Y}_i = 2673.667 + 503.1667\, D_i$$
$$\text{मानक त्रुटि} = (233.0446) \qquad (329.5749)$$
$$t = (11.4227) \qquad (-1.5267) \qquad R^2 = 0.1890$$

इस प्रकार, हम देखते हैं कि उक्त दोनों लिंगों का माध्य खाद्य उपभोग व्यय एक समान रहा है। यहाँ चर R^2 का मान भी वही है। प्रतिरूप चर गुणांक का निरपेक्ष मान और उसकी मानक त्रुटियाँ भी एक समान ही हैं। एकमात्र परिवर्तन अवरोधन पद के संख्यात्मक मान और उसके t–मान में ही देखा जाता है।

एक और प्रश्न जो हमारे सामने आ सकता है, वह यह है कि चूँकि हमारे पास दो श्रेणियाँ हैं – पुरुष और महिलाएँ – क्या हम उन्हें दो प्रतिरूप निर्दिष्ट कर सकते हैं? इसका अर्थ है कि हम इस मॉडल को निम्नवत् मानते हैं–

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_i + u_i \qquad \dots (9.4)$$

जहाँ Y खाद्य व्यय है, $D_2 = 1$ महिला के लिए व 0 पुरुष के लिए और $D_3 = 1$ पुरुष के लिए व 0 महिला के लिए। अनिवार्यतः हम यह देखने की कोशिश कर रहे हैं कि क्या हम पुरुषों और महिलाओं के लिए दो प्रतिरूप अलग–अलग निर्दिष्ट कर सकते हैं? इसका उत्तर 'नहीं' है। इसका कारण जानने के लिए दो महिलाओं और तीन पुरुषों के प्रतिदर्श हेतु आँकड़ों पर विचार करें, जिनके लिए आँकड़ा आव्यूह तालिका 9.3 में है। आप देखेंगे कि $D_2 = 1 - D_3$ अथवा $D_3 = 1 - D_2$ है। यह पूर्ण संरेखता की स्थिति है। अतः यदि गुणात्मक चर में दो श्रेणियाँ हों, जैसे कि यहाँ लिंग, तो हमें सदा केवल एक ही प्रतिरूप चर का उपयोग करना चाहिए।

तालिका 9.3: समीकरण के लिए आँकड़ा आव्यूह

लिंग	अवरोधन	D_2	D_3
पुरुष Y_1	1	0	1
पुरुष Y_2	1	0	1
महिला Y_3	1	1	0
पुरुष Y_4	1	0	1
महिला Y_5	1	1	0

एक अधिक सामान्य नियम यह है – यदि किसी मॉडल में सामान्य अवरोधन β_1 हो और गुणात्मक चर में m श्रेणियाँ हों तो हमें केवल (m – 1) प्रतिरूप चर ही प्रस्तुत करना चाहिए। यदि हम ऐसा नहीं करते हैं तो हम आकलन की समस्या में पड़ जाते हैं, जिसे 'प्रतिरूप चर जाल' कहा जाता है।

अंततः यह ध्यान देने की बात है कि जब हमारे पास केवल एक प्रतिरूप चर वाला कोई साधारण समाश्रयण मॉडल होता है, जैसा कि यहाँ माना जाता है, तो मॉडल को 'एनोवा (ANOVA) मॉडल' भी कहा जाता है। ऐसा इसलिए है कि ऐसा कोई दूसरा चर नहीं है जिससे हम आश्रित चर पर प्रभाव या परिवर्तनशीलता जानने का प्रयास कर रहे है। जब हमारे पास यह होता है तो हमें जो प्राप्त होता है उसे 'एनकोवा (ANCOVA) मॉडल' कहा जाता है। इस प्रकार के उदाहरण पर हम अगले पाठांश में चर्चा करेंगे।

9.3 सहप्रसरण का विश्लेषण (ANCOVA) मॉडल

आर्थिक विश्लेषण में व्याख्यात्मक चरों के बीच कुछ का गुणात्मक और कुछ अन्य का मात्रात्मक होना एक आम बात है। ऐसे मॉडलों को सहप्रसरण–का–विश्लेषण अर्थात् एनकोवा (ANCOVA) मॉडल कहा जाता है। यहाँ हम एक ऐसे मॉडल पर विचार करेंगे जिसमें समाश्रयियों के बीच एक मात्रात्मक और एक प्रतिरूप चर दोनों हों। आमतौर पर मात्रात्मक और गुणात्मक चरों के संयोजन वाले समाश्रयण मॉडल को ही 'एनकोवा मॉडल' कहा जाता है। यहाँ मात्रात्मक चरों को 'सहविचर' या 'नियंत्रण चर' कहा जाता है। एनकोवा मॉडल एनोवा (ANOVA) मॉडल का ही एक विस्तार है। ये किसी भी मॉडल में सहविचरों (यथा, मात्रात्मक व्याख्यात्मक चरों) के प्रभावों को सांख्यिकीय रूप से नियंत्रित करने की एक ऐसी विधि प्रदान करते हैं जिसमें एक प्रतिरूप चर के रूप में लिए जाने वाले गुणात्मक चर के साथ दोनों प्रकार के चर शामिल होते हैं।

