इकाई 4 सरल रैखिक समाश्रयण मॉडल : आकलन*

इकाई की रूपरेखा

4.0 उद्देश्य

4.1 विषय–प्रवेश

4.2 समष्टि समाश्रयण फलन (PRF)

4.2.1 नियतात्मक घटक

4.2.2 प्रसंभाव्य घटक

4.3 प्रतिदर्श समाश्रयण फलन (SRF)

4.4 पारंपरिक समाश्रयण मॉडल की अवधारणाएँ

4.5 आकलन की साधारण न्यूनतम वर्ग विधि

4.6 OLS आकलकों के बीजगणितीय गुणधर्म

4.7 निर्धारण गुणांक (coefficient of determination)

4.7.1 R^2 की संगणना का सूत्र

4.7.2 सुमानकता हेतु F–प्रतिदर्शज

4.7.3 चर F और R^2 के बीच संबंध

4.7.4 चर F और t^2 के बीच संबंध

4.8 सार–संक्षेप

4.9 बोध प्रश्नों के उत्तर अथवा संकेत

4.0 उद्देश्य

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

- पारंपरिक रैखिक समाश्रयण मॉडल की व्याख्या कर सकें;
- समष्टि समाश्रयण फलन (PRF) और प्रतिदर्श समाश्रयण फलन (SRF) के बीच अंतर स्पष्ट कर सकें;
- साधारण न्यूनतम वर्ग (OLS) आकलक ज्ञात कर सकें;
- साधारण न्यूनतम वर्ग (OLS) आकलकों के गुणधर्म बता सकें;
- समाश्रयण फलन से जुड़ी सुमानकता संबंधी संकल्पना को स्पष्ट कर सकें; तथा
- निर्धारण का गुणांक एवं उसके गुणधर्मों का वर्णन कर सकें।

4.1 विषय–प्रवेश

इस पाठ्यक्रम के अंतर्गत 'BECC 107: अर्थशास्त्र के लिए सांख्यिकीय प्रविधियाँ' की इकाई 5 में हमने सहसंबंध और समाश्रयण विषयों पर चर्चा की थी। उस इकाई में हमने समाश्रयण की अवधारणा के विषय में एक संक्षिप्त विचार प्रस्तुत किया था। हम पहले से ही जानते हैं कि

---

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

समाश्रयण विश्लेषण में दो प्रकार के चर होते है, यथा – i) आश्रित (या व्याख्यायित) चर, और ii) स्वतंत्र (या व्याख्यात्मक) चर। जैसा कि ये नाम (व्याख्यायित और व्याख्यात्मक) इंगित करते हैं, आश्रित चर की व्याख्या स्वतंत्र चर द्वारा की जाती है।

आमतौर पर हम आश्रित चर को Y और स्वतंत्र चर को X के रूप में निरूपित करते हैं। मान लीजिए कि हमने एक घरेलू सर्वेक्षण किया और X और Y में n प्रेक्षणों के जोड़े एकत्र किए। चर X और Y के बीच संबंध अनेक रूप ले सकते हैं। सामान्य व्यवहार में यह संबंध किसी गणितीय समीकरण से दर्शाया जाता है। इन समीकरणों में सबसे सरल रैखिक समीकरण होता है। इसका अर्थ है कि X और Y के बीच संबंध एक सीधी रेखा के रूप में होता है, और इसीलिए, इसे रैखिक समाश्रयण कहा जाता है। जब यह समीकरण वक्रों का प्रतिनिधित्व करता है (न कि किसी सीधी रेखा का) तो समाश्रयण को 'गैर–रेखीय' अथवा 'वक्रीय' कहा जाता है।

इस प्रकार, सामान्य नियमानुसार, उक्त संबंध को हम X और Y के बीच संबंध के रूप में व्यक्त कर सकते हैं, जैसा कि नीचे समीकरण (4.1) में दर्शाया गया है –

$$Y = f(X) \qquad \dots (4.1)$$

इस खंड में (इकाई 4, 5 और 6) हम केवल दो चरों वाले सरल रैखिक समाश्रयण मॉडलों पर ही विचार करेंगे। एक से अधिक व्याख्यात्मक चर वाले बहु समाश्रयण मॉडल पर अगले खंड में चर्चा की जाएगी।

किसी भी समाश्रयण विश्लेषण के उद्देश्य निम्नवत् हो सकते हैं–

- स्वतंत्र चरों के मान ज्ञात होने पर आश्रित चर के माध्य अथवा औसत मान का आकलन करना।
- अंतर्निहित आर्थिक सिद्धांत के संबंध में परिकल्पनाओं का परीक्षण करना। उदाहरण के लिए, हम इस परिकल्पना का परीक्षण कर सकते हैं कि माँग की मूल्य लोचता (–)1 होती है, यथा माँग पूरी तरह से लोचदार होती है, यह मानते हुए कि माँग को प्रभावित करने वाले अन्य कारकों को नियत रखा गया है।
- स्वतंत्र चरों के मान ज्ञात होने पर आश्रित चर के माध्य मान का पूर्वानुमान व्यक्त करना।

4.2 समष्टि समाश्रयण फलन (PRF)

कोई भी समष्टि समाश्रयण फलन किसी आश्रित चर और किसी स्वतंत्र या व्याख्यात्मक चरों के किसी समुच्चय के बीच एक सैद्धांतिक संबंध की परिकल्पना करता है। यह एक रैखिक फलन होता है। यह फलन ही परिभाषित करता है कि कैसे किसी चर Y की सशर्त प्रत्याशा स्वतंत्र चर X में परिवर्तनों के प्रति अपनी प्रतिक्रिया व्यक्त करती है।

$$Y_i = E(Y_i|X_i) + u_i \qquad \dots (4.2)$$

उक्त फलन में एक 'नियतात्मक' अर्थात् निर्धारणात्मक घटक $E(Y|X)$ और एक 'प्रसामान्य' अर्थात् गैर–नियतात्मक घटक, u, होता है, जैसा कि समीकरण (4.2) में दर्शाया गया है। यहाँ हमारा सरोकार आश्रित चर (X) के दिए गए मानों पर सशर्त आश्रित चर (Y) के निर्धारक तत्वों की जाँच करने से ही रहेगा।

4.2.1 नियतात्मक घटक

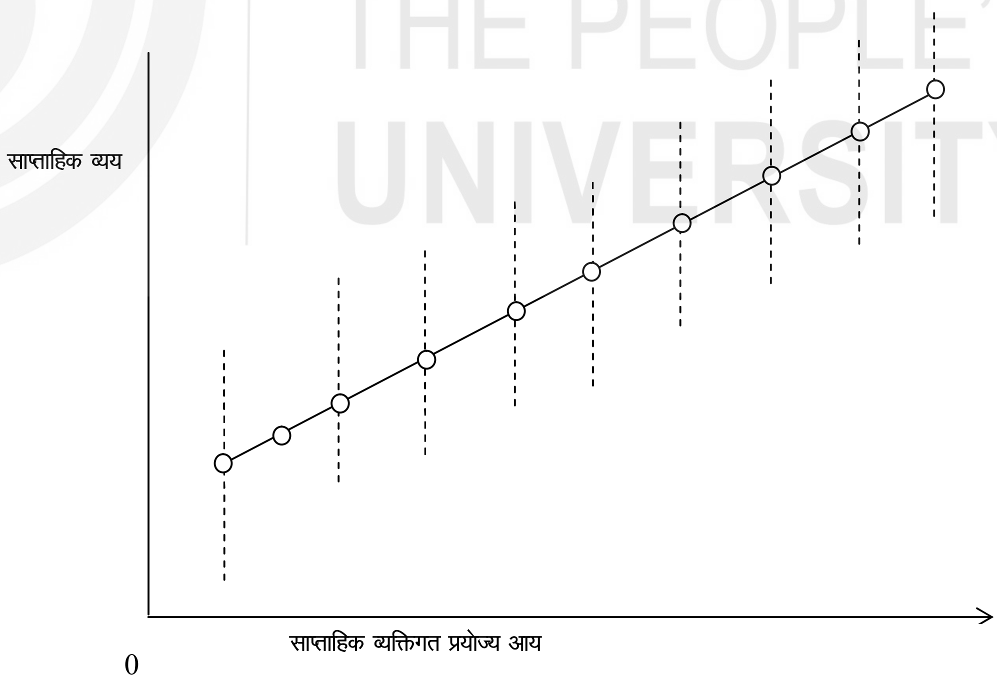
चर Y की सशर्त प्रत्याशा समाश्रयण मॉडल के नियतात्मक घटक का गठन करती है। यह एक नियतात्मक रेखा के रूप में प्राप्त होता है। इसे समष्टि समाश्रयण रेखा (population regression line - PRL) के नाम से भी जाना जाता है। दूसरी ओर, गैर–नियतात्मक या प्रसामान्य घटक को चर u_i से इंगित एक यादृच्छिक त्रुटि पद (random error term) द्वारा दर्शाया जाता है।

चलिए, एक उदाहरण लेते हैं। मान लीजिए कि हम समष्टि के एक समुच्चय के लिए साप्ताहिक व्यय पर साप्ताहिक व्यक्तिगत प्रयोज्य आय (PDI) के प्रभाव की जाँच करना चाहते हैं। इसके लिए हम साप्ताहिक PDI को स्वतंत्र चर (Y) और साप्ताहिक व्यय को आश्रित चर (X) के रूप में लेकर चलेंगे। साप्ताहिक PDI के प्रत्येक दिए गए मान के लिए साप्ताहिक व्यय का औसत मान ऊर्ध्वाधर अक्ष पर अंकित किया जाता है।

उच्च आयवर्ग के लोग अपेक्षाकृत अधिक खर्च करने की संभावना रखते हैं, अतः सहज ज्ञान से, साप्ताहिक PDI और साप्ताहिक व्यय के बीच संबंध सकारात्मक होता है। तदनुसार, निम्नलिखित समष्टि समाश्रयण रेखा प्राप्त की जाती है और एक लेखाचित्र पर अंकित कर दी जाती है, जैसा कि नीचे स्पष्ट किया गया है–

$$E(Y_i|X_i) = \beta_1 + \beta_2 X_i \qquad \dots (4.3)$$

यह ध्यान देने की बात है कि समीकरण (4.3) में β_1 और β_2 प्राचल हैं। यहाँ β_1 समष्टि समाश्रयण फलन का अवरोधन है। यह व्याख्यात्मक चर के शून्य होने पर आश्रित चर के अपेक्षित मान को इंगित करता है। इसके अलावा, β_2 समष्टि समाश्रयण फलन का प्रतिक्रमण है। यह उस परिमाण को इंगित करता है जिसके द्वारा स्वतंत्र चर में एक इकाई परिवर्तन होने पर आश्रित चर बदल जाएगा। समष्टि प्राचल समष्टि में आश्रित चर और स्वतंत्र चर के बीच संबंध का वर्णन करते हैं।



चित्र 4.1: साप्ताहिक व्यक्तिगत प्रयोज्य आय

ऊपर दिए गए चित्र 4.1 में घेरे गए बिंदुओं को ध्यान से देखें। ये बिंदु विभिन्न X_i के अनुरूप Y के माध्य या औसत मान को निरूपित करते हैं। इन्हें ही सशर्त माध्य या सशर्त अपेक्षित मान कहा जाता है। यदि हम Y के विभिन्न अपेक्षित मानों को जोड़ देते हैं तो परिणामी रेखाएँ समष्टि समाश्रयण रेखा (PRL) कहलाती हैं।

4.2.2 प्रसंभाव्य घटक

जब हम किसी प्रतिदर्श से आँकड़े एकत्र करते हैं तो हम X और Y के बीच कोई नियतात्मक संबंध नहीं बनाते हैं। उदाहरण के लिए, एक समान स्तर की आय के लिए दो व्यक्तियों का व्यय भिन्न–भिन्न हो सकता है। मान लीजिए कि दो व्यक्ति हैं, जिनकी कुल मासिक आय रु. 20,000 प्रति माह है। जबकि एक व्यक्ति का मासिक खर्च रु. 15,000 है, दूसरे व्यक्ति का मासिक खर्च रु. 19,000 हो सकता है। दूसरे व्यक्ति के मासिक खर्च में अंतर उसकी स्वास्थ्य स्थिति या जीवन शैली के कारण अधिक हो सकता है। आश्रित चर में इस प्रकार के अंतर का प्रग्रहण प्रसामान्य त्रुटि पद द्वारा किया जाता है। चित्र 4.1 में चर X के किसी विशिष्ट मान के लिए चर Y के मान को एक ऊर्ध्वाधर बिंदुकित रेखा द्वारा दर्शाया गया है। चर X के किसी विशिष्ट मान के लिए चर Y का अपेक्षित मान घेर दिया गया है (देखें चित्र 4.1)।

इस प्रकार, X और Y के बीच प्रसामान्य संबंध निर्दिष्ट करने की आवश्यकता होती है। प्रतिदर्श समाश्रयण फलन (sample regression function - SRF) का विनिर्देशन निम्नवत् होता है –

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \ldots (4.5)$$

उक्त समीकरण (4.5) में पद u_i को ही प्रसामान्य त्रुटि या यादृच्छिक त्रुटि कहा जाता है।

इस समीकरण (4.5) का पहला घटक नियतात्मक घटक $(\beta_1 + \beta_2 X_i)$ कहलाता है, जिस पर हम पहले ही चर्चा कर चुके हैं। यह नियतात्मक घटक ही विचाराधीन उदाहरण में माध्य या औसत व्यय कहलाता है। इस नियतात्मक घटक को नियमित या निर्धारणात्मक घटक भी कहा जाता है।

यहाँ दूसरा घटक u_i यादृच्छिक घटक कहलाता है (आय के अलावा अन्य कारकों द्वारा गैर–व्यवस्थित रूप से निर्धारित)। त्रुटि पद u_i को 'रव घटक' (noise component) के रूप में भी जाना जाता है। त्रुटि पद u_i एक यादृच्छिक चर होता है। अतः घटक u_i का मान नियंत्रित अथवा ज्ञात नहीं किया जा सकता है।

किसी भी समाश्रयण मॉडल में त्रुटि पद u_i को शामिल करने के तीन कारण होते हैं, यथा – (i) यह त्रुटि पद उन चरों के प्रभाव को दर्शाता है जिन्हें उस समाश्रयण मॉडल में स्पष्ट रूप से प्रस्तुत न किया गया हो। उदाहरण के लिए, ऐसे अनेक चर होते हैं जो किसी घर के उपभोग व्यय को प्रभावित करते हैं (जैसे परिवार के सदस्यों की संख्या, स्वास्थ्य की स्थिति, पास–पड़ोस, आदि)। ये चर आश्रित चर को प्रभावित करते हैं, और X एवं Y के बीच अंतर्भूत यादृच्छिकता विद्यमान होती है। (ii) मानव व्यवहार पूर्वानुमेय नहीं होता। इस प्रकार की अनियमितता का चिंतन एवं प्रग्रहण यादृच्छिक त्रुटि पद द्वारा ही किया जाता है। (iii) आँकड़ों को मापने में त्रुटियाँ सामने आती हैं, जैसे वार्षिक पारिवारिक आय का पूर्णांकन, विद्यालय से कई छात्रों की अनुपस्थिति, आदि।

उक्त अनियमितता के कारण ही आँकड़ों का वास्तविक मान या तो आश्रित चर के अपेक्षित मान से ऊपर रहता है या फिर नीचे। दूसरे शब्दों में, वास्तविक मान औसत, यथा नियमित घटक से विचलित हो जाता है।

समष्टि समाश्रयण फलन और समष्टि समाश्रयण रेखा (PRL) की प्रारंभिक अवधारणा को समझ लेने के बाद अब अगले पाठांश में प्रतिदर्श का प्रयोग कर इस रेखा (PRL) के आकलन का वर्णन किया जाएगा।

बोध प्रश्न 1

1) किसी समाश्रयण मॉडल का आकलन किन उद्देश्यों को लेकर किया जाता है?

……………………………………………………………………………………

……………………………………………………………………………………

2) आश्रित चरों का औसत मान उनके वास्तविक मान से भिन्न क्यों होता है?

……………………………………………………………………………………

……………………………………………………………………………………

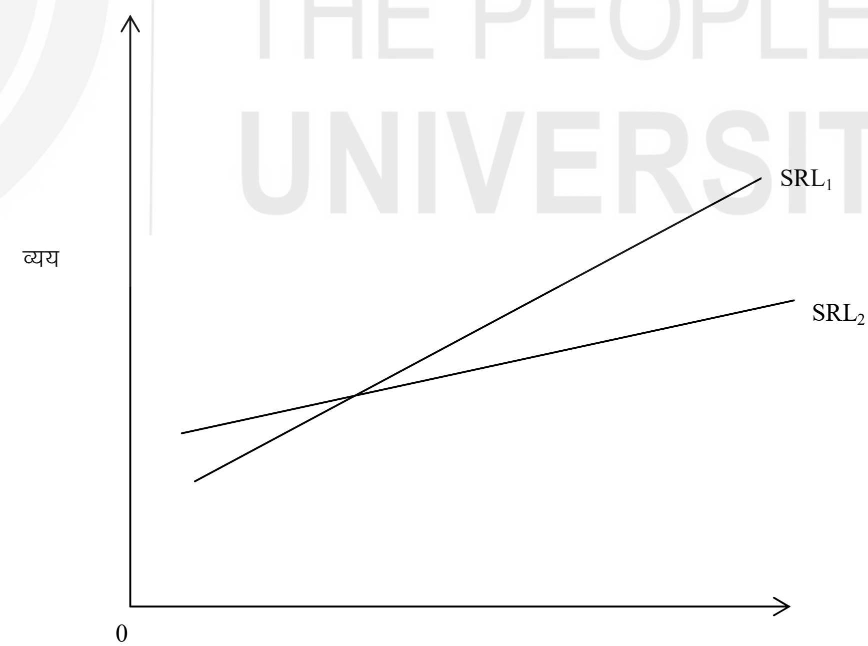
3) किसी समाश्रयण मॉडल में हम त्रुटि पद (u_i) क्यों शामिल करते हैं?

