इग्नू
जन जन का
विश्वविद्यालय

सामाजिक विज्ञान विद्यापीठ
इन्दिरा गाँधी राष्ट्रीय मुक्त विश्वविद्यालय

BECC-110
प्रारंभिक अर्थमिति

बीईसीसी–110

प्रारंभिक अर्थमिति

सामाजिक विज्ञान विद्यापीठ
इंदिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय
मैदान गढ़ी, नई दिल्ली–110068

विशेषज्ञ समीति

प्रो० अतुल शर्मा (सेवानिवृत) पूर्व निदेशक भारतीय सांख्यिकी संस्थान, नई दिल्ली।	प्रो० एम.एस. भट्ट (सेवानिवृत) जामिया मिलिया इस्लामिया नई दिल्ली।	प्रो० गोपीनाथ प्रधान (सेवानिवृत) इंदिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय नई दिल्ली।
डॉ. इद्राणी राय चौधरी, सी.एस.आर.डी, जवाहर लाल नेहरू विश्वविद्यालय, नई दिल्ली।	डॉ० मजुँला सिंह सेंट स्टीफंस कॉलेज, दिल्ली विश्वविद्यालय नई दिल्ली।	प्रो० बी. एस. प्रकाश इंदिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय नई दिल्ली।
डॉ० अनुप चटर्जी (सेवानिवृत) ए.आर.एस.डी कॉलेज, दिल्ली विश्वविद्यालय।	डॉ. एस.पी. शर्मा, श्यामलाल कॉलेज(संध्या), दिल्ली विश्वविद्यालय।	प्रो० नारायण प्रसाद इंदिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय नई दिल्ली
बी.एस. बागला (सेवानिवृत) पीजीडीएवी कॉलेज(संध्या), दिल्ली विश्वविद्यालय।	प्रो० कौस्तुभ बारिक इंदिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय नई दिल्ली	सौगतो सेन इंदिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय नई दिल्ली

पाठ्यक्रम संयोजक समीति

खण्ड / इकाई शीर्षक		इकाई लेखन एवं हिंदी अनुवादक
खण्ड 1	**अर्थमितीय सिद्धांत : मूल तत्व**	
इकाई 1	अर्थमिति : एक परिचय	डॉ. पूजा शर्मा, सहायक प्राध्यापक, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय।
इकाई 2	सांख्यिकीय अवधारणाएँ : एक विहंगावलोकन	
इकाई 3	परिकल्पना परीक्षण : विहंगावलोकन	
खण्ड 2	**समाश्रयण मॉडल : द्वि–चरीय उदाहरण**	
इकाई 4	सरल रैखिक समाश्रयण मॉडल : आकलन	डॉ. पूजा शर्मा, सहायक प्राध्यापक, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय।
इकाई 5	सरल रैखिक समाश्रयण मॉडल : अनुमिति	
इकाई 6	द्वि–चरीय समाश्रयण मॉडल का विस्तार	प्रोफेसर कौस्तुभ बारिक, इंदिरा गाँधी राष्ट्रीय मुक्त विश्वविद्यालय।
खण्ड 3	**बहु समाश्रयण मॉडल**	
इकाई 7	बहु समाश्रयण मॉडल : आकलन	डॉ. पूजा शर्मा, सहायक प्राध्यापक, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय।
इकाई 8	बहु समाश्रयण मॉडल : अनुमिति	
इकाई 9	बहु समाश्रयण मॉडल का विस्तार : प्रतिरूप चर प्रकरण	डॉ. पूजा शर्मा, सहायक प्राध्यापक, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय एवं प्रोफेसर बी, एस, प्रकाश, इंदिरा गाँधी राष्ट्रीय मुक्त विश्वविद्यालय।
खण्ड 4	**अवधारणाओं के उल्लंघन का उपचार**	
इकाई 10	बहुसंरेखता	डॉ. पूजा शर्मा, सहायक प्राध्यापक, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय।
इकाई 11	विषमविसारिता	
इकाई 12	स्व–सहसंबंध	
खण्ड 5	**अर्थमितीय मॉडल विनिर्देशन और नैदानिक परीक्षण**	
इकाई 13	मॉडल चयन मापदंड	डॉ. सहबा फातिमा, स्वतंत्र शोधकर्ता, लखनऊ।
इकाई 14	विनिर्देशन त्रुटि हेतु परीक्षण	