मान लिया जाने वाला मात्रात्मक चर आमतौर पर इस अर्थ में एक सहविचर होता है कि वह मुख्य चर के साथ घनिष्ठ संबंध रखता है। इसी कारण किसी भी मॉडल से सहविचरों के बहिष्करण के परिणामस्वरूप मॉडल विनिर्देशन त्रुटि की संभावना होती है। ऊपर लिए गए उदाहरण में हमने केवल लिंग प्रतिरूप $[Y_i = \beta_1 + \beta_2 D_i + u_i]$ पर ही 'खाद्य व्यय' को प्रतिक्रमित किया है।

आइए, अब हम किसी व्याख्यात्मक चर (X_i) के रूप में एक अन्य चर 'कर–पश्चात आय', यथा प्रयोज्य आय (खाद्य व्यय का एक सहविचर), पर विचार करते हैं। हमारा मॉडल अब निम्नवत् है–

$$Y_i = \beta_1 + \beta_2 D + \beta_3 X_i + u_i \qquad \text{... (9.5)}$$

जहाँ Y = खाद्य व्यय ($), X = कर–पश्चात आय ($), D = 1 महिला के लिए और = 0 पुरुष के लिए। चलिए, अब बेहतर मूल्यांकन के लिए तालिका 9.2 में दिए गए आँकड़ों से प्राप्त समीकरण (9.5) में समाश्रयण के परिणाम पर विचार करते हैं, जो कि निम्नवत् है–

$$\hat{Y}_c = 1506.244 - 228.9868 D_i + 0.0589 X_i$$

$$t = (8.0115) \qquad (-2.1388) \qquad (9.6417)$$

$$R^2 = 0.9284$$

प्रतिरूप चर गुणांक सांख्यिकीय रूप से महत्वपूर्ण होता है। अतएव, हम इस शून्य–स्तरीय परिकल्पना को अस्वीकार करते हैं कि पुरुषों और महिलाओं के खाद्य व्यय के औसत मान में कोई अंतर नहीं होता। दूसरे शब्दों में, हम यह निष्कर्ष निकालते हैं कि लिंग का उपभोग अथवा खाद्य व्यय पर महत्वपूर्ण प्रभाव पड़ता है।

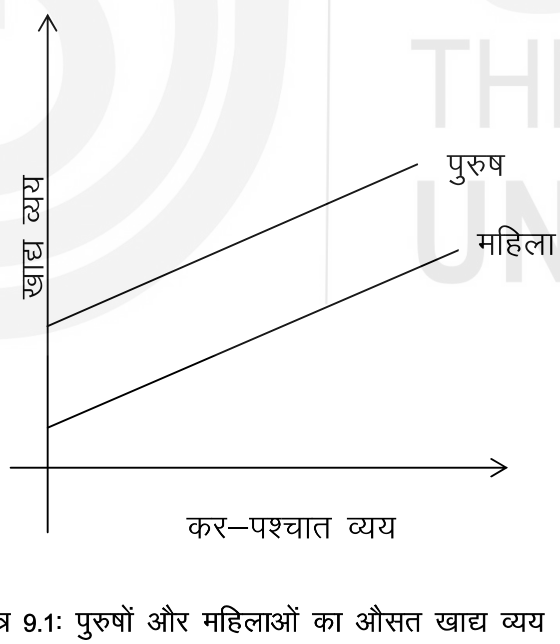
यह ध्यान देने की बात है कि उपभोग व्यय में उक्त अंतर कर–पश्चात आय के प्रभाव को नियत रखते हुए अनुमानित है। इसी प्रकार, लिंग अंतर को नियत रखते हुए, कर–पश्चात आय गुणांक भी महत्वपूर्ण होता है। 'कर–पश्चात आय' के लिए समाश्रयण गुणांक इंगित करता है कि प्रयोज्य आय में प्रत्येक अतिरिक्त डॉलर की वृद्धि के लिए औसत खाद्य व्यय [यथा, उपभोग हेतु सीमांत उपभोग प्रवृत्ति (MPC)] में 6 सेंट तक की वृद्धि होती है।

यह भी ध्यान देने की बात है कि चूँकि हमने पुरुषों के लिए '0' लिया है, अवरोधन पद पुरुष सीमांत उपभोग प्रवृत्ति से संबंध रखता है। महिला सीमांत उपभोग प्रवृत्ति के लिए हमें लिंग प्रतिरूप के गुणांक (यथा, 1506.2 − 228.9 = 1277.3) में अवरोधन मान को जोड़ना होगा। तदनुसार, महिलाओं और पुरुषों के सीमांत उपभोग प्रवृत्ति के समीकरणों को क्रमशः निम्नवत् लिखा जा सकता है–