……………………………………………………………………………………

……………………………………………………………………………………

4.3 प्रतिदर्श समाश्रयण फलन (SRF)

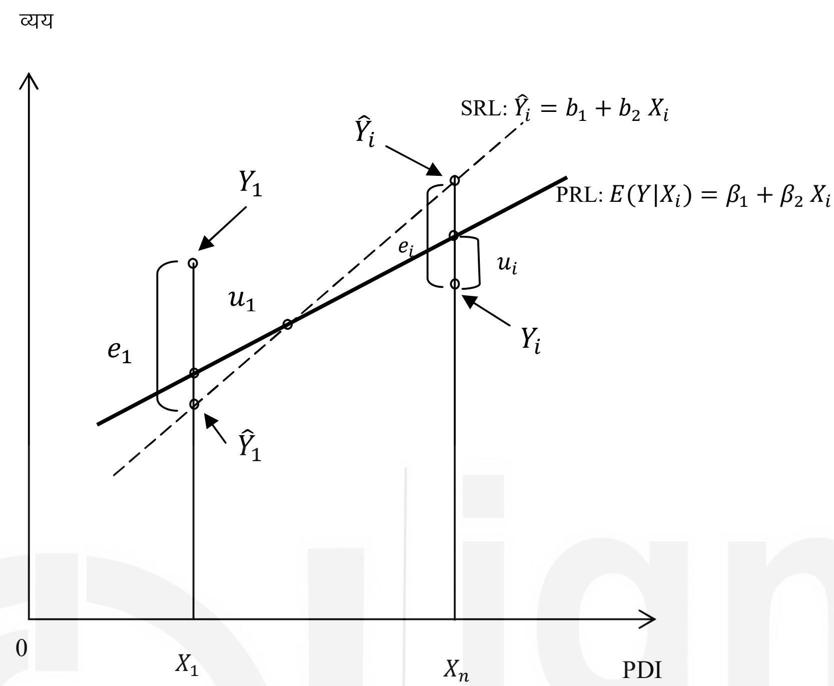
हमें कदाचित ही समस्त समष्टि से संबंधित आँकड़े उपलब्ध हो सकते हैं। हमारे पास समष्टि से लिया गया कोई एक प्रतिदर्श मात्र ही होता है। तदनुसार हमें समष्टि प्राचलों का आकलन करने के लिए प्रतिदर्श का ही प्रयोग करने की आवश्यकता पड़ती है। प्रतिदर्शन उच्चावचन अथवा प्रतिदर्शन त्रुटि के कारण हम समष्टि समाश्रयण रेखा (PRL) का पता लगाने में विफल रह सकते हैं। मान लीजिए कि हमारे पास दी गई समष्टि से लिए गए दो प्रतिदर्श हैं। इन प्रतिदर्शों का अलग–अलग प्रयोग करके हम प्रतिदर्श समाश्रयण रेखाएँ (SRL) प्राप्त करते हैं। एक प्रतिदर्श समष्टि का प्रतिनिधित्व करता है। चित्र 4.2 में हमने ये दोनों प्रतिदर्श समाश्रयण रेखाएँ SRL_1 और SRL_2 के नाम से दर्शाई हैं।



चित्र 4.2: दो प्रतिदर्श समाश्रयण रेखाएँ

उक्त दोनों प्रतिदर्श समाश्रयण रेखाएँ समष्टि समाश्रयण रेखा को निरूपित करती हैं। हालाँकि, प्रतिदर्शन में उतार–चढ़ाव के कारण इन दोनों रेखाओं (SRL) के प्रतिक्रमण और अवरोधन

भिन्न–भिन्न हैं। समष्टि समाश्रयण फलन (PRF) के अनुरूप, जो कि समष्टि समाश्रयण रेखा (PRL) को अवशायी करता है, हम प्रतिदर्श समाश्रयण रेखा (SRL) और त्रुटि पद u_i से युक्त प्रतिदर्श समाश्रयण फलन (SRF) की अवधारणा विकसित करते हैं।



चित्र 4.3: समष्टि समाश्रयण रेखा और प्रतिदर्श समाश्रयण रेखा

ऊपर दिया गया चित्र 4.3 समष्टि समाश्रयण रेखा (PRL) और प्रतिदर्श समाश्रयण रेखा (SRL) दर्शाता है। आप देखेंगे कि इन दोनों रेखाओं की प्रवणता या ढाल भिन्न–भिन्न हैं। तदनुसार,

$$b_1 \neq \beta_1 \text{ और } b_2 \neq \beta_2$$

आइए, व्याख्यात्मक चर X_1 के एक विशिष्ट मान पर विचार करें। यहाँ व्याख्यायित चर का संगत मान Y_1 है। प्रतिदर्श समाश्रयण रेखा के आधार पर हम व्याख्यायित चर का आकलित मान $\hat{Y}_1$ ज्ञात करते हैं।

चलिए, पहले त्रुटि पद (u) और अवशिष्ट (e) के बीच अंतर ज्ञात करते हैं। समष्टि समाश्रयण रेखा पर वास्तविक मान Y_1 और संगत बिंदु के बीच का अंतर u_1 होगा। यह त्रुटि u_1 हमें ज्ञात नहीं है क्योंकि हम β_1 और β_2 का मान नहीं जानते हैं। हम जो जानते हैं वह पद $\hat{Y}_1$ ही है, जिसका आकलन b_1 और b_2 के आधार पर किया जाता है। चर Y_1 और $\hat{Y}_1$ के बीच का अंतर अवशिष्ट e_1 कहलाता है।

ऊपर समीकरण (4.2) में दी गई समष्टि समाश्रयण रेखा निम्नवत् होगी –

$$Y_i = E(Y_i|X_i) + u_i$$

वह प्रतिदर्श समाश्रयण रेखा जिसका हम आकलन करते हैं, निम्नवत् दर्शायी जाती है –

$$\hat{Y}_i = b_1 + b_2 X_i \qquad \ldots (4.6)$$

समीकरण (4.6) में प्रतीक (^) को 'हैट' अथवा 'कैप' को पढ़ा जाता है। तदनुसार, $\hat{Y}_i$ को 'Y_i–हैट' पढ़ा जाएगा। यह ध्यान देने की बात है कि पद $\hat{Y}_i$ चर X_i के दिए गए मानों के लिए E(Y|X$_i$) वास्तविक माध्य मानों को कम करके ही आँकता है।

हम जिनका प्रेक्षण करते हैं वे प्रतिरूप b_1, b_2 और e होते हैं।

$$Y_i = \hat{Y}_i + e_i = b_1 + b_2X_i + e_i \qquad \text{... (4.7)}$$

जहाँ $\hat{Y}_i$ = E (Y | X_i) का आकलक है।

समीकरण (4.7) में समष्टि सशर्त माध्य का आकलक $\hat{Y}_i$ एक आकलक (अथवा एक प्रतिदर्श प्रतिदर्शज) है। किसी आकलक द्वारा प्राप्त विशिष्ट मान को ही आकल (estimate) कहा जाता है।

चर Y का वास्तविक मान अवशिष्ट पद को Y के आकलित मान में जोड़कर प्राप्त किया जाता है, जिसे 'अवशिष्ट' कहा जाता है। यह अवशिष्ट समष्टि समाश्रयण फलन के यादृच्छिक त्रुटि पद का आकलित मान होता है। समीकरण (4.7) में प्रतिदर्श समाश्रयण फलन $\hat{Y}_i$ द्वारा दी गई प्रतिदर्श समाश्रयण रेखा और आकलित अवशिष्ट पद e_i का संयोजन दर्शाता है। चित्र 4.3 में गहरी काली ऋजु रेखा ही समष्टि समाश्रयण रेखा (PRL) कहलाती है और यह निम्नलिखित समीकरण द्वारा दर्शायी जाती है–

$$E(Y \mid X) = \beta_1 + \beta_2 X_i \qquad \text{… (4.8)}$$

अतएव, अब समष्टि समाश्रयण फलन (PRF) को निम्नवत् प्रस्तुत किया जा सकता है –

$$Y_i = E(Y_i|X_i) + u_i$$

अथवा,

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \text{… (4.9)}$$

इस प्रकार, समीकरण (4.9) में दर्शाया गया समष्टि समाश्रयण फलन उक्त समष्टि समाश्रयण रेखा (PRL) $E(Y_i|X_i)$ और यादृच्छिक त्रुटि पद u_i का एक संयोजन ही है।

आप देखेंगे कि प्रतिदर्श समाश्रयण फलन (SRF) समष्टि समाश्रयण फलन (PRF) का सन्निकटन मात्र है। हम सर्वाधिक उपयुक्त प्रतिदर्श खोजने का प्रयास करते हैं, जिससे हमें आकलक के रूप में b_1 और b_2 प्राप्त होते हैं, जो कि समष्टि प्राचलों β_1 और β_2 के यथासंभव सन्निकट मान ही दर्शाते हैं। दूसरे शब्दों में, पद b_1 पद β_1 के यथासंभव सन्निकट होता है जबकि पद b_2 पद β_2 के यथासंभव सन्निकट होता है।

4.4 पारंपरिक समाश्रयण मॉडल की अवधारणाएँ

कोई भी रैखिक समाश्रयण मॉडल नीचे निर्दिष्ट कुछ अवधारणाओं पर आधारित होता है। यदि कोई समाश्रयण मॉडल कुछ पूर्वमान्यताओं का पालन करता है तो उसे पारंपरिक रैखिक समाश्रयण मॉडल (CLRM) कहा जाता है। इस मॉडल (CLRM) की ये अवधारणाएँ या पूर्व मान्यताएँ निम्नलिखित हैं–

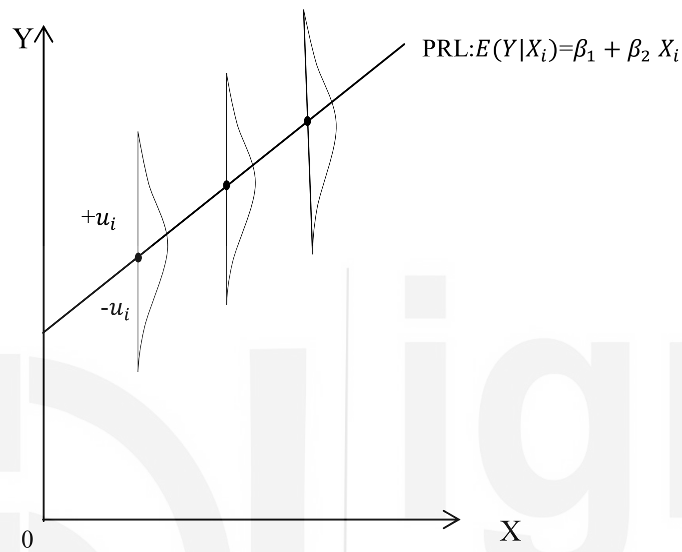
1) कोई भी समाश्रयण मॉडल अपने प्राचलों में रैखिक होता है। यह चरों में रैखिक हो भी सकता है और नहीं भी। उदाहरण के लिए, नीचे दिया गया समीकरण (4.10) प्राचलों के साथ–साथ चरों में भी रैखिक है, जैसा कि समीकरण से स्पष्ट है–

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \text{… (4.10)}$$

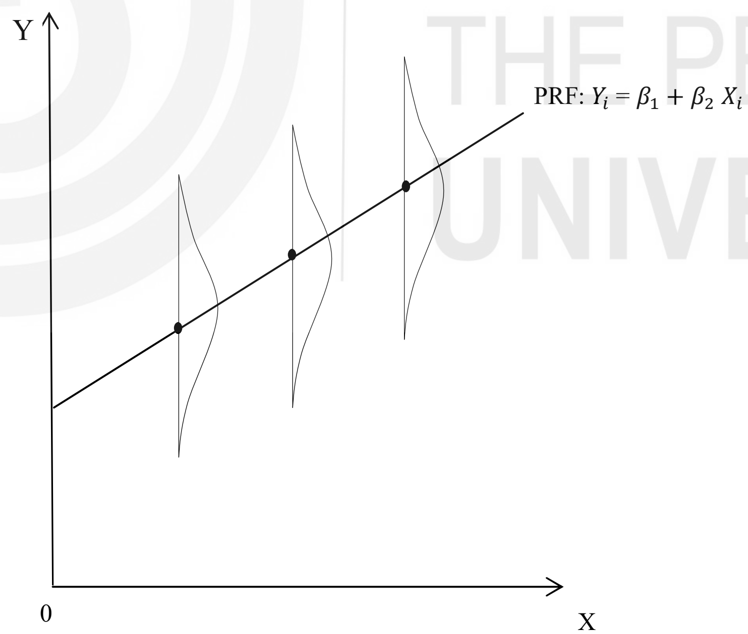
2) व्याख्यात्मक चर बाधा पद u के साथ सहसंबद्ध नहीं होता है। इस अवधारणा की अपेक्षा यह होती है कि वह $\sum u_i X_i = 0$ हो। दूसरे शब्दों में, त्रुटि पद और व्याख्यात्मक चर के बीच सहप्रसरण शून्य होता है। यदि X गैर–प्रसामान्य हो तो यह अवधारणा स्वतः ही पूरी हो जाती है। इसमें अपेक्षित होता है कि X_i मानों को आवर्ती प्रतिदर्शों में नियत रखा जाए।

3) त्रुटि पद u का अपेक्षित मान या माध्य मान शून्य होता है। संकेताक्षरों में, $E(u_i|X_i) = 0$ इसका अर्थ यह नहीं होता है कि सभी त्रुटि पद शून्य होते हैं। इसका अर्थ यह होता है कि त्रुटि पद एक दूसरे को निरस्त कर देते हैं।

4) प्रत्येक u_i का विचरण नियत होता है। संकेताक्षरों में, अंत $var(u_i) = \sigma^2$ त्रुटि पद का सशर्त बंटन चित्र 4.4 (a) में दर्शाया गया है। त्रुटि पद के विशिष्ट मान के लिए संगत त्रुटि विचरण को चित्र 4.4 (b) में दर्शाया गया है। इस चित्र से आप यह पता लगा सकते हैं कि X चर के सभी स्तरों पर त्रुटि विचरण नियत है। यह 'समविसारिता' के उदाहरण की व्याख्या करता है।

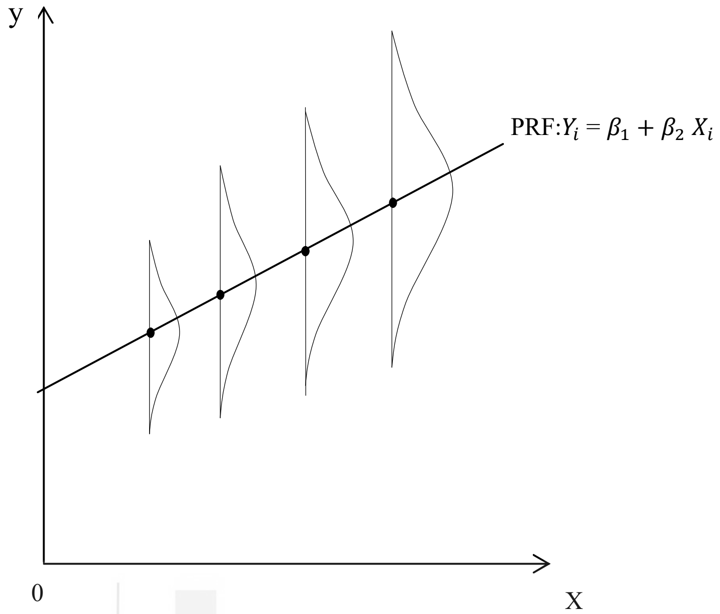


चित्र 4.4 (a) त्रुटि पद u_i का सशर्त बंटन



चित्र 4.4 (b) समविसारिता (एक समान विचरण)

नीचे दिया गया चित्र 4.5 असमान त्रुटि विचरण के उदाहरण को दर्शाता है, यथा विषमविसारिता (heteroscedasticity)। यहाँ त्रुटि पदों का विचरण X_i के मानों में भिन्न होता है।



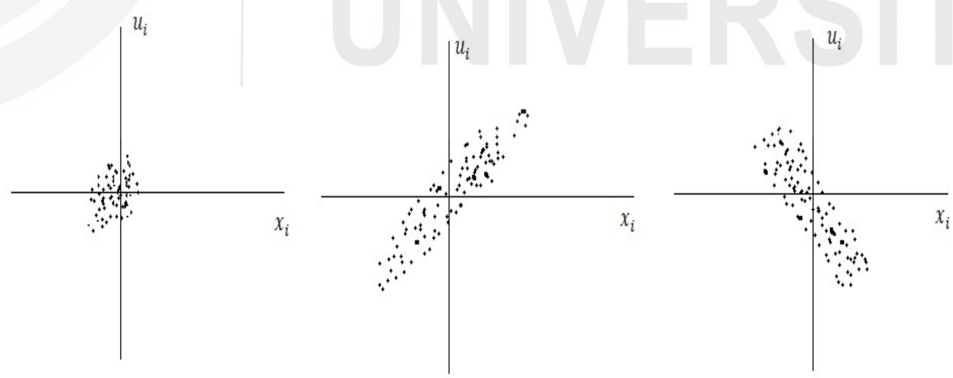
चित्र 4.5: विषमविसारिता का उदाहरण (असमान विचरण)