पाठ्यक्रम संयोजक: प्रो० कौस्तुभ बारिक

सामग्री निर्माण

सितम्बर, 2021

ISBN:

इंदिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय की तरफ से निदेशक सामाजिक विज्ञान विद्यापीठ द्वारा मुद्रित एवं प्रकाशित।

लेज़र टाईप सेट: मुकेश कुमार यादव **आवरण सज्जा :** संदीप मैनी **मुद्रण:**

विषय–सूची

पाठ्यक्रम परिचय

अर्थमिति अर्थशास्त्र, गणित और सांख्यिकी के बीच एक अंतराफलक अर्थात् एक ऐसा समन्वय बिंदु है जहाँ वे एक–दूसरे को प्रभावित करते हैं। यह मुख्य रूप से आर्थिक सिद्धांत के अनुभवजन्य आकलन से संबंध रखता है। वर्तमान पाठयक्रम मूलभूत अर्थमितीय अवधारणाओं और तकनीकों की व्यापक भूमिका प्रस्तुत करता है। पाठयक्रम को 14 इकाइयों वाले पाँच खंडों में विभाजित किया गया है।

खण्ड 1 : अर्थमितीय सिद्धांत : मूल तत्व में तीन इकाइयाँ शामिल की गई हैं। **इकाई 1** अपनी प्रकृति में परिचयात्मक है। यह अर्थमिति को परिभाषित करती है और अर्थमितीय अध्ययन में हमारे द्वारा अनुसरण किए जाने वाले चरणों को सूचीबद्ध करती है। **इकाई 2** अर्थमिति में प्रायः प्रयोग की जाने वाली अवधारणाओं का एक सिंहावलोकन प्रस्तुत करती है। **इकाई 3** में हमने परिकल्पना परीक्षण की अवधारणा और प्रक्रिया को परिभाषित किया है।

खण्ड 2 : समाश्रयण मॉडल : दो–चर उदाहरण में तीन इकाइयाँ हैं। प्रथम, **इकाई 4**, की शुरुआत साधारण न्यूनतम वर्ग (OLS) विधि द्वारा समाश्रयण मॉडल की आकलन प्रक्रिया से की गई है। इसमें OLS आकलकों के गुणधर्मों और समाश्रयण मॉडलों की सुमानकता पर भी प्रकाश डाला गया है। **इकाई 5** सरल समाश्रयण मॉडल के साथ चर्चा जारी रखती है और परिकल्पना के परीक्षण की प्रक्रिया का वर्णन करती है। इस संदर्भ में यह समाश्रयण मॉडलों के साथ पूर्वानुमानन प्रक्रिया की भी व्याख्या करती है। **इकाई 6** लघुगुणक–रैखिक मॉडलों और किसी समाश्रयण मॉडल में चरों की मापन इकाइयों को बदलने के अनुसार सरल समाश्रयण मॉडलों का विस्तार करती है।

खण्ड 3 : बहु समाश्रयण मॉडल में उन उदाहरणों पर विचार किया गया है जिनमें एक से अधिक व्याख्यात्मक चर होते हैं। इस खंड में भी तीन इकाइयाँ हैं। **इकाई 7** बहु समाश्रयण मॉडलों के आकलन से संबंध रखती है। **इकाई 8** बहु समाश्रयण मॉडल के उदाहरण में परिकल्पना परीक्षण से संबंधित है। **इकाई 9** समाश्रयण मॉडलों की संरचनात्मक स्थिरता पर विचार करती है और बहु समाश्रयण मॉडलों में व्याख्यात्मक चरों के रूप में प्रतिरूप चर शामिल करती है।