महिला औसत खाद्य व्यय : $\hat{Y}_i = 1277.2574 + 0.0589X_i$

पुरुष औसत खाद्य व्यय : $\hat{Y}_i = 1506.2440 + 0.0589X_i$

चूँकि सीमांत उपभोग प्रवृत्ति अथवा प्रतिक्रमण दोनों लिंगों के लिए एक समान है, दोनों समाश्रयण समानांतर हैं, जैसा कि नीचे चित्र 9.1 में दर्शाया गया है।



चित्र 9.1: पुरुषों और महिलाओं का औसत खाद्य व्यय

उक्त मॉडल किसी आश्रित चर की व्याख्या करने में दोनों प्रकार के चरों (मात्रात्मक एवं गुणात्मक) की भूमिका और प्रभाव को दर्शाता है। विशेष रूप से, उपर्युक्त उदाहरण में कर–पश्चात व्यय पुरुषों और महिलाओं दोनों के खाद्य व्यय को प्रभावित करता है।

बोध प्रश्न 1

[प्रत्येक प्रश्न का उत्तर लगभग 50—100 शब्दों में दें।]

1) किसी गुणात्मक चर को आप कैसे परिभाषित करेंगे?

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

2) किसी एकल प्रतिरूप चर के साथ समाश्रयण मॉडल निर्दिष्ट करें। आकलित गुणांकों की व्याख्या की दृष्टि से उसकी विशेषताओं का उल्लेख भी करें।

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

3) यदि समीकरण (9.1) के अनुसार सरल समाश्रयण मॉडल में प्रतिरूप चर, जैसे लिंग, के लिए आधार मान को पुनर्निर्दिष्ट कर दिया जाता है तो क्या होता है?

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

4) 'प्रतिरूप चर जाल' से आप क्या समझते हैं? हम इससे कैसे बचें?

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

5) एनोवा (ANOVA) मॉडल और एनकोवा (ANCOVA) मॉडल के बीच अंतर स्पष्ट करें।

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

…………………………………………………………………….

6) एनकोवा (ANCOVA) मॉडल का क्या लाभ होता है? किसी एनोवा मॉडल में किसी सहविचर को शामिल करने से बचने का क्या परिणाम होता है?

………………………………………………………………………….…

………………………………………………………………………….…

………………………………………………………………………….…

………………………………………………………………………….…

7) एक गुणात्मक और एक मात्रात्मक चर के साथ किसी एनकोवा (ANCOVA) मॉडल के सामान्य रूप को निर्दिष्ट करें। उदाहरणार्थ लिए गए मात्रात्मक चर के लिए समाश्रयण गुणांक सामान्यतया क्या दर्शाता है?

………………………………………………………………………….…

………………………………………………………………………….…

………………………………………………………………………….…

………………………………………………………………………….…

9.4 दो समाश्रयण मॉडलों के बीच तुलना

ऊपर दिए गए उदाहरण में, यथा एनोवा और एनकोवा दोनों मॉडलों के लिए, हमने देखा कि समाश्रयण गुणांक एक समान थे परंतु उनके अवरोधन भिन्न–भिन्न थे। इससे यह प्रश्न उठता है कि क्या प्रतिक्रमण भी भिन्न–भिन्न हो सकते हैं? यदि हमारी रुचि समाश्रयण गुणांकों में अंतर के लिए भी परीक्षण करने की हो तो हम ऐसे मॉडल को कैसे तैयार करें? इसे समझने के लिए हम एक 'स्लोप ड्रिफ्टर' अर्थात् प्रतिगामी प्रक्षेपक प्रस्तुत करते हैं। ऊपर दिए गए पुरुष या महिला के उपभोग व्यय के उदाहरण के लिए, आइए, अब हम मॉडल को प्रतिरूपों के साथ निर्दिष्ट करके लिंगानुसार उपभोग व्यय में अंतर की तुलना करने के लिए आगे बढ़ते हैं–

$$Y_i = \beta_1 + \beta_2 D_i + \beta_3 X_i + \beta_4 (D_i X_i) + u_i \qquad \text{... (9.6)}$$

आप देखेंगे कि जोड़ा गया अतिरिक्त चर $D_i X_i$ है, जो कि गुणक अथवा अन्योन्य रूप में है। समीकरण (9.6) में हमने पुरुषों के लिए $D_i = 0$ और महिलाओं के लिए $D_i = 1$ लिया है। अब पुरुषों के लिए 'औसत खाद्य व्यय' निम्नवत् दर्शाया जाएगा–

$$E(Y_i \mid D_i = 0, X_i) = \beta_1 + \beta_3 X_i \qquad \text{… (9.7)}$$

{चूँकि $D_i = 0$ }

जबकि महिलाओं के लिए 'औसत खाद्य व्यय' निम्नवत् दर्शाया जाएगा –

$$E(Y_i \mid D_i = 1, X_i) = \beta_1 + \beta_2 D_i + (\beta_3 + \beta_4 D_i) X_i$$