5) दो त्रुटि पदों के बीच कोई सहसंबंध नहीं होता है। इसी को 'शून्य स्वसहसंबंध' संबंधी अवधारणा कहा जाता है।

$$cov(u_i, u_j) = 0 \qquad i \neq j$$

इसका अर्थ है कि दो त्रुटि पदों के बीच कोई नियमित संबंध नहीं है। इस अवधारणा का निहितार्थ यह है कि त्रुटि पद u_i यादृच्छिक हैं। चूँकि दो त्रुटि पदों को सहसंबद्ध नहीं माना जाता है, कोई भी दो Y मान असंबद्ध ही सिद्ध होंगे, अर्थात् $cov(Y_i, Y_j) = 0.$

नीचे चित्र 4.6 (i) में शून्य स्वसहसंबंध की स्थिति को दर्शाया गया है। यहीं चित्र 4.6 (ii) धनात्मक स्वसहसंबंध को दर्शाता है, और चित्र 4.7 (iii) ऋणात्मक स्वसहसंबंध को दर्शाता है।



(i) शून्य स्वसहसंबंध (ii) धनात्मक स्वसहसंबंध (iii) ऋणात्मक स्वसहसंबंध

चित्र 4.6: शून्य स्वसहसंबंध के विभिन्न उदाहरण

6) आपका समाश्रयण मॉडल सही ढंग से निर्दिष्ट हो अर्थात् मॉडल में कोई विनिर्देशन त्रुटि न हो। यदि कोई प्रासंगिक चर शामिल न हो अथवा या कोई अप्रासंगिक चर इस

मॉडल में शामिल हो तो इससे ही विनिर्देशन त्रुटि होती है। उदाहरण के लिए, मान लीजिए कि हम मोटरवाहनों की माँग का अध्ययन करते हैं। यदि हम केवल मोटरवाहनों की कीमत लेते हैं और उपभोक्ता आय की आय को शामिल नहीं करते हैं तो यहाँ एक विनिर्देशन त्रुटि है। इसी प्रकार, यदि हम विज्ञापन, वित्त पोषण, गैसोलीन की कीमतें आदि को लागत में शामिल नहीं करते हैं तो हम कोई विनिर्देशन त्रुटि ही करेंगे (विनिर्देशन त्रुटि के विषय पर हम इकाई 13 में चर्चा करेंगे)।

4.5 आकलन की साधारण न्यूनतम वर्ग विधि

जैसा कि इस पाठ्यक्रम की इकाई 1 में बताया गया, हमें अपने समाश्रयण मॉडल के प्राचलों का आकलन करने की आवश्यकता पड़ती है। इन प्राचलों के आकलन के लिए अनेक विधियाँ प्रचलन में हैं। बहरहाल, हम यहाँ ऐसी दो ही विधियों के विषय में चर्चा करेंगे, यथा – (i) न्यूनतम वर्ग, और (ii) अधिकतम संभावना। नीचे पहले साधारण न्यूनतम वर्ग (OLS) विधि परविस्तृत चर्चा की गई है।

साधारण न्यूनतम वर्ग (OLS) विधि में वर्गों के त्रुटि योगफल (ESS) को न्यूनतम करके ही किसी रैखिक समाश्रयण मॉडल के प्राचलों का आकलन किया जाता है। दूसरे शब्दों में, यह प्रेक्षित आश्रित चर (Y_i) और आश्रित चर $(\hat{Y}_i)$ के आकलित अथवा अपेक्षित मानों के बीच अंतर के वर्गों के योगफल को न्यूनतम करता है।

संकेताक्षरों में,

$$e_i = (\mathrm{Y_i} - \hat{Y}_i)$$

$$e_i^2 = (\mathrm{Y_i} - \hat{Y}_i)^2$$

$$\textstyle\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (\mathrm{Y_i} - \hat{Y}_i)^2 \qquad \ldots (4.11)$$

उक्त (OLS) विधि में हम $\sum_{i=1}^{n} e_i^2$ को न्यूनतम करते हैं।

हमें ज्ञात है कि –

$$\hat{Y}_i = b_1 + b_2 X_i$$

यदि $\hat{Y}_i$ के मान को समीकरण (4.11) में प्रतिस्थापित कर दिया जाए तो हमें निम्नलिखित समीकरण प्राप्त होता है –

$$\textstyle\sum_{i=1}^{n} e_i^2 = \Sigma(Y_i - b_1 - b_2 X_i)^2$$

न्यूनतमीकरण की प्रथम कोटि शर्त पूरी करने के लिए आवश्यक है कि आंशिक अवकलक शून्य के बराबर हों। यहाँ ध्यान दें कि हमें b_1 और b_2 के मानों पर निर्णय इस तरह लेना होता है कि ESS न्यूनतम रहे। तदनुसार, हमने b_1 और b_2 के संबंध में आंशिक अवकलक लिया है। इसका अर्थ यह है कि –

$$\frac{\partial \Sigma e_i^2}{\partial b_1} = 0 \qquad \ldots (4.13)$$

तथा $$\frac{\partial \Sigma e_i^2}{\partial \mathrm{b}_2} = 0 \qquad \ldots (4.14)$$

समीकरण (4.13) से हमें ज्ञात होता है कि –

$$2\Sigma(Y_i - b_1 - b_2 X_i)(-1) = 0$$

उपर्युक्त समीकरण में आए पदों को पुनर्व्यवस्थित कर हमें प्राप्त होता है –

$$\Sigma Y_i = nb_1 + b_2 \Sigma X_i \qquad \dots (4.15)$$

समीकरण (4.15) में ध्यान दें कि n ही प्रतिदर्श आमाप है।

समीकरण (4.14) से हमें ज्ञात होता है कि –

$$2\Sigma(Y_i - b_1 - b_2X_i)\,(-X_i) = 0$$

उपर्युक्त समीकरण में आए पदों को पुनर्व्यवस्थित कर हमें प्राप्त होता है –

$$\Sigma X_i Y_i = b_1 \Sigma X_1 + b_2 \Sigma X_i^2 \qquad \dots (4.16)$$

समीकरण (4.15) और (4.16) 'प्रसामान्य समीकरण' कहलाते हैं। हमारे पास दो अज्ञात चरों (b_1 और b_2) वाले दो ही समीकरण हैं। तदनुसार, इन दो प्रसामान्य समीकरणों को हल करके हम चर b_1 और b_2 के अद्वितीय मान ज्ञात कर सकते हैं।

प्रसामान्य समीकरण (4.15) और (4.16) हल करके हम निम्नलिखित समीकरण प्राप्त करते हैं–

$$\boxed{b_1 = \bar{Y} - b_2\bar{X} \qquad \dots (4.17)}$$

तथा

$$b_2 = \frac{n\Sigma XY - \Sigma X \Sigma Y}{n\Sigma X^2 - (\Sigma X)^2} = \frac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{\Sigma(X_i - X)^2}$$

आइए, अब चर X और Y को विचरण रूप में लेते हैं, यथा –

$$x_i = X_i - \bar{X} \qquad\qquad y_i = Y_i - \bar{Y}$$

तदनुसार,

$$\boxed{b_2 = \frac{\Sigma x_i y_i}{\Sigma x_i^2} \qquad \dots (4.18)}$$

जैसा कि हम b_2 के सूत्र से देख सकते हैं, समाश्रयण गुणांक के आकलक को विचलन रूप में लिखना अपेक्षाकृत सरल होता है। किसी भी चर के मान को उसके माध्य मान में व्यक्त करने से मानों के पदानुक्रम में कोई परिवर्तन नहीं होता क्योंकि हम प्रत्येक मान से वही नियतांक घटा रहे होते हैं। यह ध्यान देने की बात है कि चर b_1 और b_2 को प्रतिदर्श से परिकलित मात्राओं के रूप में व्यक्त किया जाता है, जो कि सूत्र द्वारा समीकरण (4.17) और (4.18) में निष्पीड़न के रूप में दिया गया है।

आकलक b_1 और b_2 का विचरण और मानक त्रुटि ज्ञात करने के सूत्रों का उल्लेख नीचे किया गया है –

$$Var(b_1) = \sigma_{b_1}^2 = \frac{\Sigma X_i^2}{n\Sigma x_i^2}\sigma^2 \qquad \dots (4.19)$$

$$SE(b_1) = \sqrt{Var(b_1)} \qquad \dots (4.20)$$

$$Var(b_2) = \sigma_{b_2}^2 = \frac{\sigma^2}{\Sigma x_i^2}$$

$$SE(b_2) = \sqrt{\text{var}}(b_2) \qquad \dots (4.21)$$

$$\hat{\sigma}^2 = \frac{\Sigma e_i^2}{n-2} = \frac{RSS}{n-2} = \frac{RSS}{d.f.} \qquad \dots (4.22)$$

अवशिष्ट (e_i) की मानक त्रुटि $= \sqrt{\hat{\sigma}^2}$ $\qquad \dots (4.23)$

समीकरणों (4.19), (4.20), (4.21), (4.22) और (4.23) में उल्लिखित सूत्र आकलित प्राचल b_1 और b_2 का विचरण और मानक त्रुटियाँ ही दर्शाते हैं।

पद $\widehat{\sigma}^2$ का मान जितना कम होगा, वास्तविक Y–मान उसके आकलित मान के उतना ही सन्निकट होगा। स्मरण करें कि किसी सामान्य रूप से बंटित चर के किसी भी रैखिक फलन को स्वयं भी सामान्य रूप से ही बंटित किया जाता है। उदाहरणार्थ, यदि चर b_1और b_2 सामान्य रूप से बंटित चर u_i के रैखिक फलन हों तो वे स्वयं भी सामान्य रूप से ही बंटित होते हैं। तदनुसार,

$$b_1 \sim N\left(\beta_1, \sigma^2_{b_1}\right) \quad \ldots (4.24)$$

$$b_2 \sim N\left(\beta_2, \sigma^2_{b_2}\right) \quad \ldots (4.25)$$

बोध प्रश्न 2

1) उपयुक्त आरेख की सहायता से त्रुटि पद और अवशिष्ट के बीच अंतर स्पष्ट करें।

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

2) सिद्ध कीजिए कि प्रतिदर्श समाश्रयण रेखा X और Y के माध्य मानों से होकर गुजरती है।

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

4.6 OLS आकलकों के बीजगणितीय गुणधर्म

साधारण न्यूनतम वर्ग (OLS) आकलक b_1 और b_2 कुछ महत्वपूर्ण गुणधर्म दर्शाते हैं, जो कि निम्नलिखित हैं –

a) साधारण न्यूनतम वर्ग (OLS) विधि से प्राप्त प्रतिदर्श समाश्रयण फलन (SRF) चर X और Y के प्रतिदर्श माध्य मानों से गुजरता है। इसका मुख्य रूप से अर्थ यह है कि बिंदु $(\overline{X}, \overline{Y})$ प्रतिदर्श समाश्रयण रेखा से होकर गुजरता है।

$$\overline{Y} = b_1 + b_2\overline{X} \quad \ldots (4.26)$$

अवशिष्टों का माध्य मान $\overline{e}$ सदैव शून्य होता है, यथा $\overline{e} = \frac{\Sigma e_i}{n} = 0$ इसका अर्थ यह है कि औसतन धनात्मक और ऋणात्मक अवशिष्ट पद एक–दूसरे को निरस्त कर देते हैं।

b) $\Sigma e_i X_i = 0 \quad \ldots (4.27)$

अवशिष्टों के गुणनफल e_i और व्याख्यात्मक चर X के मानों का योगफल शून्य होता है, यथा ये दोनों चर सहसंबंध नहीं होते हैं।

c) $\Sigma e_i \hat{Y}_i = 0 \quad \ldots (4.28)$

अवशिष्टों के गुणनफल e_i और आकलित चर $\hat{Y}_i$ का योगफल शून्य होता है, यथा $e_i\hat{Y}_i = 0$.

4.7 निर्धारण गुणांक

आइए, इस समाश्रयण मॉडल पर विचार करें –

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

समीकरण (4.7) से स्मरण करें कि

$$Y_i = \hat{Y}_i + e_i$$

यदि हम उपर्युक्त समीकरण के दोनों पक्षों से $\bar{Y}$ घटा दें तो हमें निम्नलिखित समीकरण प्राप्त होगा –

$$(Y_i - \bar{Y}) = (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i) \qquad \ldots (4.29)$$

[चूँकि $e_i = Y_i - \hat{Y}_i$]

समीकरण (4.20) में तीन पद हैं, यथा (i) $(Y_i - \bar{Y})$ जो कि चर Y_i में विचरण है; (ii) $(\hat{Y}_i - \bar{Y})$ जो कि अव्याख्यायित विचरण है; और (iii) $(Y_i - \hat{Y}_i)$ जो कि अस्पष्टीकृत या अवशिष्ट विचरण है।

चलिए, अब किसी भी चर के माध्य से विचलन को इंगित करने के लिए रोमन लिपि के छोटे अक्षरों का प्रयोग करते हैं। तदनुसार, समीकरण (4.30) को निम्नवत् लिखा जा सकता है–

$$y_i = \hat{y}_i + e_i \qquad \ldots (4.30)$$

चूँकि $\sum e_i = 0$, हमें मान $\bar{e} = 0$ प्राप्त होता है।

अतएव, हमें $\bar{Y} = \bar{\hat{Y}}$ प्राप्त होता है, यथा वास्तविक Y और आकलित Y के औसत मान एक समान ही होते हैं।

स्मरण करें कि

$$Y_i = b_1 + b_2 X_i + e_i \qquad \ldots (4.7)$$

तथा

$$\bar{Y} = b_1 + b_2 \bar{X} \qquad \ldots (4.26)$$

यदि हम समीकरण (4.7) में से समीकरण (4.26) घटा दें तो हमें निम्नलिखित समीकरण प्राप्त होगा –

$$y_i = b_2 x_i + e_i \qquad \ldots (4.31)$$

यदि समीकरण (4.31) का OLS आकलक ज्ञात करना हो तो हमें यह समीकरण प्राप्त होगा –

$$\hat{y}_i = b_2 x_i$$

अतएव,

$$y_i = \hat{y}_i + e_i \qquad \ldots (4.32)$$

आइए, अब हम समीकरण (4.32) के दोनों पक्षों के वर्ग लेते हैं और इसे प्रतिदर्श के मान में जोड़ देते हैं। अब पदों को पुनर्व्यवस्थित करने पर हमें निम्नलिखित समीकरण प्राप्त होता है –

$$\Sigma y_i^2 = \Sigma \hat{y}_i^2 + \Sigma e_i^2 \qquad \ldots (4.33)$$

अथवा, समतुल्य रूप से,

$$\Sigma y_i^2 = b_2^2 \Sigma x_i^2 + \Sigma e_i^2 \qquad \ldots (4.34)$$

अब समीकरण (4.34) को निम्नलिखित रीति से व्यक्त किया जा सकता है –

TSS = ESS + RSS … (4.35)

जहाँ TSS = वर्गों का कुल योग

ESS = वर्गों का व्याख्यायित योगफल

RSS = वर्गों का अवशिष्ट योगफल

चलिए, समीकरण (4.35) को TSS से भाग देते हैं। इससे हमें यह समीकरण प्राप्त होता है –

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS} \qquad \ldots (4.36)$$

आइए, अब समीकरण को परिभाषित करते हैं –

$$R^2 = \frac{ESS}{TSS} \qquad \ldots (4.37)$$

उक्त पद R^2 ही निर्धारण गुणांक (coefficient of determination) कहलाता है। इसे किसी भी समाश्रयण मॉडल की 'सुमानकता' को मापदंड के रूप में लिया जाता है। कुल मिलाकर 'सुमानकता' ही है जो हमें यह बताती है कि आकलित समाश्रयण रेखा वास्तविक Y–मानों के कितने अनुरूप है।

4.7.1 R^2 की संगणना का सूत्र

समीकरण (4.37) में दी गई R^2 की परिभाषा का प्रयोग करके हम समीकरण (4.36) को इस प्रकार लिख सकते हैं–

$$1 = R^2 + \frac{RSS}{TSS} = R^2 + \frac{\Sigma e_i^2}{\Sigma y_i^2} \qquad \ldots (4.40)$$

अतएव,

$$R^2 = 1 - \frac{\Sigma e_i^2}{\Sigma y_i^2} \qquad \ldots (4.38)$$

इस बात पर ध्यान दें कि पद R^2 एक रीति (ESS) द्वारा स्पष्ट की गई दूसरी रीति (TSS) का प्रतिशतता दर्शाता है। तदनुसार, यदि $R^2 = 0.75$ हो तो हम कह सकते हैं कि आश्रित चर में 75 प्रतिशत विचरण को समाश्रयण मॉडल में व्याख्यात्मक चर द्वारा स्पष्ट किया गया है। अब पद R^2 का मान या निर्धारण गुणांक 0 और 1 के बीच रहता है। इसका मुख्य कारण यह है कि यह वर्गों के कुल योग और वर्गों के व्याख्यायित योगफल का अनुपात निरूपित करता है।

आइए, अब हम पद R^2 के बीजीय गुणधर्मों को देखते हैं और उन्हीं के अनुसार इसकी व्याख्या करते हैं। जब $R^2 = 0$ हो तो हमें ESS = 1 प्राप्त होता है। यह इंगित करता है कि वर्गों के व्याख्यायित योगफल द्वारा आश्रित चर में विचरण के किसी भी अनुपात की व्याख्या नहीं की गई है। यदि $R^2 = 1$ हो तो प्रतिदर्श समाश्रयण एकदम उपयुक्त होता है। यदि $R^2 = 1$ हो तो सभी प्रेक्षण आकलित समाश्रयण रेखा पर स्थित होंगे। पद R^2 के किसी भी उच्चतर मान का अर्थ होगा कि वह समाश्रयण मॉडल अपेक्षाकृत अधिक उपयुक्त है।