$$= (\beta_1 + \beta_2) + (\beta_3 + \beta_4) X_i \qquad \text{… (9.8)}$$

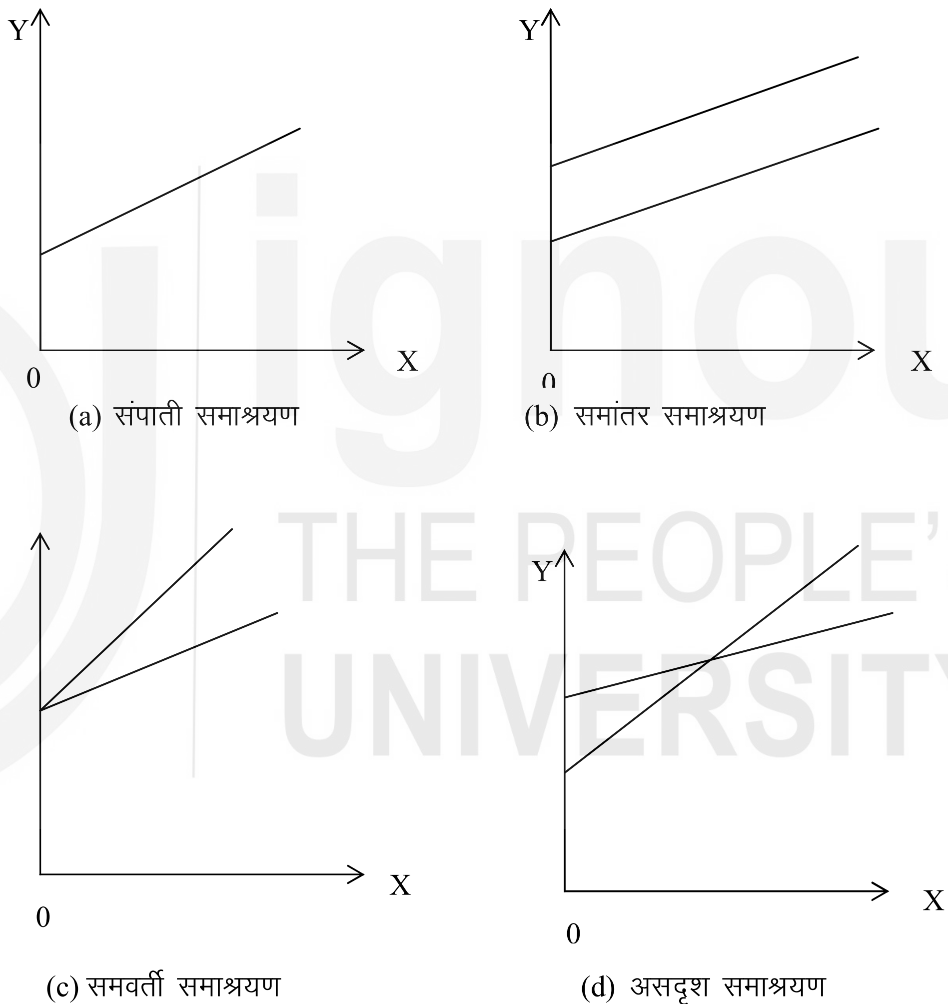
{चूँकि $D_i = 1$}

समीकरण (9.8) में $(\beta_1 + \beta_2)$ उस श्रेणी के लिए Y का माध्य मान देता है जो X के शून्य होने पर 1 का प्रतिरूप मान प्राप्त करता है। साथ ही, $(\beta_3 + \beta_4)$ उस श्रेणी के लिए आय चर का समाश्रयण गुणांक देता है जो 1 का प्रतिरूप मान प्राप्त करता है। आप देखेंगे कि

'योग t रूप' में प्रतिरूप चर का समावेशन हमें उक्त दो वर्गों के अवरोधन पदों के बीच अंतर करने में सक्षम बनाता है। इसी प्रकार, अन्योन्य (अथवा गुणक) रूप (यथा, D_iX_i) का समावेशन हमें उक्त दो वर्गों के समाश्रयण गुणांकों (अथवा पदों) के बीच अंतर करने में सक्षम बनाता है।

अवकल अवरोधन गुणांक, β_2, और अवकल समाश्रयण गुणांक, β_4, के सांख्यिकीय महत्व के आधार पर हम अनुमान लगा सकते हैं कि क्या महिला और पुरुष खाद्य व्यय फलन उनके अवरोधन मान अथवा उनके समाश्रयण मान, अथवा दोनों में भिन्न–भिन्न हैं।

नीचे दिए गए चित्र 9.2 के अनुसार यहाँ चार संभावनाएँ हो सकती हैं। चित्र 9.2 (a) से ज्ञात होता है कि उक्त दोनों खाद्य व्यय समाश्रयणों के अवरोधन अथवा समाश्रयण गुणांक में कोई अंतर नहीं है। ऐसे समाश्रयण समीकरणों को 'संपाती समाश्रयण' कहा जाता है।



चित्र 9.2: समाश्रयण समीकरणों की तुलना

चित्र 9.2 (b) से पता चलता है कि दो समाश्रयण गुणांक एक समान हैं, परंतु अवरोधन भिन्न हैं। इस प्रकार के समाश्रयण को 'समानांतर समाश्रयण' के रूप में जाना जाता है। चित्र 9.2 (c) से पता चलता है कि दो समाश्रयणों में एक ही अवरोधन है, परंतु प्रतिक्रमण भिन्न–भिन्न हैं। इस प्रकार के समाश्रयण को 'समवर्ती समाश्रयण' कहा जाताहै। चित्र 9.2

(d) से ज्ञात होता है कि दो अवरोधन और दो समाश्रयण गुणांक दोनों भिन्न–भिन्न हैं। इस प्रकार के समाश्रयण को 'असदृश समाश्रयण' कहा जाता है।

9.5 विविध प्रतिरूप और अन्योन्य प्रतिरूप

हमें प्रायः एक से अधिक प्रतिरूप चरों पर विचार करने की आवश्यकता पड़ती है। इसके अलावा ऐसे मामले भी हो सकते हैं जहाँ हमें प्रतिरूप चर अन्योन्य क्रियाओं के प्रभाव देखने में दिलचस्पी हो सकती है। आइए, नीचे दिए गए एक उदाहरण पर विचार करें–

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 X_i + u_i \qquad \text{... (9.7)}$$

जहाँ Y आय है; X स्कूली शिक्षा के वर्षों की संख्या में मापी जाने वाली शिक्षा है; D_2 लिंग है – 0 यदि पुरुष है, 1 यदि महिला है; तथा D_3 वह आरक्षित खंड या समूह (यथा, अनुसूचित जाति/अनुसूचित जनजाति/अन्य पिछड़ा वर्ग) में है अथवा नहीं – 0 यदि वह 'आरक्षित वर्ग में नहीं' अर्थात् सामान्य वर्ग में है, 1 यदि वह 'आरक्षित वर्ग में' है।

यहाँ लिंग (D_2) और आरक्षण (D_3) गुणात्मक चर हैं और X मात्रात्मक चर है। इस निरूपण में (उदाहरण के लिए, समीकरण 9.7) हमने एक अंतर्निहित धारणा बना ली है कि आरक्षण के उक्त दो खंडों में लिंग का अवकल प्रभाव नियत है। हम यह भी मानकर चल रहे हैं कि दो लिंगों में आरक्षण का अवकल प्रभाव नियत है। इसका अर्थ यह है कि यदि पुरुषों की औसत आय महिलाओं की तुलना में अधिक है तो ऐसा इसलिए है कि व्यक्ति सामान्य वर्ग में है अथवा आरक्षण खंड में है।

इसी प्रकार, यहाँ यह भी माना जाता है कि यदि दो आरक्षण खंडों के बीच औसत आय भिन्न–भिन्न है तो यह लिंग पर ध्यान दिए बिना है। बहरहाल, कई मामलों में ऐसी अवधारणाएँ मान्य नहीं हो सकती हैं। इसका अर्थ है कि लिंग और आरक्षण प्रतिरूपों के बीच अन्योन्य क्रिया हो सकती है। दूसरे शब्दों में, औसत आय पर उनका प्रभाव (समीकरण 9.7) की भाँति केवल योगात्मक नहीं हो सकता है, परंतु गुणक अवश्य हो सकता है। यदि हम इस अन्योन्य क्रिया के प्रभाव पर विचार करना चाहते हैं तो हमें अपने मॉडल को निम्नानुसार निर्दिष्ट करना होगा–

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 (D_{2i}\, D_{3i}) + \beta_5 X_i + u_i \qquad \text{... (9.8)}$$

समीकरण (9.8) में प्रतिरूप चर $D_{2i}D_{3i}$ को 'अंतः क्रियात्मक या अन्योन्य प्रतिरूप' कहा जाता है। यह दो गुणात्मक चरों के संयुक्त अथवा समकालिक प्रभाव को निरूपित करता है। समीकरण (9.8) के दोनों पक्षों की प्रत्याशा लेते हुए, यथा लिंग और आरक्षण में आय पर औसत प्रभाव पर विचार करते हुए, हम निम्नलिखित समीकरण प्राप्त करते हैं–

$$E(Yi \mid D_{2i} = 1, D_{3i} = 1, X_i) = \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 X_i \qquad \text{... (9.9)}$$

समीकरण (9.9) महिला आरक्षित श्रेणी के श्रमिकों के लिए औसत आय फलन है, जहाँ β_2 महिला होने का अवकल प्रभाव है, β_3 आरक्षित खंड में होने का अंतर प्रभाव है, और β_4 महिला और आरक्षित वर्ग दोनों में होने का अन्योन्य प्रभाव है। अब विभिन्न प्रतिरूपों की सांख्यिकीय सार्थकता के आधार पर हमें उपयुक्त निष्कर्ष निकालने की आवश्यकता है।