4.7.2 सुमानकता हेतु F–प्रतिदर्शज

किसी भी समाश्रयण मॉडल की सांख्यिकीय सार्थकता का परीक्षण F–प्रतिदर्शज द्वारा किया जाता है जबकि t–परीक्षण का प्रयोग कर हम अपने समाश्रयण मॉडल के किसी विशिष्ट प्राचल की सांख्यिकीय सार्थकता का परीक्षण कर सकते हैं। उदाहरण के लिए, शून्य–स्तरीय परिकल्पना (null hypothesis) $H_0: \beta_2 = 0$ का अर्थ होगा कि समष्टि में Y और X के बीच कोई संबंध नहीं है। इसके अलावा, F–प्रतिदर्शज का प्रयोग कर हम शून्य–स्तरीय परिकल्पना का यह परीक्षण कर सकते हैं कि मॉडल में सभी प्राचल शून्य होते हैं। अतएव, हम सुमानकता जाँचने के लिए हम F–प्रतिदर्शज का ही प्रयोग करते हैं।

सुमानकता हेतु F–प्रतिदर्शज को निम्नलिखित समीकरण से दर्शाया जाता है –

$$F = \frac{ESS/(k-1)}{RSS/(n-k)} \qquad \ldots (4.39)$$

जहाँ k समाश्रयण समीकरण में प्राचलों की संख्या है और n प्रतिदर्श आमाप है।

4.7.3 चर F और R^2 के बीच संबंध

समीकरण (4.39) से हम जानते हैं कि $F = \frac{ESS/(k-1)}{RSS/(n-k)}$ होता है। समीकरण (4.39) से हम जानते हैं कि

$$F = \frac{ESS/TSS/(k-1)}{RSS/TSS/(n-k)} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \qquad \ldots (4.40)$$

यह ध्यान देने की बात है कि F–प्रतिदर्शज पद R^2 का एक वर्धमान फलन होता है। पद R^2 के मान में वृद्धि का अर्थ होता है – अंश में वृद्धि और हर में कमी।

चलिए, अब हम समीकरण (4.41) में प्राप्त F–प्रतिदर्शज की व्याख्या की व्याख्या करते हैं। किसी भी आँकड़ा समुच्चय में समीकरण (4.41) लागू करने से प्राप्त मान F का परिकलित मान अथवा F–परिकलित कहलाता है। हम इस मान की तुलना पुस्तक के अंत में दिए गए F के तालिकाबद्ध मान या क्रांतिक मान से करते हैं। तुलना के उद्देश्य से स्वतंत्रता की कोटियाँ $((k-1), (n-k))$ प्रयोग की जाती हैं।

यदि F–आकलित F–क्रांतिक से अधिक होता है तो हम शून्य–स्तरीय परिकल्पना $H_0: \beta_2 = 0$ को अस्वीकार कर देते हैं। इस कथन का एक निहितार्थ यह भी है कि स्वतंत्र चर आश्रित चर की व्याख्या करते हैं। दूसरे शब्दों में, चर Y और X के बीच सांख्यिकीय रूप से सार्थक संबंध विद्यमान होता है। दूसरी ओर, यदि F–आकलित F–क्रांतिक से कम होता है तो हम शून्य–स्तरीय $H_0: \beta_2 = 0$ परिकल्पना को स्वीकार कर लेते हैं। इस प्रकार, चर Y और X के बीच कोई सार्थक संबंध नहीं होता है।

4.7.4 F और t^2 के बीच संबंध

समाश्रयण मॉडल में F–प्रतिदर्शज और t–प्रतिदर्शज के बीच होता संबंध है। मान लीजिए कि व्याख्यात्मक चरों की संख्या $k = 2$ है। तदनुसार,

$$F = \frac{ESS/(k-1)}{RSS/(n-2)}$$

किसी भी दो–चर मॉडल के लिए,

$$F = \frac{ESS/(2-1)}{RSS/(n-2)} = \frac{ESS}{RSS/(n-2)} \qquad \ldots (4.41)$$

हमें ज्ञात है कि ESS = $\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$ और $RSS = \sum_{i=1}^{n} e_i^2$

अतएव,

$$F = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} e_i^2/(n-2)} = \frac{\sum_{i=1}^{n}([b_1 + b_2 X_i] - [b_1 + b_2 \bar{X}])^2}{\hat{\sigma}^2} \qquad \ldots (4.42)$$

त्रुटि विचरण का आकलन $= \hat{\sigma}^2 = \frac{RSS}{n-k} = \frac{\Sigma e_i^2}{n-2}$ … (4.43)

$$F = \frac{1}{\hat{\sigma}^2} \cdot \sum_{i=1}^{n} b_2^2 (X_i - \overline{X})^2 \quad \dots (4.44)$$

हमें ज्ञात है कि

$$var(b_2) = \frac{\hat{\sigma}^2}{\Sigma x_i^2}$$

समीकरण (4.43) को समीकरण (4.44) में प्रतिस्थापित करने पर हमें निम्नलिखित समीकरण प्राप्त होता है –

$$F = \frac{b_2^2}{\hat{\sigma}^2} = \frac{b_2^2}{var(b_2)} = \frac{b_2^2}{[SE(b_2)]^2} = t^2 \quad \dots (4.45)$$

अतएव, F–प्रतिदर्शज t–प्रतिदर्शज के वर्ग के बराबर होता है, यथा $(F = t^2)$। बहरहाल, उपर्युक्त परिणाम केवल दो–चर मॉडल के लिए ही सही होता है। यदि समाश्रयण मॉडल में व्याख्यात्मक चरों की संख्या बढ़ जाती है तो उपर्युक्त परिणाम सिद्ध नहीं हो सकता है।

बोध प्रश्न 3

1) क्या निर्धारण के गुणांक के आधार पर F–परीक्षण करना संभव होता है? कारण स्पष्ट करें।

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

2) क्या निर्धारण का गुणांक 1 से अधिक हो सकता है? कारण स्पष्ट करें।

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

4.8 सार–संक्षेप

इस इकाई में हमने पारंपरिक रैखिक समाश्रयण मॉडल के विषय में चर्चा की, जो कि कुछ अवधारणाओं पर आधारित है। हमने समष्टि समाश्रयण फलन और प्रतिदर्श समाश्रयण फलन के बीच स्पष्ट अंतर किया। हमने स्पष्ट किया कि समाश्रयण समीकरण में प्रसामान्य त्रुटि पद क्यों जोड़ा जाता है। हमने पारंपरिक समाश्रयण मॉडल की प्रत्येक अवधारणा का अर्थ भी स्पष्ट किया। इसमें समाश्रयण मॉडल के OLS आकलक प्राप्त करने की प्रक्रिया सविस्तार दी गई है।

इकाई के अंत में सुमानकता संबंधी अवधारणा और R^2 की अवधारणा पर प्रकाश डाला गया है।

4.9 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1) किसी भी समाश्रयण मॉडल को क्रियान्वित करने के उद्देश्य इस प्रकार हो सकते हैं–

- स्वतंत्र चरों के मान दिए गए हों तो आश्रित चर के माध्य या औसत मान का आकलन करना।
- अंतर्निहित आर्थिक सिद्धांत के संबंध में परिकल्पनाओं का परीक्षण करना। उदाहरण के लिए, आप इस परिकल्पना का परीक्षण कर सकते हैं कि माँग की मूल्य लोचता (–)1 है।
- स्वतंत्र चर के मान को देखते हुए आश्रित चर के माध्य मान का पूर्वानुमानन अथवा भविष्यवाणी करना।

2) चर X और Y के बीच संबंध अपनी प्रकृति में प्रसंभाव्य होता है। समाश्रयण समीकरण में एक त्रुटि पद जोड़ा जाता है। यादृच्छिक त्रुटि पद को शामिल करने से अपेक्षित मूल्य और आश्रित चर के वास्तविक मूल्य के बीच अंतर उत्पन्न होता है।

3) समाश्रयण मॉडल में त्रुटि पद को शामिल करने के तीन कारण होते हैं। विस्तृत विवरण के लिए पाठांश 4.2.2 देखें।

बोध प्रश्न 2

1) पाठांश 4.3 का अध्ययन करें। आपको चित्र 4.3 का प्रयोग कर त्रुटि पद और अवशिष्ट के बीच के अंतर को स्पष्ट करना चाहिए।

2) साधारण न्यूनतम वर्ग (OLS) विधि में हम Σe_i^2 के आंशिक अवकलों को शून्य के बराबर करके कम करते हैं। शर्त $\frac{\partial \Sigma e_i^2}{\partial b_1} = 0$ हमें पहला सामान्य यह समीकरण $Y_i = nb_1 + b_2 \Sigma X_i$ देती है। यदि हम इस समीकरण को प्रतिदर्श आमाप n से विभाजित कर देते हैं तो हमें समीकरण $\bar{Y} = b_1 + b_2\bar{X}$ होता है। इस प्रकार, आकलित समाश्रयण बिंदु $\bar{X}, \bar{Y}$ से होकर गुजरता है।

बोध प्रश्न 3

1) हाँ, हम R^2 मान के आधार पर F–परीक्षण कर सकते हैं। समीकरण (4.40) का अध्ययन करें।

2) चर R^2 का मान या निर्धारण का गुणांक 0 और 1 के बीच स्थित होता है। इसका मुख्य कारण यह है कि यह ESS और TSS के अनुपात को दर्शाता है। यह चर Y में विचरण के अनुपात को इंगित करता है जिसे व्याख्यात्मक चर द्वारा स्पष्ट किया गया है। किसी समीकरण का अंश ESS वर्गों के कुल योग (TSS) से अधिक नहीं हो सकता। अतएव, R^2 का मान 1 से अधिक नहीं हो सकता।

इकाई 5 सरल रैखिक समाश्रयण मॉडल : अनुमिति*

इकाई की रूपरेखा

5.0 उद्देश्य

5.1 विषय–प्रवेश

5.2 परिकल्पना का परीक्षण

5.3 विश्वास्यता अंतराल

5.4 सार्थकता का परीक्षण

5.5 विचरण का विश्लेषण (ANOVA)

5.6 गॉस–मार्कोव प्रमेय

5.7 पूर्वानुमान

5.7.1 व्यष्टिक पूर्वानुमान

5.7.2 माध्य पूर्वानुमान

5.8 सार–संक्षेप

5.9 बोध प्रश्नों के उत्तर अथवा संकेत

5.0 उद्देश्य

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

- परिकल्पना के परीक्षण संबंधी अवधारणा की व्याख्या कर सकें;
- किसी सरल रैखिक समाश्रयण मॉडल में समाश्रयण गुणांक के लिए विश्वास्यता अंतराल अवकलित कर सकें;
- आकलित समाश्रयण गुणांक पर परिकल्पना के परीक्षण हेतु 'सार्थकता के परीक्षण' संबंधी उपागम की व्याख्या कर सकें;
- विचरण के विश्लेषण (ANOVA) संबंधी अवधारणा का वर्णन कर सकें;
- गॉस–मार्कोव प्रमेय उसके गुणधर्मों सहित स्पष्ट कर सकें; तथा
- किसी सरल रैखिक समाश्रयण मॉडल में Y के आकलित मान के लिए विश्वास्यता अंतराल अवकलित कर सकें।

5.1 विषय–प्रवेश

इस पाठ्यक्रम की इकाई 4 में हमने प्राचलों के मानों के आकलन की प्रक्रिया पर चर्चा की थी। अब इस इकाई में हम प्राप्त प्राचलों के आकलनों के आधार पर निष्कर्ष निकालने की विधि पर ध्यान केंद्रित करेंगे। यहाँ हम केवल एक स्वतंत्र चर के साथ एक सरल रैखिक समाश्रयण मॉडल पर विचार करेंगे। इसका अर्थ है कि हमारे पास स्वतंत्र चर से सहबद्ध एक समाश्रयण गुणांक और एक ही अवरोधन पद होगा। इकाई का आरंभ हम 'परिकल्पना परीक्षण' के मूल तत्वों को संक्षेप में दोहरा कर करेंगे।

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

5.2 परिकल्पना का परीक्षण

परिकल्पना के परीक्षण से हमारा अभिप्राय यह आकलन करना होता है कि हमारा प्रेक्षण अथवा अनुमिति उक्त परिकल्पना के अनुकूल है अथवा नहीं। अनुकूलता या सुसंगतता शब्द का अर्थ परिकल्पित मान के ''पर्याप्त रूप से निकट'' होना माना जाता है। यह आगे इंगित करता है कि हम कथित परिकल्पना को निरस्त नहीं करते हैं। इस कथित परिकल्पना को 'शून्य–स्तरीय परिकल्पना' (null hypothesis) भी कहा जाता है और इसे H_0 से दर्शाया जाता है।

उक्त निराकरणीय परिकल्पना का परीक्षण आमतौर पर 'वैकल्पिक परिकल्पना' (alternative hypothesis) की तुलना में किया जाता है, जिसे 'अनुरक्षित परिकल्पना' के नाम से भी जाना जाता है। दूसरी ओर, वैकल्पिक परिकल्पना को H_1 से निरूपित किया जाता है। उदाहरण के लिए, मान लीजिए कि दिया गया समष्टि समाश्रयण फलन निम्नलिखित समीकरण से दर्शाया गया है–

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \text{... (5.1)}$$

जहाँ X_i व्यक्तिगत प्रयोज्य आय (PDI) है और Y_i व्यय है। अब शून्य–स्तरीय परिकल्पना निम्नवत् होगी–

$$H_0 : \beta_2 = 0 \qquad \text{... (5.2)}$$

जबकि वैकल्पिक परिकल्पना निम्नवत् होगी –

$$H_1 : \beta_2 \neq 0 \qquad \text{... (5.3)}$$

हम जानबूझकर शून्य–स्तरीय परिकल्पना को 'शून्य' पर नियत करते हैं ताकि यह पता लगाया जा सके कि चर Y चर X से बिल्कुल भी संबंधित है अथवा नहीं। यदि चर X वास्तव में मॉडल से संबंधित होगा तो हम वैकल्पिक परिकल्पना $H_1 \Rightarrow \beta_2 \neq 0$ के पक्ष में शून्य–निराकरणीय परिकल्पना H_0 को निरस्त किए जाने की पूरी आशा करेंगे।

वैकल्पिक परिकल्पना का अर्थ है कि समाश्रयण गुणांक शून्य से भिन्न होता है। यह धनात्मक हो सकता है या फिर यह ऋणात्मक भी हो सकता है। इसी प्रकार, निम्नवत् शून्य–स्तरीय परिकल्पना की स्थापना करके वास्तविक समष्टि अवरोधन का परीक्षण किया जा सकता है–

$$H_1 : \beta_1 = 0 \qquad \text{… (5.4)}$$

जबकि वैकल्पिक परिकल्पना निम्नवत् होगी–

$$H_1 : \beta_1 \neq 0 \qquad \text{… (5.5)}$$

उक्त शून्य–स्तरीय परिकल्पना के अनुसार वास्तविक समष्टि अवरोधन शून्य के बराबर होता है, जबकि वैकल्पिक परिकल्पना के अनुसार वह शून्य के बराबर नहीं होता है। बहरहाल, दोनों ही शून्य–स्तरीय परिकल्पनाओं के मामले में, यानी वास्तविक समष्टि प्राचल अथवा समाश्रयण और अवरोधन के लिए, जैसा कि कहा गया है, शून्य–स्तरीय परिकल्पना एक 'सरल परिकल्पना' सिद्ध होती है।

दूसरी ओर, वैकल्पिक परिकल्पना सम्मिश्रित होती है। इसे **द्विपक्षीय परिकल्पना** (two-sided hypothesis) के रूप में भी जाना जाता है। इस प्रकार की दो–पक्षीय वैकल्पिक परिकल्पना इस तथ्य को दर्शाती है कि हमारे पास उस दिशा के विषय में कोई सुदृढ़ अग्रिम अथवा सैद्धांतिक प्रत्याशा नहीं होती है जिसमें वैकल्पिक परिकल्पना शून्य–स्तरीय

परिकल्पना से इतर जा सके। बहरहाल, जब हमारे पास पिछले कुछ शोध या अनुभवजन्य कार्यों के आधार पर कोई सुदृढ़ अग्रिम अथवा सैद्धांतिक प्रत्याशाएँ होती हैं तो वैकल्पिक परिकल्पना दो–पक्षीय की बजाय एकपक्षीय अथवा एकदिशीय हो सकती है। उदाहरण के लिए, यदि हम सुनिश्चित हैं कि समाश्रयण गुणांक का वास्तविक समष्टि मान धनात्मक है तो उक्त दोनों परिकल्पनाओं को व्यक्त करने का सबसे अच्छा तरीका निम्नवत् होगा–

$$H_0: \beta_2 = 0$$

$$H_1: \beta_2 > 0$$

चलिए, समष्टि अर्थशास्त्र से एक उदाहरण लेते हैं। अभिभावी आर्थिक सिद्धांत के अनुसार, उपभोग करने की सीमांत प्रवृत्ति सकारात्मक होती है। इसका अर्थ है कि समाश्रयण गुणांक धनात्मक होता है। अब मान लीजिए कि दिए गए समष्टि समाश्रयण फलन का आकलन साधारण न्यूनतम वर्ग (OLS) आकलन को अपना कर किसी प्रतिदर्श समाश्रयण का प्रयोग करके किया जाता है। चलिए, हम यह भी मान लेते हैं कि प्रतिदर्श समाश्रयण के परिणाम आकलित समाश्रयण गुणांक का मान $b_2 = 0.0814$ के रूप में देते हैं। यह संख्यात्मक मान प्रतिदर्श से प्रतिदर्श में बदलता जाता है। हमें ज्ञात है कि β_2 सामान्य बंटन का अनुसरण करता है, यथा –