इस विनिर्देशन को एक से अधिक मात्रात्मक चरों और दो से अधिक गुणात्मक चरों के लिए सरलता से सामान्यीकृत किया जा सकता है।

बोध प्रश्न 2

[प्रत्येक प्रश्न का उत्तर लगभग 50—100 शब्दों में दें।]

1) 'स्लोप ड्रिफ्टर' का क्या अर्थ है? इसे कब और किस प्रयोजन से प्रस्तुत किया जाता है? इस प्रकार के प्रतिगामी प्रक्षेपक वाला एक सामान्य मॉडल निर्दिष्ट करें और प्रस्तुत किए गए अतिरिक्त चर पर टिप्पणी करें।

………………………………………………………………………………….

………………………………………………………………………………….

………………………………………………………………………………….

………………………………………………………………………………….

2) समीकरण (9.6) के प्रकार के किसी मॉडल में दो प्रतिगामी प्रक्षेपकों β_2 और β_4 के साथ उस पर विचार करते समय प्राप्त होने वाले चार प्रकार के समाश्रयणों के बीच अंतर स्पष्ट करें।

………………………………………………………………………………….

………………………………………………………………………………….

………………………………………………………………………………….

………………………………………………………………………………….

3) इस इकाई में विचाराधीन रहे विभिन्न उदाहरणों अथवा स्थितियों को समायोजित करने के लिए प्रतिरूप चरों वाले चार प्रकार के समाश्रयण मॉडलों की सूची प्रस्तुत करें। उनके नाम और अभिलक्षणों से उनके बीच अंतर भी निर्दिष्ट करें।

………………………………………………………………………………….

………………………………………………………………………………….

………………………………………………………………………………….

………………………………………………………………………………….

9.6 सार—संक्षेप

यह इकाई गुणात्मक और मात्रात्मक चरों के बीच अंतर स्पष्ट करती है। इसमें तीन प्रकार के मॉडलों पर विचार किया गया है, जिनमें समाश्रयण मॉडल में गुणात्मक चरों को शामिल करने पर ध्यान केंद्रित किया गया है। इस प्रकार के प्रथम मॉडल को एक साधारण समाश्रयण मॉडल माना जाता है। इस मॉडल में हमने समाश्रयण समीकरण के दाहिने पक्ष पर एक स्वतंत्र चर के रूप में केवल एक प्रतिरूप चर पर विचार किया। यह समीकरण निम्नलिखित रूप में दिखाई देता है–

$$Y_i = \beta_1 + \beta_2 D_i + u_i$$

इस रूप में विश्लेषण को एनोवा (ANOVA) कहा जाता है। इस रूप में समाश्रयण मॉडल पर विचार करते समय हम अक्सर ही किसी विनिर्देशन को अभिनत कर रहे होते हैं। ऐसा इसलिए होता है कि चर Y_i स्पष्ट रूप से किसी चर X_i से संबंधित होगा, जो कि एक मात्रात्मक चर होता है। इसे समायोजित करने के लिए हमने दूसरे प्रकार के मॉडल पर विचार किया, जहाँ हमने समाश्रयण समीकरण में एक सहविचर (X_i) शामिल किया, यथा–

$$Y_i = \beta_1 + \beta_2 D + \beta_3 X_i + u_i$$

इस रूप में विश्लेषण को एनकोवा (ANCOVA) कहा जाता है।

उक्त दोनों प्रकार के मॉडलों में हमारा ध्यान केवल अवरोधन में अंतर के महत्व को देखने पर केंद्रित रहा। किंतु व्यवहार में हम ऐसी अनेक स्थितियों का सामना करते हैं जिनमें न केवल अवरोधन बल्कि प्रतिक्रमण भी श्रेणियों के बीच भिन्न–भिन्न हो सकता है। इस प्रकार की स्थिति पर विचार करने के लिए हमने तीसरे प्रकार के मॉडल पर विचार किया, जिसमें हमने 'मात्रात्मक चर वाले प्रतिरूप चर' यथा D_iX_i के अन्योन्य प्रभाव के लिए समीकरण समायोजित किया। इस प्रकार के विश्लेषण के लिए माना जाने वाला समाश्रयण मॉडल निम्नलिखित प्रकार का होता है–

$$Y_i = \beta_1 + \beta_2 D_i + \beta_3 X_i + \beta_4 (D_i X_i) + u_i$$

इस स्थिति में हमने देखा कि हमारे समक्ष चार प्रकार की संभावनाएँ आ सकती हैं – संपाती, समानांतर, समवर्ती और असदृश समाश्रयण।