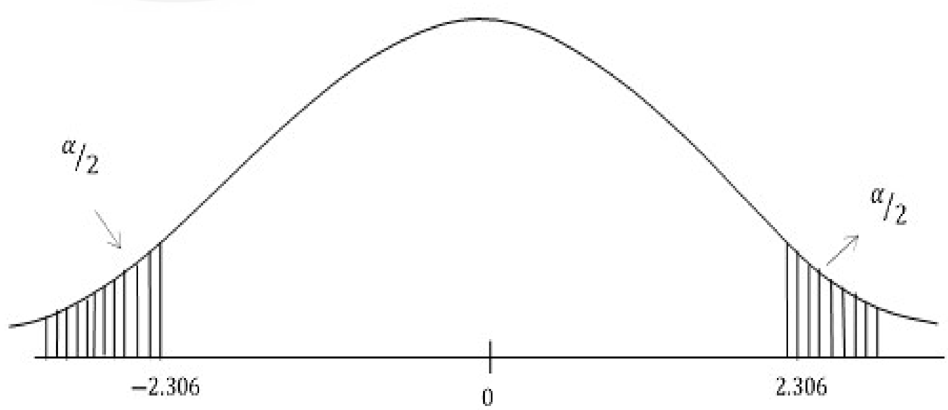
$$b_2 \sim N\left(\beta_2, \frac{\sigma^2}{\Sigma x_i^2}\right)$$

इस शून्य–स्तरीय परिकल्पना का परीक्षण करने के दो तरीके हैं कि वास्तविक समष्टि समाश्रयण गुणांक शून्य के बराबर होता है। इस इकाई के अगले दो पाठांशों में समाश्रयण प्राचलों की परिकल्पना के परीक्षण संबंधी दो विधियों का वर्णन किया जाएगा।

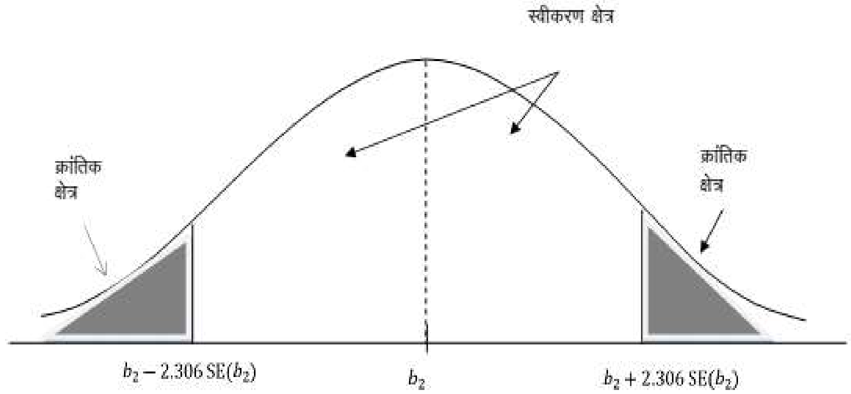
5.3 विश्वास्यता अंतराल

इस पाठांश में हम ऊपर दिए गए समीकरण (5.1) में समाश्रयण प्राचल के लिए विश्वास्यता अंतराल अवकलित करेंगे। यह ध्यान देने की बात है कि विश्वास्यता अंतराल उपागम परिकल्पना के परीक्षण की ही एक विधि है। ऐसा इसलिए है कि वह इस प्रायिकता को इंगित करता है कि समष्टि प्राचल दी गई तालिका से लिए जाने वाले क्रांतिक मानों की शृंखला में ही शामिल होगा।

हम दो अवधारणाएँ लेकर चलते हैं, यथा – (i) टाइप–I त्रुटि करने की प्रायिकता पर सार्थकता का स्तर α, स्तर 5 प्रतिशत पर तय किया जाता है, और (ii) वैकल्पिक परिकल्पना द्विपक्षीय होती है।



चित्र 5.1: *t*–बंटन



चित्र 5.2: β_2 हेतु विश्वास्यता अंतराल

दी गई t–तालिका से हम पाते हैं कि $(\alpha = 5\%)$ पर $(n-k)$ स्वतंत्रता की कोटि (d.f) में चर t का क्रांतिक मान निम्नवत् है–

$$P(-2.306 \le t \le 2.306) = 0.95 \qquad ...(5.6)$$

अब चर 't' का प्रतिस्थापन कर समीकरण (5.6) को निम्नवत् लिखा जा सकता है

$$P\left(2.306 \le \frac{b_2-\beta_2}{\hat{\sigma}/\sqrt{\Sigma x_i^2}} \le 2.306\right) = 0.95 \qquad ...(5.7)$$

अतएव, यह प्रायिकता कि चर t का मान सीमा (–2.306, +2.306) के भीतर ही है, 0.95 अथवा 95% है। यही क्रांतिक t–मान कहलाते हैं। अब चर t का मान समीकरण (5.6) में प्रतिस्थापित कर और पदों को समीकरण (5.7) में पुनर्व्यवस्थित कर हमें निम्नलिखित समीकरण प्राप्त होता है–

$$\mathrm{P}[(\mathrm{b}_2 - 2.306.\mathrm{SE}(\mathrm{b}_2) \le \beta_2 \le \mathrm{b}_2 + 2.306\mathrm{SE}(\mathrm{b}_2))] = 0.95$$

उपर्युक्त समीकरण β_2 के लिए एक 95% विश्वास्यता अंतराल प्रदान करता है। इस तरह के विश्वास्यता अंतराल को स्वीकरण के क्षेत्र (H_0 के लिए) के रूप में जाना जाता है और विश्वास्यता अंतराल के बाहर के क्षेत्र को अस्वीकरण के क्षेत्र [(H_0) के लिए] के रूप में जाना जाता है। यदि इस अंतराल β_2 में का मान शामिल हो तो हम परिकल्पना को अस्वीकार नहीं करते हैं, परंतु यदि यह मान विश्वास्यता अंतराल से बाहर हो तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर देते हैं।

बोध प्रश्न 1

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

1) किसी सरल और किसी सम्मिश्रित परिकल्पना के बीच अंतर स्पष्ट करें।

....

....

....

....

2) शून्य–स्तरीय परिकल्पना सरल और सम्मिश्रित परिकल्पना का संसूचक होती है। क्या यह कथन सही है? औचित्य प्रतिपादन करें।

....

....

....

....

3) किसी 'विश्वास्यता अंतराल' से आप क्या समझते हैं?

....

....

....

....

4) हम ऐसा क्यों कहते हैं कि अंतराल में वास्तविक समष्टि प्राचल का परिकल्पित मान निहित होता है?

....

....

....

....

5.4 सार्थकता का परीक्षण

सार्थकता का परीक्षण उपागम परिकल्पना के परीक्षण का एक अन्य तरीका है। चर H_0 को स्वीकार अथवा अस्वीकार करने का निर्णय t–परीक्षण के मान के आधार पर किया जाता है। इसका संगणन प्रतिदर्श आँकड़ों से लिए गए किसी प्रतिदर्शज से किया जाता है, यथा–

$$t = \frac{b_2 - \beta_2}{SE(b_2)} \quad \dots (5.8)$$

समीकरण (5.8) $(n - k)$ df वाले t–बंटन का अनुसरण करता है। यहाँ हम जिस शून्य–स्तरीय परिकल्पना का परीक्षण कर रहे हैं, निम्नवत् है–

$$H_0 : \beta_2 = \beta_2^* \quad \dots (5.9)$$

यह ध्यान देने की बात है कि चर β_2^* प्रतिदर्श प्रेक्षणों के आधार पर संगणित चर β_2 का कोई विशिष्ट संख्यात्मक मान होता है। तदनुसार, परीक्षण–प्रतिदर्शज 't' का परिकलित मान निम्नवत् होगा–

$$t = \frac{b_2 - \beta_2^*}{SE(b_2)} \quad \dots (5.10)$$

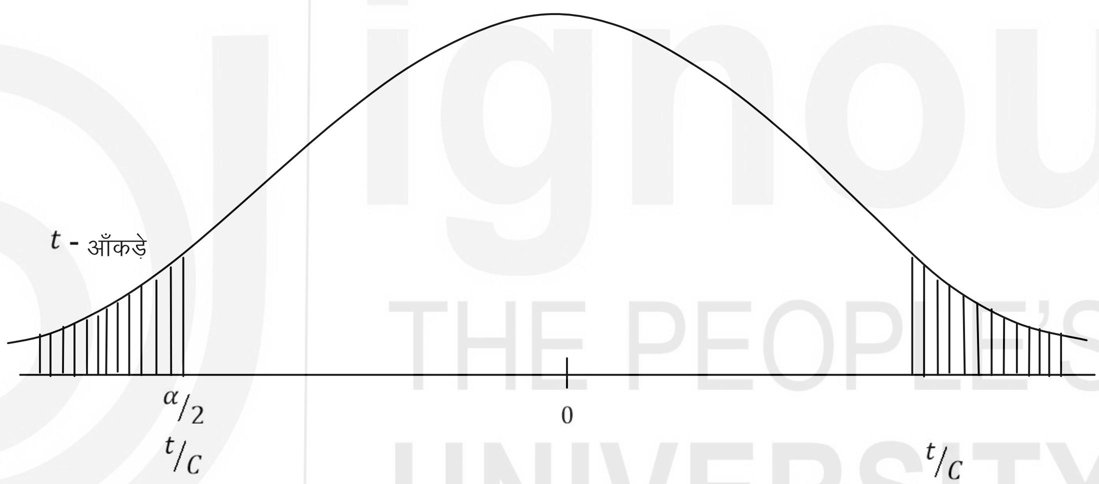
= [(आकलित मान) – (परिकल्पित मान) ÷ (आकलक की मानक त्रुटि)]

इसका संगणन प्रतिदर्श आँकड़ों से किया जा सकता है क्योंकि सभी मान उपलब्ध हैं। समीकरण (5.10) से संगणित t–मान $(n – k)$ स्वतंत्रता की कोटि (d.f) वाले t–बंटन का

अनुसरण करता है। इस परीक्षण प्रक्रिया को ही t–परीक्षण कहा जाता है। चित्र 5.3 अस्वीकरण के क्षेत्र और स्वीकरण के क्षेत्र को दर्शाता है। परीक्षण के परिणाम पर निर्णय लेने की एक विधि तालिका मान (जिसे 'क्रांतिक मान' भी कहा जाता है) के साथ संगणित मान की तुलना करना भी है।

यदि चर t का संगणित मान t के क्रांतिक मान से अधिक हो तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर देते हैं। इसका अर्थ यह है कि हम ऐसी परिकल्पना को निरस्त कर रहे हैं जिसका वास्तविक समष्टि प्राचल, अथवा समाश्रयण गुणांक, शून्य है। इसका निहितार्थ यह हुआ कि व्याख्यात्मक चर आश्रित चर के निर्धारण में एक सार्थक भूमिका निभाता है।

दूसरी ओर, यदि परिकलित t–मान t के क्रांतिक मान से कम हो तो हम इस शून्य–स्तरीय परिकल्पना को अस्वीकार नहीं करते हैं कि समष्टि प्राचल (अथवा समाश्रयण गुणांक) का वास्तविक मान शून्य होता है। शून्य–स्तरीय परिकल्पना को अस्वीकार न करने का अर्थ है कि समाश्रयण गुणांक का मान शून्य है और व्याख्यात्मक चर आश्रित चर को निर्धारित करने में कोई सार्थक भूमिका नहीं निभाता है।



चित्र 5.3: सार्थकता का परीक्षण

वर्तमान समय में, जब समाश्रयण के परिणाम कंप्यूटर से प्राप्त किए जाते हैं, हम आमतौर पर संगणित आँकड़ों के लिए p–मान प्राप्त करते हैं। यह p–मान इस संभावना को इंगित करता है कि शून्य–स्तरीय परिकल्पना सत्य हो सकती है। यदि $p < 0.05$ हो तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर देते हैं और वैकल्पिक परिकल्पना को स्वीकार कर लेते हैं। दूसरी ओर, यदि $p > 0.05$ हो तो हम शून्य–स्तरीय परिकल्पना को स्वीकार कर लेते हैं।

इसका अर्थ है कि हम अपने परीक्षण परिणाम को सार्थकता के 5 प्रतिशत स्तर पर आधारित करते हैं। इसका अर्थ यह भी है कि 100 स्वतंत्र प्रतिदर्शों में से 95 में हमारा परीक्षण परिणाम एक समान होगा। दूसरे शब्दों में, 100 में से 5 उदाहरणों में हम किसी त्रुटिपूर्ण निष्कर्ष पर आ सकते हैं।

5.5 विचरण का विश्लेषण (ANOVA)

विचरण का विश्लेषण (ANOVA) एक सांख्यिकीय उपकरण है, जिसका उपयोग अनेक कारकों के कारण उत्पन्न विचरणों के लिए दिए गए आँकड़ों का विश्लेषण करने के लिए

किया जाता है। इन कारकों को दो भागों में बाँटा गया है – एक को नियतात्मक (अथवा सुव्यवस्थित) भाग कहा जाता है और दूसरे को यादृच्छिक भाग कहा जाता है। विचरण का विश्लेषण करने की यह विधि वर्ष 1918 में रोनाल्ड फिशर द्वारा विकसित की गई थी। इसीलिए इसे 'फिशर का विचरण विश्लेषण' भी कहा जाता है।

उक्त एनोवा (ANOVA) विधि आँकड़ों में देखे गए विचरण को विभिन्न घटकों में श्रेणीबद्ध करती है। इसका उपयोग समाश्रयण विश्लेषण में आश्रित चर पर स्वतंत्र चर के प्रभाव को निर्धारित करने के लिए किया जाता है। इस समाश्रयण विश्लेषण में एनोवा किसी भी समाश्रयण के भीतर विचरणशीलता की पहचान करता है। यह ध्यान देने की बात है कि आश्रित चर की कुल विचरणशीलता को दो भागों में निम्नानुसार व्यक्त किया जा सकता है–

$$(Y_i - \bar{Y}) = (Y_i - \hat{Y}_i) + (\hat{Y}_i - \bar{Y}) = (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i) \qquad \text{... (5.11)}$$

समीकरण (5.11) आश्रित चर Y में कुल विचरण को दो भागों में वितरित कर देता है, यथा माध्य में विचरण और शेष मान। इस समीकरण में प्रत्येक पद का वर्ग करने और सभी n प्रेक्षणों को जोड़ने पर हमें निम्नलिखित समीकरण प्राप्त होता है–

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2 \qquad \text{... (5.12)}$$

उपर्युक्त समीकरण को इस प्रकार लिखा जा सकता हैः TSS = ESS + RSS जहाँ TSS वर्गों का कुल योग है, ESS वर्गों का प्रतिपादित योगफल है और RSS वर्गों का शेष योगफल है।

उक्त RSS को 'त्रुटि के कारण वर्गों का योगफल' भी कहा जाता है। अनुपात ESS / TSS को निर्धारण के गुणांक R^2 के रूप में परिभाषित किया गया है। पद R^2 समाश्रयण मॉडल द्वारा स्पष्ट किए गए वर्गों के कुलयोग के अनुपात को इंगित करता है।

कोई भी एनोवा विश्लेषण एक तालिका (तालिका 5.1) की मदद से किया जाता है। विचरण के विश्लेषण की ऐसी तालिका से F–आँकड़े संगणित किए जा सकते हैं, यथा ESS/RSS मान। इन F–आँकड़ों का प्रयोग मॉडल के सार्थकता के समग्र स्तर का परीक्षण करने के लिए किया जाता है। एनोवा का प्रयोग कर समग्र सार्थकता के परीक्षण के लिए शून्य–स्तरीय परिकल्पना और वैकल्पिक परिकल्पना निम्नलिखित पदों से दर्शाई जाती हैं –

H_0: समाश्रयण गुणांक शून्य है

H_1: समाश्रयण गुणांक शून्य के बराबर नहीं है।

तालिका 5.1: एक विशिष्ट एनोवा तालिका का प्रारूप

स्रोत	स्वतंत्रता की कोटि (df)	वर्गों का योगफल	माध्य वर्ग	F-आँकड़े = ESS /RSS
मॉडल	1	$\sum(\hat{Y}_i - \bar{Y})^2$	ESS / df	
त्रुटि	*n-2*	$\sum(Y_i - \hat{Y}_i)^2$	RSS / df	
कुल योग	*n-1*	$\sum(Y_i - \bar{Y})^2$	TSS / df	

$F = \frac{ESS/(k-1)}{RSS/(n-k)}$ प्रेक्षित मान दर्शाता है।

स्वतंत्रता की $(k - 1)$ और $(n - k)$ कोटियों पर F–क्रांतिक मान सांख्यिकीय तालिका से ज्ञात किया जा सकता है।

जब F का परिकलित मान F–क्रांतिक मान से अधिक ($F > F$–क्रांतिक) हो तो शून्य–स्तरीय परिकल्पना अस्वीकृत हो जाती है। चूँकि वैकल्पिक परिकल्पना स्वीकार कर ली जाती है, निष्कर्ष यह है कि व्याख्यात्मक चर आश्रित चर को निर्धारित करने में महत्वपूर्ण भूमिका निभाता है। इसी प्रकार, जब जब F का परिकलित मान F–क्रांतिक मान से कम ($F < F$–क्रांतिक) हो तो शून्य–स्तरीय परिकल्पना अस्वीकृत नहीं की जाती है। इस स्थिति में, यह परिकल्पना स्वीकार कर ली जाती है कि व्याख्यात्मक चर आश्रित चर को निर्धारित करने में कोई भूमिका नही ंनिभाता है।

पुनः, यहाँ भी, हम अपने निष्कर्ष को p–मान के पर आधारित मान सकते हैं। इसका अर्थ है कि यदि $p < 0.05$ हो तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर देंगे।

बोध प्रश्न 2

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

1) परिकल्पना परीक्षण के लिए 'सार्थकता का परीक्षण' उपागम से आप क्या समझते हैं?

....

....

....

....

2) 'सार्थकता का स्तर' क्या इंगित करता है?