इकाई के अंत में हमने उस उदाहरण पर विचार किया है जिसमें एक से अधिक गुणात्मक चरों को समायोजित करने के लिए कोई समाश्रयण मॉडल निरूपित करना पड़ सकता है और कोई ऐसा उदाहरण भी जहाँ हम दो गुणात्मक चरों के अन्योन्य प्रभाव की जाँच करने में रुचि रखते हों। इसके लिए हमने निम्नलिखित की भाँति मॉडलों पर विचार किया–

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 (D_{2i}\, D_{3i}) + \beta_5 X_i + u_i.$$

9.6 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1. गुणात्मक चर एक ऐसा चर होता है जिसमें कोई स्पष्ट प्रतिक्रिया होती है, जैसे कि हाँ/नहीं अथवा रोजगार प्राप्त/बेरोजगार अथवा पुरुष/महिला। यदि यह प्रतिक्रिया दो तक ही सीमित होती है, जैसा कि इन उदाहरणों में है, तो इसे द्विभाजित चर कहा जाता है। ये प्रतिक्रियाएँ दो से अधिक भी हो सकती हैं, परंतु तब उन्हें 1, 2, 3, ... के रूप में ही वर्गीकृत किया जा सकेगा। ऐसी प्रतिक्रियाएँ असंदिग्ध या सुस्पष्ट होती हैं। अतः गुणात्मक चर को प्रतिरूप चर अथवा श्रेणीगत चर भी कहा जाता है।
2. इस उदाहरण में मॉडल इस प्रकार हो सकता है : $Y_i = \beta_1 + \beta_2 D_i + u_i$ यहाँ हम आश्रित चर Y_i को मात्रात्मक चर मानकर चल रहे हैं। तदनुसार, चर D_i नियत है और गैर–प्रसंभाव्य है। साथ ही, इसको एक द्विभाजित माना जाता है, यथा यह 0 और 1 मान लेता है। ऐसे उदाहरणों में वह कारक या इकाई जिसे मान 0 दिया जाता है, आधार श्रेणी कहलाती है। चर Y_i, के माध्य का आकलित मान, जबकि $D_i = 0$, ज्ञात है, चर β_1 से दर्शाया गया है। यहाँ चर β_2 नियमनिष्ठतः समाश्रयण गुणांक नहीं है बल्कि 'अवकल अवरोधन गुणांक' है। चर Y_i, के माध्य का आकलित मान, जबकि $D_i = 1$ ज्ञात है, $\beta_1 + \beta_2$ दर्शाया गया है।

3. दो लिंग वर्गों के लिए Y_i का माध्य मान, उसका R^2 मान, आकलित प्रतिरूप चर गुणांक का निरपेक्ष मान और मानक त्रुटियाँ एक समान होंगी। अवरोधन पद का संख्यात्मक मान और उसका t–मान बदल जाएगा।
4. प्रतिरूप चर की प्रतिक्रियाओं की संख्या को प्रतिक्रिया की 'श्रेणियाँ' कहा जाता है। यदि प्रतिरूप चर प्रतिवादी के लिंग को इंगित करता है तो प्रतिक्रिया की दो श्रेणियाँ होती हैं, यथा पुरुष और महिला। यदि हम ऐसे उदाहरणों में दो भिन्न–भिन्न प्रतिरूप आवंटित करते हैं तो हमें पूर्ण संरेखता की स्थिति का सामना करना पड़ता है। अतः हमें अद्वितीय आकलन प्राप्त नहीं होंगे अथवा दो प्राचलों में से किसी एक का आकलन नहीं किया जा सकेगा। इस स्थिति को 'प्रतिरूप चर जाल' कहा जाता है। इस स्थिति से बचने के लिए सामान्य नियम यह है कि यदि हमारे पास m श्रेणियाँ हों तो हम प्रतिरूपों की संख्या को '$m - 1$' तक सीमित कर देते हैं। इन मॉडलों में कोई सामान्य अवरोधन β_1 भी होना चाहिए।
5. यदि लिए गए समाश्रयण मॉडल में सामान्यतः केवल एक ही स्वतंत्र चर हो और वह भी एक प्रतिरूप चर हो, जैसा कि यहाँ विशेष रूप से माना जाता है, तो विचरण या प्रसरणशीलता के वे स्रोत जिनकी पहचान आश्रित चर के लिए की जाती है, उस एक चर तक ही सीमित रहते हैं। ऐसे उदाहरणों में समाश्रयण मॉडल को 'एनोवा मॉडल' कहा जाता है। यदि लिए गए स्वतंत्र चरों की संख्या दो हो और उनमें से एक को प्रतिरूप चर मानकर चला गया हो और दूसरे चर को प्रतिरूप चर से ही संबद्ध मान लिया गया हो तो ऐसे मॉडल को 'एनकोवा मॉडल' कहा जाता है। दूसरे शब्दों में, ऐसे समाश्रयण मॉडल जिनमें कुछ स्वतंत्र चर गुणात्मक हों और कुछ अन्य मात्रात्मक हों तो वे मॉडल 'एनकोवा मॉडल' कहलाते हैं।
6. एनकोवा मॉडल का लाभ यह है कि यह मॉडल सहविचरों के प्रभावों को सांख्यिकीय रूप से नियंत्रित करने की एक विधि प्रदान करता है। किसी सहविचर को इस मॉडल में शामिल करने से बचने का परिणाम यह होता है कि मॉडल 'विनिर्देशन त्रुटि' से ग्रस्त हो जाता है। इन विनिर्देश त्रुटियों का परिणाम यह होता है कि OLS आकलकों के दक्ष होने के लिए आवश्यक आदर्श अवधारणाओं का उल्लंघन करना पड़ता है। फलतः वे अपने दक्षता संबंधी गुणधर्मों से वंचित हो जाते हैं।
7. इस मॉडल का सामान्य रूप इस प्रकार होता है : $Y_i = \beta_1 + \beta_2 D + \beta_3 X_i + u_i$ समाश्रयण गुणांक 'सीमांत उपभोग प्रवृत्ति (MPC)' में वृद्धि (अथवा कमी) की दर को इंगित करता है। ऐसा तब होता है जब आश्रित चर Y भोजन पर व्यय जैसे उपभोग चर से संबंधित होता है और मात्रात्मक स्वतंत्र चर यहाँ मानकर चली गई प्रयोज्य आय की भाँति होता है।