....

....

....

....

3) किसी एनोवा (ANOVA) से हमारे किस उद्देश्य की पूर्ति होती है?

....

....

....

....

4) किसी समाश्रयण मॉडल में t–परीक्षण और F–परीक्षण के बीच अंतर स्पष्ट करें।

....

....

....

....

5.6 गॉस–मार्कोव प्रमेय

यह एक महत्वपूर्ण प्रमेय है, जो हमें वह अवस्था प्रदान करती है जिसके तहत न्यूनतम वर्ग आकलक सर्वोत्तम आकलक कहलाता है। जब पारंपरिक रैखिक समाश्रयण मॉडल की अवधारणाओं का उल्लंघन नहीं किया जाता है तो न्यूनतम वर्ग आकलक कुछ इष्टतम गुणधर्मों को संतुष्ट करता है। इन गुणधर्मों को ही गॉस–मार्कोव प्रमेय में सारबद्ध किया गया है, जिसे निम्नवत् व्यक्त किया गया है–

गॉस–मार्कोव प्रमेय : पारंपरिक रैखिक समाश्रयण मॉडल की अवधारणाओं को ध्यान में रखते हुए, न्यूनतम वर्ग आकलक, सभी अनभिनत रैखिक आकलकों के वर्ग में, न्यूनतम विचरण दर्शाते हैं, यथा वे सर्वोत्तम रैखिक अनभिनत आकलक (BLUE) सिद्ध होते हैं। इन आकलकों (BLUE) की विशिष्टता दर्शाती है कि साधारण न्यूनतम वर्ग (OLS) विधि से प्राप्त आकलक में निम्नलिखित गुणधर्म देखे जाते हैं–

a) यह रैखिक होता है अर्थात् यह आकलक एक यादृच्छिक चर का एक रैखिक फलन होता है (उदाहरणार्थ, समाश्रयण मॉडल में आश्रित चर Y)।

b) यह अनभिनत होता है अर्थात् इसका औसत अथवा प्रत्याशित मान वास्तविक मान के बराबर $[E(b_2) = \beta_2]$ होता है।

c) इस प्रकार के सभी रैखिक अनभिनत आकलकों के वर्ग में इसका न्यूनतम विचरण होता है। दूसरे शब्दों में, न्यूनतम विचरण वाला ऐसा कोई भी आकलक एक दक्ष आकलक कहलाता है।

इस प्रकार, समाश्रयण के संदर्भ में, साधारण–न्यूनतम–वर्ग (OLS) आकलक सर्वोत्तम रैखिक अनभिनत आकलक (BLUE) होते हैं। यही गॉस–मार्कोव प्रमेय का सार है।

5.7 पूर्वानुमान

अब तक हमने समष्टि प्राचलों के आकलन के बारे में ही चर्चा की है। इस चर्चा के दौरान द्विचर मॉडल में हमने अवरोधन (β_1) और समाश्रयण (β_2) प्राचलों के साधारण न्यूनतम वर्ग (OLS) आकलक अवकलित करना सीखा। पूर्वानुमान स्वतंत्र चर के किसी विशिष्ट मान पर आश्रित मान के आकलन को इंगित करता है। दूसरे शब्दों में, हम X के किसी ज्ञात मान के अनुरूप Y के मान का पूर्वानुमान व्यक्त करने के लिए आकलित समाश्रयण मॉडल प्रयोग करते हैं।

पूर्वानुमानन हमारे लिए दो कारणों से महत्वपूर्ण होता है– पहला, यह हमें नीति–निर्माण में मदद करता है। अर्थमितीय मॉडल के आधार पर हम आश्रित चर पर व्याख्यात्मक चर में परिवर्तनों के प्रभाव का पता लगा सकते हैं। दूसरा, हम अपने आकलित मॉडल की सुदृढ़ता का पता लगा सकत हैं। यदि हमारा अर्थमितीय मॉडल सही है तो पूर्वानुमानित मान और आश्रित चर के वास्तविक मान के बीच की त्रुटि निरर्थक ही होगी।

पूर्वानुमान दो प्रकार का हो सकता है, जैसा कि नीचे बताया गया है –

5.7.1 व्यष्टिक पूर्वानुमान

यदि हम व्याख्यात्मक चर के किसी विशिष्ट मान के अनुरूप आश्रित चर के एक व्यष्टिक मान का पूर्वानुमान व्यक्त करते हैं तो हमें व्यष्टिक पूर्वानुमान प्राप्त होता है। चलिए, चर X का एक विशिष्ट मान, माना $X = X_0$ लेते हैं।

उक्त मान $X = X_0$ पर Y का व्यष्टिक पूर्वानुमान निम्नानुसार ज्ञात किया जाएगा–

$$Y_0 = \beta_1 + \beta_0 X_0 + u_0 \quad \ldots (5.13)$$

हमें ज्ञात है कि चर b_1 और b_2 चर β_1 और β_2 के अनभिनत आकलक हैं। अतः पद $\hat{Y}_0$ पद $E(Y \mid X_0)$ का एक अनभिनत प्राग्सूचक होगा।

अतएव,

$$\hat{Y}_0 = b_1 + b_2 X_0 \quad \ldots (5.14)$$

चूँकि $\hat{Y}_0$ एक आकलक है, वास्तविक मान Y_0 पद $\hat{Y}_0$ से भिन्न होगा और उसमें कोई 'पूर्वानुमान त्रुटि' निहित होगी।

पद $\hat{Y}_0 - Y_0$, में पूर्वानुमान त्रुटि निम्नवत् दर्शायी जाती है–

$$\hat{Y}_0 - Y_0 = (b_1 + b_2 X_0) - (\beta_1 + \beta_2 X_0 + u_0) \quad \ldots (5.15)$$

समीकरण (5.15) के पदों को पुनर्व्यवस्थित कर हम निम्नलिखित समीकरण प्राप्त कर सकते हैं –

$$\hat{Y}_0 - Y_0 = (b_1 - \beta_1) + (b_2 - \beta_2) X_0 - u_0$$

चलिए, समीकरण (5.15) का प्रत्याशित मान निम्नवत् मान लेते हैं–

$$\mathrm{E}(\hat{Y}_0 - Y_0) = E(b_1 - \beta_1) + E(b_2 - \beta_2) X_0 - E(u_0) \quad \ldots (5.16)$$

हमें ज्ञात है कि $E(b_1) = \beta_1, E(b_2) = \beta_2$ और $E(u_0) = 0$ होता है।

तदनुसार, हम पाते हैं कि पूर्वानुमान त्रुटि का प्रत्याशित मान शून्य होता है।

आइए, अब यहाँ पूर्वानुमान त्रुटि के विचरण का पता लगाते हैं।

पूर्वानुमान त्रुटि का विचरण,

$$\mathrm{V}(\hat{Y}_0 - Y_0) = V(b_1 - \beta_1) + V(b_2 - \beta_2) X_0$$
$$+2\, X_0\, cov(b_1 - \beta_1, b_2 - \beta_2) + V(u_0) \quad \ldots (5.17)$$

हमें ज्ञात है कि

$$V(b_1) = \sigma^2 \frac{\Sigma X_i^2}{n \Sigma x_i^2} \quad \ldots (5.18)$$

$$V(b_2) = \frac{\sigma^2}{\Sigma x_i^2} \quad \ldots (5.19)$$

$$Cov(b_1, b_2) = -\bar{X} \left(\frac{\sigma^2}{\Sigma x_i^2} \right) \quad \ldots (5.20)$$

ऊपर प्राप्त तीनों समीकरणों को जोड़कर और योगफल के पदों को पुर्नव्यवस्थित कर हम निम्नलिखित समीकरण प्राप्त कर सकते हैं –

$$V(\hat{Y}_0 - Y_0) = \sigma^2 \left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X})}{\Sigma x_i^2}\right] \quad \ldots (5.21)$$

इस प्रकार, Y_0 माध्य $\beta_1 + \beta_0 X_0$ और विचरण $\sigma^2 \left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\Sigma x_i^2}\right]$ के साथ प्रसामान्य बंटन का अनुसरण करता है।

यदि पद σ^2 के लिए आकलक लें तो हमें निम्नलिखित समीकरण प्राप्त होगा–

$$t = \frac{\hat{Y}_0 - (\beta_1 + \beta_0 X_0)}{SE(\hat{Y}_0)} \quad \ldots (5.22)$$

समीकरण (5.22) के आधार पर हम पद $\hat{Y}_0$ के लिए विश्वास्यता अंतराल की रचना कर सकते हैं। चूँकि

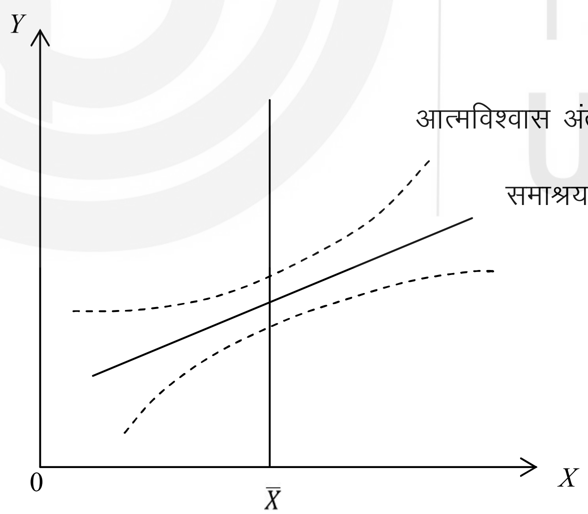
$t = \frac{\hat{Y}_0 - E(Y/\alpha_0)}{SE(\hat{Y}_0)}$ है, हमें प्राप्त होता है–

$$P\left[-t_{\alpha/2} \leq t \leq t_{\alpha/2}\right] = 1 - \alpha$$

तदनुसार, पद $\hat{Y}_0$ का विश्वास्यता अंतराल होगा–

$$P\left[(b_1 + b_2 X_0) - t_{\alpha/2} SE(\hat{Y}_0) \leq (\beta_1 + \beta_2 X_0) \leq (b_1 + b_2 X_0) + t_{\alpha/2} SE(\hat{Y}_0)\right] = 1 - \alpha \quad \ldots (5.23)$$

चलिए, समीकरण (5.21) पर एक बार फिर से दृष्टि डालते हैं। हम पाते हैं कि पद $\hat{Y}_0$ का विचरण पद $(X_0 - \bar{X})^2$ के साथ बढ़ता है। तदनुसार, विचरण में वृद्धि होती है, बशर्ते चर X_0 प्रतिदर्श के उस माध्य $\bar{X}$ से और दूर चला जाए जिसके आधार पर चर b_1 और b_2 परिकलित किए गए हों। चित्र 5.4 में हमने पद $\hat{Y}_0$ के लिए विश्वास्यता अंतराल दर्शाया है (देखें बिंदुकित रेखा)।



चित्र 5.4: व्यष्टिक पूर्वानुमानन हेतु विश्वास्यता अंतराल

5.7.2 माध्य पूर्वानुमान

माध्य पूर्वानुमान Y_0 के प्रत्याशित मानों के पूर्वानुमान को इंगित करता है, न कि व्यष्टिक मान को। दूसरे शब्दों में, यहाँ हम निम्नलिखित समीकरण का पूर्वानुमान व्यक्त कर रहे हैं–

$$\hat{Y}_0 = b_1 + b_2 X_0$$

तदनुसार, त्रुटि पद u_0 को नहीं जोड़ा जाता है।

माध्य पूर्वानुमान के उदाहरण में पद $[\hat{Y}_0 - Y_0]$ की पूर्वानुमान त्रुटि निम्नवत् दर्शायी जाती है –

$$\hat{Y}_0 - Y_0 = (b_1 + b_2X_0) - (\beta_1 + \beta_2X_0) \qquad \ldots (5.24)$$

समीकरण (5.24) के पदों को पुनर्व्यवस्थित कर हम निम्नलिखित समीकरण प्राप्त कर सकते हैं –

$$\hat{Y}_0 - Y_0 = (b_1 - \beta_1) + (b_2 - \beta_2)X_0$$

समीकरण (5.24) का प्रत्याशित मान लेने पर हमें निम्नलिखित समीकरण प्राप्त होता है–

$$\mathrm{E}(\hat{Y}_0 - Y_0) = E(b_1 - \beta_1) + E(b_2 - \beta_2)X_0 \qquad \ldots (5.25)$$

इस प्रकार, हम देखते हैं कि पूर्वानुमान त्रुटि का प्रत्याशित मान शून्य है।

आइए, अब हम माध्य पूर्वानुमान के उदाहरण में पूर्वानुमान त्रुटि के विचरण का पता लगाते हैं।

पूर्वानुमान त्रुटि का विचरण,

$$\mathrm{V}(\hat{Y}_0 - Y_0) = V(b_1 - \beta_1) + V(b_2 - \beta_2)X_0$$
$$+2\,X_0\,cov(b_1 - \beta_1, b_2 - \beta_2) \qquad \ldots (5.26)$$

यदि हम समीकरणों (5.17) और (5.26) की तुलना करते हैं तो हमें एक महत्वपूर्ण परिवर्तन दिखाई देता है – पद $V(u_0)$ समीकरण (5.26) में नहीं है। तदनुसार, माध्य पूर्वानुमान के उदाहरण में पूर्वानुमान त्रुटि का विचरण व्यष्टिक पूर्वानुमान की तुलना में कम है। तथापि, माध्य पूर्वानुमान के मामले में पद $\hat{Y}_0$ के विचरण में परिवर्तन होता है। माध्य पूर्वानुमान के उदाहरण में पूर्वानुमान त्रुटि का विचरण निम्नवत् दर्शाया जाता है–

$$V(\hat{Y}_0 - Y_0) = \sigma^2\left[\frac{1}{n} + \frac{(X_0 - \bar{X})}{\Sigma x_i^2}\right] \qquad \ldots (5.27)$$

पुनः, पूर्वानुमान त्रुटि के विचरण में वृद्धि होती है, बशर्ते चर X_0 प्रतिदर्श के उस माध्य $\grave{X}$ से और दूर चला जाए जिसके आधार पर चर b_1 और b_2 परिकलित किए गए हों। यह कुछ–कुछ उस विश्वास्यता अंतराल जैसा दिखाई देगा जो हमने चित्र 5.4 में दर्शाया था, परंतु विश्वास्यता अंतराल का विस्तार निरर्थक ही सिद्ध होगा।

उपर्युक्त से हम जो निष्कर्ष निकालते हैं, वह यह है कि हम व्याख्यात्मक चर (X_0) के किसी भी विशिष्ट मान के लिए आकलित समाश्रयण समीकरण के आधार पर आश्रित चर के मान का पूर्वानुमान व्यक्त कर सकते हैं। बहरहाल, यदि X का विशिष्ट मान चर $\grave{X}$ के मान से अधिक भिन्न हुआ तो हमारे पूर्वानुमान की विश्वसनीयता कम हो जाएगी।

बोध प्रश्न 3

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

1) गॉस–मार्कोव प्रमेय लिखिए।

……………………………………………………………………………………

……………………………………………………………………………………

……………………………………………………………………………………

……………………………………………………………………………………

2) पूर्वानुमानन में दो प्रकार की पूर्वानुमान संभावनाओं के बीच अंतर स्पष्ट करें।

……………………………………………………………………………………

……………………………………………………………………………………

……………………………………………………………………………………

……………………………………………………………………………………

5.8 सार–संक्षेप

इस इकाई में हमने जाना कि किस प्रकार किसी साधारण समाश्रयण मॉडल के आकलित परिणामों से निष्कर्ष निकाला जाए। मूल बातों को दोहराने के लिए परिकल्पना परीक्षण का लेखा–जोखा प्रस्तुत करने के बाद, इसमें आकलित परिणामों के सत्यापन पर निर्णय लेने के लिए दो उपागमों की व्याख्या की गई।

उक्त दो उपागम हैं– विश्वास्यता अंतराल उपागम और सार्थकता परीक्षण उपागम। मॉडल के समग्र सार्थकता के परीक्षण को एनोवा (ANOVA) तकनीक द्वारा समझाया गया है। यहाँ F–प्रतिदर्शज के अनुप्रयोग को समझाया गया है।

पारंपरिक रैखिक समाश्रयण मॉडल की अवधारणा कुछ अद्वितीय गुणधर्मों का लाभ उठाने वाले आकलित मापदंडों की ओर ले जाती है। इसके आलोक में आकलनो को सर्वोत्तम रैखिक अनभिनत आकलन (BLUE) कहा जाता है। इस तथ्य को गॉस–मार्कोव प्रमेय नामक परिणाम में स्पष्ट किया गया है।

इकाई पूर्वानुमानन की अवधारणा के विस्तृत विवरण के साथ समाप्त होती है। पुनः, यह एक ऐसी तकनीक है जिसमें हम एक ऐसा विश्वास्यता अंतराल प्रस्तुत करते हैं जिसमें आश्रित चर के आकलित अथवा पूर्वानुमानित मान की स्थिति दर्शायी जाती है।

5.11 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1) दोनों शून्य–स्तरीय परिकल्पनाओं (यथा, समाश्रयण और अवरोधन के वास्तविक समष्टि प्राचलों के लिए) के उदाहरण में से शून्य–स्तरीय परिकल्पना एक सरल परिकल्पना है, जबकि वैकल्पिक परिकल्पना सम्मिश्रित है। पूर्ववर्ती को आमतौर पर शून्य के बराबर माना जाता है (जब तक कि उसे किसी ज्ञात मान के बराबर न माना जाए) और परवर्ती का उल्लेख असमानता के संदर्भ में किया जाता है। परवर्ती को 'बराबर नहीं है' के पदों में व्यक्त किए जाने पर द्विपक्षीय परिकल्पना के रूप में भी जाना जाता है। यदि उसे $>$ या $<$ के पदों में व्यक्त किए जाने पर एकपक्षीय ही माना जाता है।

2) नहीं, यह कथन सही नहीं है। यह वैकल्पिक परिकल्पना ही है जो यह तय करती है कि कोई परिकल्पना सम्मिश्रित है अथवा एकपक्षीय। यदि वैकल्पिक परिकल्पना को शून्य के बराबर नहीं बताया जाता है तो वह सम्मिश्रित अथवा द्विपक्षीय परीक्षण होगा। दूसरे प्रकार से, यथा यदि वैकल्पिक परिकल्पना को धनात्मक अथवा ऋणात्मक पदों में व्यक्त किया जाता है तो यह एकपक्षीय परीक्षण ही कहलाएगा।

3) विश्वास्यता अंतराल उपागम परिकल्पना के परीक्षण की एक विधि है। यह इस प्रायिकता को इंगित करता है कि कोई समष्टि प्राचल दी गई तालिका से निकाले गए क्रांतिक मानों की शृंखला में ही शामिल होगा।

4) हम कहते हैं कि परिकल्पित मान अंतराल में निहित है क्योंकि अंतराल का मान प्रतिदर्श अथवा आकलन के लिए प्रयोग किए गए आँकड़ों पर निर्भर करता है। वास्तविक समष्टि प्राचल मान नियत होता है परंतु प्रतिदर्श के आधार पर अंतराल बदलता है।

बोध प्रश्न 2

1) सार्थकता का परीक्षण उपागम परिकल्पना के परीक्षण की ही एक अन्य विधि है। पद H_0 को स्वीकार अथवा अस्वीकार करने का निर्णय प्रतिदर्श आँकड़ों से प्राप्त परीक्षण आँकड़ों के मान के आधार पर किया जाता है। यह परीक्षण प्रतिदर्शज समीकरण $t = \frac{b_2 - \beta_2}{SE(b_2)}$ से दर्शाया जाता है और यह $(n - k)$ d.f. के साथ t–बंटन का अनुसरण करता है।

2) जब शून्य–स्तरीय परिकल्पना को निरस्त कर दिया जाता है तो यह साक्ष्य की शक्ति के मापन के रूप में सामने आता है। इससे यह निष्कर्ष निकालता है कि प्रभाव सांख्यिकीय रूप से सार्थक है। यह सत्य सिद्ध होने पर शून्य–स्तरीय परिकल्पना को अस्वीकार कर दिए जाने की संभावना होती है। यह एक गंभीर त्रुटि है और इसलिए इसे 1 प्रतिशत अथवा 5 प्रतिशत जैसे किसी छोटे मापन में चुना जाता है।

3) विचरण का विश्लेषण (ANOVA) एक ऐसी तकनीक या उपकरण है जिसका उपयोग दिए गए आँकड़ों का दो प्रकार से अथवा दो दिशाओं में विश्लेषण करने के लिए किया जाता है। एक को नियतात्मक कारकों के लिए उत्तरदायी माना जाता है, और व्याख्या किया गया भाग अथवा सुव्यवस्थित भाग भी कहा जाता है। दूसरे को यादृच्छिक अथवा अस्पष्टीकृत भाग कहा जाता है। विचरण विधि के विश्लेषण की इस पद्धति को वर्ष 1918 में रोनाल्ड फिशर द्वारा विकसित किया गया था।

4) अर्थमिति में t– परीक्षण का प्रयोग का उपयोग आकलित व्यष्टिक गुणांक की सार्थकता का परीक्षण करने के लिए किया जाता है। इसे $(k - 1)$ स्वतंत्रता की कोटियों (d.f) के साथ t के रूप में बंटित किया जाता है, जहाँ k अवरोधन पद सहित आकलित प्राचलों की संख्या है। तदनुसार, किसी भी सरल रैखिक समाश्रयण के लिए यह $[n - (2 - 1)] = n - 1$ होगी। इससे भिन्न, F–परीक्षण का प्रयोग समग्र मॉडल के सार्थकता परीक्षण के लिए किया जाता है। इसके दो प्राचल होते हैं। किसी भी F–परीक्षण के लिए स्वतंत्रता की कोटियाँ सामान्यतः $(k - 1)$ और $(n - k)$ होती हैं। चर k अवरोधन पद शामिल होता है। अतः किसी भी सरल रैखिक समाश्रयण में चर F के लिए स्वतंत्रता की कोटियाँ $(2 - 1)$ और $(n - 2)$ अथवा 1 और $(n - 2)$ होती हैं। यह ध्यान देने की बात है कि किसी भी सरल रैखिक समाश्रयण में t–परीक्षण और F–परीक्षण समतुल्य होते हैं क्योंकि स्वतंत्र चर की संख्या केवल एक ही होती है।

बोध प्रश्न 3

1) गॉस–मार्कोव प्रमेय में कहा गया है कि साधारण न्यूनतम वर्ग (OLS) आकलक भी सर्वोत्तम रैखिक अनभिनत आकलक (BLUE) सिद्ध होते हैं। उक्त गुणधर्म (BLUE) की विद्यमानता का अर्थ है कि उक्त विधि (OLS) से प्राप्त आकलक निम्नलिखित गुणधर्मों को बनाए रखता है – (i) वह रैखिक होता है अर्थात् आकलक किसी यादृच्छिक चर का एक रैखिक फलन होता है, जैसे समाश्रयण मॉडल में आश्रित चर Y; (ii) यह अनभिनत होता है अर्थात् इसका औसत अथवा प्रत्याशित मान इस अर्थ में वास्तविक मान के बराबर होता है कि $E(b_2) = \beta_2$; (iii) ऐसे सभी रैखिक अनभिनत आकलकों के

वर्ग में इसका न्यूनतम विचरण होता है। न्यूनतम विचरण वाले ऐसे आकलक को एक दक्ष आकलक के रूप में भी जाना जाता है।

2) पूर्वानुमानन का अर्थ दो प्रकार के मानों की भविष्यवाणी करना होता है – एक, सशर्त माध्य का पूर्वानुमान, यथा $E(Y \mid X_0) \rightarrow$ समष्टि समाश्रयण रेखा पर कोई बिंदु। इसे 'माध्य पूर्वानुमान' कहा जाता है। दूसरे, व्यष्टिक Y मान का पूर्वानुमान, संवादी $f(X_0)$ को 'व्यक्तिगत पूर्वानुमान' कहा जाता है।

इकाई 6 द्वि–चरीय समाश्रयण मॉडल का विस्तार *

इकाई की रूपरेखा

6.0 उद्देश्य

6.1 विषय–प्रवेश

6.2 मूल बिंदु के माध्यम से समाश्रयण

6.3 मापन इकाइयों में परिवर्तन

6.4 अर्ध–लघुगणकीय मॉडल

6.5 लघुगणकीय–रैखिक मॉडल

6.6 फलनिक रूप का चयन

6.7 सार–संक्षेप

6.8 बोध प्रश्नों के उत्तर अथवा संकेत

6.0 उद्देश्य

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

- मूल बिंदु से गुजरने वाले समाश्रयण मॉडलों की व्याख्या कर सकें;
- आकलों पर आश्रित और स्वतंत्र चरों के मापन की इकाई में परिवर्तनों के प्रभाव की व्याख्या कर सकें;
- अर्ध–लघुगणक और लघुगणक–रैखिक समाश्रयण मॉडल में प्राचलों की व्याख्या कर सकें; तथा
- समाश्रयण मॉडल के सही फलनिक रूप की पहचान कर सकें।

6.1 विषय–प्रवेश

पिछली दो इकाइयों में हमने इस बात चर्चा की कि कैसे किसी दो–चर समाश्रयण मॉडल का आकलन किया जा सकता है और आकलित समाश्रयण समीकरण के आधार पर निष्कर्ष कैसे निकाले जा सकते हैं। इस संदर्भ में हमने आकलन की साधारण न्यूनतम वर्ग (OLS) विधि के विषय में चर्चा की। स्मरण करें कि इस विधि (OLS) से प्राप्त आकलक इस अर्थ में सर्वोत्तम रैखिक अनभिनत आकलक (BLUE) होते हैं कि वे रेखीय समाश्रयण मॉडल के वर्ग में सर्वश्रेष्ठ सिद्ध होते हैं।

इस दो–चर समाश्रयण मॉडल का फलन निम्नानुसार होता है –

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \text{... (6.1)}$$

जहाँ Y आश्रित चर है और X स्वतंत्र चर है। हमने अपने समाश्रयण मॉडल में एक प्रसंभाव्य त्रुटि पद (u_i) जोड़ा है। हमने अपने समाश्रयण मॉडल में त्रुटि पद को शामिल करने के तीन कारणों का हवाला दिया –

(i) यह मॉडल में बहिष्कृत चर का ध्यान रखता है,

(ii) यह मॉडल में अननुमेय मानव स्वभाव को शामिल करता है, और

---

* प्रोफेसर कौस्तुभ बारिक, इंदिर गाँधी राष्ट्रीय मुक्त विश्वविद्यालय।

(iii) यह मापन त्रुटि, त्रुटिपूर्ण फलन रूप, आदि प्रभावों को समाहित कर लेता है।

हम यह मानकर चले कि समाश्रयण मॉडल सही ढंग से निर्दिष्ट होता है। इस मॉडल में सभी प्रासंगिक चर शामिल किए जाते हैं। समाश्रयण मॉडल में कोई अप्रासंगिक चर शामिल नहीं होता है।

अब इस इकाई में हम पिछली दो इकाइयों की ही भाँति दो चरों के उदाहरण को जारी रखेंगे। हम उन अवधारणाओं का अनुसरण भी यहाँ जारी रखेंगे जिनका उल्लेख इकाई 4 में किया गया है।

आइए, ऊपर समीकरण (6.1) में दिए गए समाश्रयण मॉडल को देखें।. आप देखेंगे कि समाश्रयण मॉडल अपने प्राचलों में रैखिक है। हमारे पास प्राचलों के जटिल रूप नहीं हैं, जैसे कि प्राचलों के रूप में β_2^2 अथवा $\beta_1\beta_2$ आदि। इसके अलावा, इस समाश्रयण मॉडल में रैखिक चर हैं। हमारे पास X^2 अथवा $\log X$ व्याख्यात्मक चर के रूप में नहीं है।

क्या हमारे पास किसी समाश्रयण मॉडल में इस प्रकार के चर हो सकते हैं? यदि ऐसे चर विद्यमान हों तो हम अपने समाश्रयण मॉडल की व्याख्या कैसे करते हैं? हम समीकरण (6.1) में दिए गए सरल समाश्रयण मॉडल का विस्तार करेंगे और स्पष्ट करेंगे कि किंचित परिवर्तनों के साथ मॉडल की व्याख्या कैसे बदल जाती है।

6.2 मूल बिंदु के माध्यम से समाश्रयण

आइए, अब समीकरण (6.1) में दिए गए सरल समाश्रयण मॉडल का विश्लेषण करते हैं। इस समाश्रयण मॉडल में दो प्राचल हैं – β_1 और β_2, जहाँ चर β_1 अवरोधन प्राचल कहलाता है, जबकि चर β_2 समाश्रयण प्राचल कहलाता है। जब व्याख्यात्मक चर मान शून्य लेता है तो अवरोधन β_1 आश्रित चर के मान को इंगित करता है, यथा $E(Y_0|X_0) = \beta_1$.

मान लीजिए कि हमारा समाश्रयण मॉडल निम्नलिखित रूप लेता है –

$$Y_i = \beta_2 X_i + u_i \quad \text{... (6.2)}$$

समीकरण (6.2) में केवल एक प्रवण प्राचल β_2 है। यहाँ कोई अवरोधन नहीं है। इसका निहितार्थ यह है कि यहाँ समाश्रयण रेखा मूल बिंदु से होकर गुजरती है। हमारा समष्टि समाश्रयण फलन $Y = \beta_2 X_i + u_i$ है और प्रतिदर्श समाश्रयण फलन $Y_i = b_2 X_i + e_i$ है।

अब हम OLS विधि प्रयोग करते हैं और OLS आकलक b_2 का पता लगाते हैं। जैसा कि हम इकाई 4 से जानते हैं, OLS विधि में वर्गों के त्रुटि योगफल (ESS) को न्यूनतम किया जाता है। तदनुसार, हम वर्गों के त्रुटि योगफल को न्यूनतम कर प्राप्त करते हैं –

$$ESS = \sum e_i^2 = \sum (Y_i - b_2 X_i)^2 \quad \text{... (6.3)}$$

हम वर्गों के त्रुटि योगफल (ESS) के अवकलज को लेकर उसे शून्य के समतुल्य करते हैं –

$$\frac{d\sum e_i^2}{db_2} = 0 \quad \text{... (6.4)}$$

$$\frac{d\sum e_i^2}{db_2} = 2\sum (Y_i - b_2 X_i)(-X_i) = 0 \quad \text{... (6.5)}$$

इसका निहितार्थ है कि –

$$-2\sum e_i\left(X_i Y_i - b_2 X_i^2\right) = 0$$

$$\sum X_i Y_i - b_2 \sum X_i^2 = 0$$

$$b_2 = \frac{\sum X_i Y_i}{\sum X_i^2} \quad \text{... (6.6)}$$

समीकरण (6.6) में दिया गया आकलक अनभिनत होता है। आकलक का विचरण निम्नवत् दर्शाया जाता है –

$$var(b_2) = \frac{\sigma^2}{\Sigma X_i^2} \quad \text{... (6.7)}$$

अब उपर्युक्त आकलक की तुलना समाश्रयण मॉडल $Y_i = \beta_1 + \beta_2 X_i + u_i$ हेतु आकलक से करें (इकाई 4 में समीकरण (4.18) देखें)।

$$b_2 = \frac{\Sigma x_i y_i}{\Sigma x_i^2} \quad \text{... (6.8)}$$

तथा

$$var(b_2) = \frac{\sigma^2}{\Sigma x_i^2} \quad \text{... (6.9)}$$

यहाँ ध्यान दें कि समीकरण (6.6) में चर विचलन रूप में नहीं हैं। तदनुसार जब हमारे पास समाश्रयण मॉडल में अवरोधन नहीं होता है तो समाश्रयण प्राचल का आकलक अवरोधन वाले समाश्रयण मॉडल से भिन्न होता है। ये दोनों आकलक एक समान होंगे यदि और केवल यदि $\bar{X} = 0$ होगा।

नीचे दी गई तालिका 6.1 अवरोधन सहित और अवरोधन रहित समाश्रयण मॉडल के बीच तुलना प्रस्तुत करती है।

तालिका 6.1: अवरोधन के बिना समाश्रयण मॉडल के अभिलक्षण

अवरोधन सहित समाश्रयण मॉडल	अवरोधन रहित समाश्रयण मॉडल
$b_2 = \frac{\sum x_i y_i}{\sum x_i^2}$	$b_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$
$\mathrm{var}(b_2) = \frac{\sigma^2}{\sum x_i^2}$	$\mathrm{var}(b_2) = \frac{\sigma^2}{\sum X_i^2}$
$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2}$	$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-1}$
R^2 ऋणेतर है	R^2 ऋणात्मक हो सकता है

आकलित समाश्रयण मॉडल निम्नवत् दर्शाया जाता है –

$$\hat{Y}_i = b_2 X_i \quad \text{... (6.10)}$$

आप देखेंगे कि निर्धारण का गुणांक R^2 अवरोधन के बिना समाश्रयण मॉडल के लिए उपयुक्त नहीं होता है। यदि समाश्रयण मॉडल में अवरोधन सांख्यिकीय रूप से सार्थक न हो तो हम मूल बिंदु के माध्यम से समाश्रयण कर सकते हैं। अन्यथा, यह विनिर्देशन त्रुटि की ओर जाता है। यहाँ कोई प्रासंगिक चर छूट जाता है।

6.3 मापन इकाइयों में परिवर्तन

मान लीजिए कि आपको विगत 30 वर्षों के भारत के सकल घरेलू उत्पाद और कुल उपभोग व्यय पर काल–श्रेणी आँकड़े दिए गए हैं। आपको एक ऐसा समाश्रयण मॉडल चलाने के लिए कहा जाता है जिसमें उपभोग व्यय आश्रित चर के रूप में और आय स्वतंत्र चर के रूप में होगी। इसका उद्देश्य भारत के कुल उपभोग फलन का आकलन करना है।

मान लीजिए कि आपने सकल घरेलू उत्पाद और उपभोग व्यय को 'करोड़ रुपये' में दर्शाया है। अब आपको आकलित समाश्रयण समीकरण निम्नवत् प्राप्त होता है –

$$Y_i = 237 + 0.65X_i \quad \text{... (6.11)}$$

जब आपने अपने वरिष्ठ जन के सामने ये परिणाम प्रस्तुत किए तो उन्होंने कहा कि सकल घरेलू उत्पाद और उपभोग व्यय का मापन 'मिलियन रुपये' में होना चाहिए था ताकि यह भारत के बाहर भी लोगों को समझ में आए। यदि आप चरों को परिवर्तित करके परिणामों पुनः आकलित करते हैं तो क्या आकल एक समान होंगे? या फिर क्या आप आकलों में किसी परिवर्तन की आशा करेंगे? आइए, इस मुद्दे पर कुछ विस्तार से चर्चा करें।

मान लीजिए कि हम आश्रित और स्वतंत्र दोनों चरों को निम्नानुसार रूपांतरित करते हैं –

$Y_i^* = w_1 Y_i$ और $X_i^* = w_1 X_i$ ... (6.12)

समाश्रयण मॉडल (6.1) को अब निम्नवत् परिवर्तित किया जा सकता है –

$$Y_i^* = \beta_1 + \beta_2 X_i^* + u_i \quad \text{... (6.13)}$$

अब OLS विधि से समीकरण (6.13) का आकलन करने से निम्नलिखित आकल प्राप्त होते हैं –

$$b_1^* = \bar{Y}^* - b_2^* \bar{X}^* \quad \text{... (6.14)}$$

$$b_2^* = \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}} \quad \text{... (6.15)}$$

इसी प्रकार हम चर b_1^* और b_2^* का विचरण और त्रुटि विचरण का आकलक भी ज्ञात कर सकते हैं।

समीकरण (6.15) से हम यह पता लगा सकते हैं कि

$$b_2^* = \frac{w_1}{w_2} b_2 \quad \text{... (6.16)}$$

तथा

$$b_1^* = w_1 b_1 \quad \text{... (6.17)}$$

आइए, अब हम एक–एक कर उपर्युक्त के निहितार्थ देखते हैं –

(i) चलिए, हम आश्रित चर Y_i से आरंभ करते हैं। मान लीजिए कि चर Y_i दोगुना ($w_1 = 2$) है और चर X_i अपरिवर्तित ($w_2 = 1$) है। अब चर b_1 और b_2 का क्या होगा? चर w_1 और w_2 के मान समीकरण (6.16) और (6.17) में प्रतिस्थापित करें। आप पाएँगे कि दोनों आकलन दोगुने हो गए हैं। तदनुसार, यदि आश्रित चर को किसी नियत c से गुणा किया जाता है तो सभी OLS गुणांकों को c से ही गुणा किया जाएगा।

(ii) अब हम स्वतंत्र चर का उदाहरण लेते हैं। मान लीजिए कि चर X_i दोगुना ($w_2 = 2$) है और चर Y_i अपरिवर्तित ($w_1 = 1$) है। चर w_1 और w_2 के मानों को समीकरणों (6.16) और (6.17) में प्रतिस्थापित करने पर हम पाते हैं कि समाश्रयण गुणांक (b_2) आधा हो जाता है, परंतु अवरोधन (b_1) अपरिवर्तित ही रहता है।

(iii) यदि हम दोनों चरों X_i और Y_i को दोगुना कर देते हैं तो समाश्रयण गुणांक (b_2) अपरिवर्तित ही रहेगा, परंतु अवरोधन बदल जाएगा। याद रखें कि आश्रित चर के मापन स्तर में परिवर्तन से अवरोधन बदल जाता है।

अब प्रश्न यह उठता है कि क्या मॉडल के t–अनुपात और F–मान में कोई बदलाव आएगा? नहीं, t और F प्रतिदर्शज किसी भी चर के मापन स्तर में किसी भी परिवर्तन से प्रभावित नहीं होते हैं।

बोध प्रश्न 1

1) हमें किस स्थिति में मूल बिंदु के माध्यम से समाश्रयण चलाना चाहिए?