बोध प्रश्न 2

1. एक अवरोधन और एक ही समाश्रयण गुणांक वाले समाश्रयण मॉडल में हमारी रुचि यह जानने के लिए हो सकती है कि– (i) क्या अवरोधन पद सांख्यिकीय रूप से भिन्न हैं, और (ii) क्या समाश्रयण गुणांक सांख्यिकीय रूप से भिन्न हैं? दूसरे प्रश्न की जाँच के लिए हमें 'स्लोप ड्रिफ्टर' के रूप में जाने वाले प्रतिगामी प्रक्षेपक प्रस्तुत करना होगा। ऐसे प्रक्षेपक के साथ निर्दिष्ट मॉडल इस प्रकार होगा : $Y_i = \beta_1 + \beta_2 D_i + \beta_3 X_i + \beta_4 (D_i X_i) + u_i$ यहाँ प्रस्तुत किया गया अतिरिक्त चर $DiXi$ है। यह अन्योन्य रूप में एक गुणक चर है। यहाँ β_2 और β_4 प्रतिक्रमण वाले ऐसे दो प्रक्षेपक हैं जो हमें

क्रमशः अवरोधन चर और समाश्रयण चर में सांख्यिकीय अंतर का आकलन करने में मदद करते हैं।

2. जब अवरोधन और समाश्रयण दोनों में कोई अंतर नहीं होता है तो हमें एक 'संपाती समाश्रयण' प्राप्त होता है। जब दो अवरोधन पद भिन्न हों परंतु दो समाश्रयण गुणांक एक समान हों तो हमें एक 'समानांतर समाश्रयण' प्राप्त होता है। जब दो समाश्रयणों में एक ही अवरोधन होता हो परंतु भिन्न–भिन्न प्रतिक्रमण हों तो हमें एक 'समवर्ती समाश्रयण' प्राप्त होता है। जब अवरोधन पद और समाश्रयण गुणांक दोनों ही भिन्न–भिन्न हों तो हमें दो 'असदृश समाश्रयण' प्राप्त होते हैं।
3. (i) $Y_i = \beta_1 + \beta_2 D_i + u_i$ (ii) $Y_i = \beta_1 + \beta_2 D + \beta_3 X_i + u_i$ (iii) $Y_i = \beta_1 + \beta_2 D_i + \beta_3 X_i + \beta_4 (D_i X_i) + u_i$ (iv) $Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 (D_{2i}\, D_{3i}) + \beta_5 X_i + u_i$. यहाँ प्रथम एनोवा मॉडल है, जिसमें हमने केवल एक एकल प्रतिरूप चर को स्वतंत्र चर माना है। दूसरा एनकोवा मॉडल है, जिसमें हमने एक गुणात्मक प्रतिरूप चर और एक अन्य ऐसे मात्रात्मक बहिर्जात चर पर विचार किया है जो प्रतिरूप चर से संबंधित है और जिसको छोड़ देने से इसमें 'विनिर्देशन अभिनति' दिखाई देगी। तीसरे में एक अन्योन्य चर $(D_i X_i)$ शामिल है, जिसमें हम यह देखने की कोशिश करते हैं कि क्या समाश्रयण और अवरोधन गुणांक दोनों भिन्न–भिन्न हैं। इसमें चार भिन्न–भिन्न प्रकार के समाश्रयण होने की संभावना होती है, यथा संपाती, समानांतर, समवर्ती और असदृश प्रतिगमन। मानी जाने वाली चौथी स्थिति में एक अन्योन्य प्रतिरूप चर शामिल है, यथा $Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 (D_{2i}\, D_{3i}) + \beta_5 X_i + u_i$.