....

....

....

....

2) मूल बिंदु के माध्यम से किसी समाश्रयण मॉडल के निहितार्थ क्या होते हैं? स्पष्ट करें।

....

....

....

....

3) यदि व्याख्यात्मक चर के मापन स्तर में कोई परिवर्तन होता है तो आकलों पर क्या प्रभाव पड़ता है? स्पष्ट करें।

....

....

....

....

4) यदि आश्रित चर के मापन स्तर में कोई परिवर्तन होता है तो आकलों पर क्या प्रभाव पड़ता है? स्पष्ट करें।

....

....

....

....

6.4 अर्ध–लघुगण्कीय मॉडल

कुछ उदाहरणों में समाश्रयण मॉडल अरैखिक होता है, फिर भी समाश्रयण समीकरण के दोनों पक्षों पर लघुगणक लेने से हमें कोई रैखिक मॉडल ही प्राप्त होता है। यदि कोई मॉडल अरैखिक हो परंतु अपने चर रूपांतरित हो जाने के बाद वह रैखिक हो जाता हो तो मॉडल को आंतरिक रूप से रैखिक कहा जाता है। तदनुसार, अर्ध–लघुगणक और लघुगणक–रैखिक मॉडल आंतरिक रूप से रैखिक मॉडल ही होते हैं। इस पाठांश में हम अर्ध–लघुगणक मॉडल के विषय में ही चर्चा करेंगे। लघुगणक–रैखिक मॉडल के विषय में चर्चा हम अगले पाठांश में करेंगे।

आइए, हम एक फलनिक रूप के साथ निम्नानुसार आरंभ करते हैं –

$$Y_t = e^{\beta_1 + \beta_2 X_t + u_t} \quad \dots (6.18)$$

यह समाश्रयण मॉडल, अपने वर्तमान स्वरूप में, अरैखिक है। अतः OLS विधि से इसका आकलन नहीं किया जा सकता है। तथापि, यदि हम इसके दोनों पक्षों के प्राकृतिक लघुगणक लेते हैं तो हमें निम्नलिखित समीकरण प्राप्त होता है –

$\ln Y_t = \beta_1 + \beta_2 t + u_t$ ... (6.19)

यह एक अर्ध–लघुगणक समीकरण में बदल जाता है। इसे ही अर्ध–लघुगणक मॉडल कहा जाता है क्योंकि एक चर लघुगणक रूप में होता है।

यदि हम मान $\ln Y_t = Y_t^*$ लेते हैं तो समीकरण (6.19) को निम्नवत् लिखा जा सकता है –

$Y_t^* = \beta_1 + \beta_2 t + u_t$ ... (6.20)

समीकरण (6.20) का आकलन सरल है। यह समीकरण अपने प्राचलों और चरों में रैखिक है। तदनुसार, हम प्राचलों का आकलन करने के लिए OLS विधि प्रयोग कर सकते हैं। समाश्रयण मॉडल (6.20) का निहितार्थ, हालाँकि, समाश्रयण मॉडल (6.1) से बहुत भिन्न है।

यदि हम समाश्रयण मॉडल $Y_t = \beta_1 + \beta_2 X_t + u_t$ का अवकलन लेते हैं तो हमें निम्नलिखित समीकरण प्राप्त होता है –

$\frac{dY}{dt} = \beta_2$ ... (6.21)

समीकरण (6.21) दर्शाता है कि समाश्रयण समीकरण की प्रवणता नियत है। उपर्युक्त का एक निहितार्थ यह भी है कि स्वतंत्र चर में इकाई वृद्धि के लिए आश्रित चर में निरपेक्ष परिवर्तन पूरे प्रतिदर्श में नियत रहता है। यदि चर X में एक इकाई वृद्धि होती है तो चर Y में β_2 इकाई वृद्धि होती है।

आइए, अब हम समाश्रयण मॉडल $\ln Y_t = \beta_1 + \beta_2 t + u_t$पर विचार करते हैं। यदि हम समीकरण (6.19) का अवकलन लें तो पाते हैं कि

$\frac{d\ln Y_t}{dt} = \beta_2$ होता है,

जिसका अर्थ है कि

$\frac{1}{Y_t}\frac{dY_t}{dt} = \beta_2$ ... (6.22)

समीकरण (6.22) का एक निहितार्थ यह भी है कि समाश्रयण मॉडल की प्रवणता परिवर्तनशील होती है। तदनुसार इसकी व्याख्या समाश्रयण मॉडल $Y_t = \beta_1 + \beta_2 X_t + u_t$ से भिन्न होती है।

समीकरण (6.19) के लिए हम समाश्रयण गुणांक (β_2) की व्याख्या इस प्रकार करते हैं – चर X में प्रत्येक इकाई वृद्धि के लिए चर Y में β_2 प्रतिशत वृद्धि हुई है। इस प्रकार, अर्ध–लघुगणकीय मॉडल के लिए आश्रित चर में परिवर्तन प्रतिशतता के पदों में प्राप्त होता है। अतएव, अर्ध–लघुगणकीय मॉडल विकास दर का आकलन करने के लिए उपयोगी सिद्ध होता है।

6.5 लघुगणकीय–रैखिक मॉडल

आइए, हम निम्नलिखित समाश्रयण समीकरण पर विचार करते हैं –

चलिए, सबसे पहले निम्नलिखित अरैखिक मॉडल का उदाहरण लेते हैं –

$Y = \beta_1 X^{\beta_2}$ ... (6.23)

यह मॉडल आंतरिक रूप से रैखिक होगा यदि इसे निम्नवत् परिवर्तित किया जा सकता हो –

$Y^* = \beta_1 + \beta_2 X^* + u$ ... (6.24)

समीकरण (6.23) में प्रत्येक चर के लघुगणक का प्रयोग करके हम निम्नलिखित रूपांतरित समीकरण प्राप्त करते हैं –

$\ln Y_i = \beta_1 + \beta_2 \ln X_i + u_i$ ... (6.25)

समीकरण (6.25) में दर्शाए गए समाश्रयण मॉडल को ही लघुगणक–रैखिक मॉडल कहा जाता है (क्योंकि यह अपने चरों के लघुगणक में रैखिक होता है) अथवा दोहरा–लघुगणक मॉडल (क्योंकि इसमें दोनों चर लघुगणक रूप में होते हैं)।

आइए, अब निम्नानुसार चर X_i के सन्दर्भ में समीकरण (6.25) का अवकलन लेते हैं –

$$\frac{d(\ln Y_i)}{dX_i} = \frac{1}{Y_i}.\frac{\mathrm{d}Y_i}{dX_i} \qquad \text{... (6.26)}$$

$$\frac{\mathrm{d}Y_i}{dX_i} = \frac{\beta_2}{X_i} \qquad \text{... (6.27)}$$

समीकरण (6.26) और (6.27) का योग करके हमें निम्नलिखित समीकरण प्राप्त होता है –

$$\frac{\mathrm{d}Y_i}{dX_i} = \frac{Y_i}{X_i}\beta_2$$

अथवा,

$$\frac{\mathrm{d}Y_i}{dX_i}\frac{X_i}{Y_i} = \beta_2 \qquad \text{... (6.28)}$$

समीकरण (6.28) को ध्यानपूर्वक देखने से ज्ञात होता है कि समाश्रयण प्राचल चर Y और चर X के बीच लोचता को निरूपित करता है।

लघुगणक–रैखिक मॉडल की इस आकर्षक विशेषता ने इसे व्यावहारिक कार्य में लोकप्रिय बना दिया है। समाश्रयण गुणांक β_2 चर X के संदर्भ में चर Y की लोचता, यथा चर X में एक प्रतिशत परिवर्तन के लिए चर Y में प्रतिशतता परिवर्तन को मापता है। तदनुसार, यदि चर Y किसी माँग की गई वस्तु की मात्रा और चर X उसके इकाई मूल्य को निरूपित करता हो तो चर β_2 माँग की कीमत लोचता को मापता है।

6.6 फलनिक रूप का चयन

आपने देखा ही होगा कि दो–चर समाश्रयण मॉडल के तीन फलनिक रूप हो सकते हैं, जैसा कि नीचे दर्शाया गया है –

(I) $Y_i = \beta_1 + \beta_2 X_i + u_i$

(II) $lnY_i = \beta_1 + \beta_2 X_i + u_i$

(III) $\ln Y_i = \beta_1 + \beta_2 \ln x_i + u_i$

अब एक प्रश्न यह उठता है कि इनमें सबसे अच्छा मॉडल कौन–सा है? वस्तुतः फलनिक रूप का चुनाव हमारे उद्देश्य पर निर्भर करता है। हमें वह मॉडल चुनना चाहिए जो हमें हमारे प्रश्नों का प्रासंगिक उत्तर दे सके। मान लीजिए कि हमारा उद्देश्य आश्रित चर पर स्वतंत्र चर में परिवर्तन के प्रभाव का आकलन करना है। इस उदाहरण में हम मॉडल–I का प्रयोग कर सकते हैं।

दूसरी ओर, यदि हमारा उद्देश्य स्वतंत्र चर में परिवर्तन के परिणामस्वरूप आश्रित चर में वृद्धि दर का आकलन करना हो तो हमें अर्ध–लघुगणक मॉडल (मॉडल–II) का विकल्प चुनना चाहिए। यदि हमारा उद्देश्य दो चरों के बीच लोचता का आकलन करना हो तो हम लघुगणक–रैखिक मॉडल चुनते हैं।

उक्त तीन समाश्रयण मॉडल (मॉडल–I, II, III) प्राचलों के भिन्न–भिन्न आकल देंगे। आकलकों की मानक त्रुटि भी भिन्न ही होगी। इसके अलावा, निर्धारण का गुणांक R^2 भी तीनों मॉडलों के लिए भिन्न–भिन्न ही होगा। क्या हम इन मॉडलों के R^2 की तुलना कर सकते हैं और कह सकते हैं कि उच्चतम R^2 वाला मॉडल ही सुमानक है? नहीं। हम R^2 के मान की तुलना विभिन्न आश्रित चरों वाले समाश्रयण मॉडलों से प्राप्त मान से नहीं कर सकते। बहरहाल, हम इन मॉडलों के R^2 की

तुलना एक ही आश्रित चर और एक ही आकलन विधि वाले समाश्रयण मॉडलों से कर सकते हैं। तदनुसार मॉडल–I और मॉडल–II के R^2 मान की तुलना नहीं की जा सकती है। हम मॉडल–II और मॉडल–III की तुलना उनकी सुमानकता के संदर्भ में कर सकते हैं।

यदि दो समाश्रयण मॉडल अपने निर्धारण गुणांक, आकलकों की सांख्यिकीय सार्थकता और नैदानिक परीक्षण (जिन पर चर्चा इकाई 13 और 14 में की जाएगी) के संदर्भ में लगभग एक समान हों तो हम सरलतर मॉडल को प्राथमिकता देते हैं। सरलतर मॉडल को समझना अपेक्षाकृत सरल होता है और आमतौर पर दूसरों द्वारा स्वीकार किया जाता है।

लघुगणक–रैखिक समाश्रयण मॉडल के कुछ विशिष्ट लाभ हैं, यथा –

(i) प्राचल मान परिवर्तन के प्रति अचल होते हैं क्योंकि वे प्रतिशत परिवर्तनों को मापते हैं,

(ii) मॉडल सीधे लोचता के आँकड़े देता है, और

(iii) मॉडल कुछ हद तक विषमविसारिता की समस्या को नियंत्रित कर लेता है। (विषमविसारिता की समस्या के लिए इकाई 11 देखें)।

बोध प्रश्न 2

1) अर्ध–लघुगणक मॉडल में आप समाश्रयण गुणांक की व्याख्या कैसे करते हैं? स्पष्ट करें।

....

....

....

....

....

2) लघुगणक–रैखिक समाश्रयण मॉडल के समाश्रयण प्राचल का आकलन कैसे किया जाता है? संक्षिप्त वर्णन करें।

....

....

....

....

....

3) लघुगणक–रैखिक मॉडल के क्या लाभ हैं? स्पष्ट करें।

....

....

....

....

....

4) आंतरिक रूप से रैखिक मॉडल का क्या अर्थ होता है? क्या आप किसी आंतरिक रूप से रैखिक मॉडल के परिणामों की तुलना किसी रैखिक मॉडल से कर सकते हैं? क्यों अथवा क्यों नहीं? स्पष्ट करें।

....

....

....

....

....

6.7 सार–संक्षेप

इस इकाई में हमने ऐसे फलनिक रूपों के विषय में चर्चा की जिन्हें दो–चर समाश्रयण मॉडल में समायोजित किया जा सकता हो।

हमने इकाई का आरंभ मूल बिंदु से (जहाँ कोई अवरोधन नहीं होता) गुजरने वाले समाश्रयण मॉडल के साथ किया।

हमने चरों के मान परिवर्तन के प्रभाव की ओर भी इशारा किया। इसके बाद हमने तीन फलनिक रूपों पर विचार किया – मूल मॉडल, अर्ध–लघुगणक मॉडल और लघुगणक–रैखिक मॉडल।

इकाई में इन सभी तीन फलनिक रूपों में प्राचलों की व्याख्या पर विस्तृत चर्चा की गई है।

6.8 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1. किसी भी समाश्रयण प्रतिगमन मॉडल से अवरोधन पद के बहिष्करण के गंभीर निहितार्थ होते हैं। इसे तभी छोड़ा जाना चाहिए जब अप्रतिबंधित मॉडल में अवरोधन पद सांख्यिकीय रूप से महत्वपूर्ण न हो।
2. हमने तालिका 6.1 में अवरोधन पद की चूक के निहितार्थों को सूचीबद्ध किया है। इसे ध्यानपूर्वक पढ़ें और उत्तर दें।
3. जब व्याख्यात्मक चर के मापदंड में कोई परिवर्तन होता है तो संबद्ध आकलन प्रभावित होता है। यदि X को c से गुणा किया जाता है तो प्राचल को भी c से ही विभाजित किया जाता है।
4. यदि Y को c से गुणा किया जाता है तो मॉडल के सभी प्राचलों को c से ही गुणा किया जाता है।

बोध प्रश्न 2

1. किसी भी अर्ध–लघुगणक मॉडल में समाश्रयण प्राचल विकास दर को इंगित करता है। यदि X के मान में 1 इकाई की वृद्धि होती है, तो Y का अपेक्षित मान β प्रतिशत बढ़ जाता है।
2. लघुगणक–रैखिक मॉडल का आकलन साधारण समाश्रयण मॉडल के समान ही होता है, सिवाय इसके कि चर रूपांतरित हो जाते हैं। किसी भी समाश्रयण मॉडल के आकलन में अपनाए जाने वाले चरणों को स्वयं लिखें।

3. हमने पाठ में तीन लाभों का उल्लेख किया है –

   (i) प्राचल मान परिवर्तन के प्रति अचल होते हैं क्योंकि वे प्रतिशत परिवर्तनों को मापते हैं, (ii) मॉडल सीधे लोचता के आँकड़े देता है, और
   (iii) मॉडल विषमविसारिता की समस्या को कुछ हद तक कम करता है।
4. आप दो समाश्रयण मॉडलों के परिणामों की तुलना तब तक नहीं कर सकते जब तक कि आश्रित चर समान न हों।