खण्ड 4 : अवधारणाओं के उल्लंघन का उपचार अपने विषय पर चर्चा अगली तीन इकाइयों में प्रस्तुत करता है। **इकाई 10** बहुसंरेखता के विषय को लेकर चलती है। यह बहुसंरेखता के परिणामों, उसके संसूचन और उपचारात्मक उपायों की रूपरेखा प्रस्तुत करती है। **इकाई 11** विषमविसारिता – उसके परिणामों, उसके संसूचन एवं उपचारात्मक उपायों के विषय से संबंधित है। **इकाई 12** बहु समाश्रयण मॉडलों में एक अन्य महत्वपूर्ण समस्या, यथा स्वसहसंबंध, पर चर्चा करती है। यह स्वसहसंबंध के परिणामों, उसके संसूचन एवं उपचारात्मक उपायों पर विस्तृत चर्चा प्रस्तुत करती है।

खण्ड 5 : अर्थमितीय मॉडल विनिर्देशन और नैदानिक परीक्षण नामक इस अंतिम खंड में मात्र दो ही इकाइयाँ हैं। **इकाई 13** मॉडल चयन मापदंड से संबंधित है। इस इकाई में हमने प्रासंगिक चरों के बहिष्करण और अप्रासंगिक चरों के समावेश जैसे विषयों पर चर्चा की है। **इकाई 14** की विषय–वस्तु विनिर्देशन की त्रुटियों के लिए परीक्षण है। इस संदर्भ में यह इकाई एकाइके सूचना मानदंड (AIC), श्वार्ज सूचना मानदंड (SIC), और मॉलो के मानदंड की रूपरेखा प्रस्तुत करती है।

इकाई 1 अर्थमिति : एक परिचय *

इकाई की रूपरेखा

1.0 उद्देश्य

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

- अर्थशास्त्र के क्षेत्र में अर्थमिति के महत्व की व्याख्या कर सकें;
- अर्थमिति, गणितीय अर्थशास्त्र और आर्थिक सांख्यिकी के बीच अंतर स्पष्ट कर सकें;
- किसी भी अर्थमितीय अध्ययन में अनुसरण किए जाने वाले चरणों का वर्णन कर सकें; तथा
- साहचर्य (association) और कारण–कार्य संबंध (cause and effect relationship) के बीच अंतर स्पष्ट कर सकें।

1.1 विषय–प्रवेश

आर्थिक आँकड़ों के विश्लेषण हेतु विशिष्ट विधियों का अनुप्रयोग अर्थात् अर्थमिति इस यथार्थ जगत को अनेक विद्यमान आर्थिक सिद्धांतों से जोड़ती है। अर्थशास्त्र की यह शाखा आर्थिक संबंधों एवं विभिन्न आर्थिक सिद्धांतों के परीक्षण के लिए सांख्यिकीय विधियों के विकास पर आधारित है। जहाँ तक कि विभिन्न चरों के बीच संबंध का सरोकार है, अर्थमिति हमारी दो प्रकार से सहायता करती है, यथा –

1. चरों के बीच पिछले संबंध की व्याख्या करना, और
2. किसी एक चर के मानक अन्य चरों के मानों के आधार पर पूर्वानुमान व्यक्त करना।

अर्थमिति को अर्थशास्त्र, गणित एवं सांख्यिकी के बीच एक अंतराफलक माना जा सकता है।

यह मुख्य रूप से आर्थिक सिद्धांतों के अनुभवजन्य आकलन से संबंध रखती है।

एक व्यापक अर्थ में हम कह सकते हैं कि अर्थमिति सामाजिक विज्ञान की एक ऐसी शाखा है जो गणित और सांख्यिकीय आकलन के उपकरणों को परस्पर जोड़ती है, और फिर इन उपकरणों को आर्थिक दृश्य घटनाओं का विश्लेषण करने के लिए प्रयोग किया जाता है।

अर्थशास्त्र की यह शाखा समाश्रयण तकनीक का उपयोग करती है, जो विभिन्न चरों के बीच साहचर्य अथवा सह–संबंध स्थापित करती है। यह ध्यान देने की बात है कि ऐसे संबंधों का अर्थ कारण–कार्य संबंध (यथा, कारण एवं प्रभाव संबंध) नहीं होता है। कारण–कार्य संबंध की अवधारणा अर्थशास्त्र के ही किसी सिद्धांत से उत्पन्न हुई होती है।

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

1.2 अर्थमिति का अर्थ

जैसा कि ऊपर उल्लेख किया गया है, अर्थमिति 'आर्थिक माप' से संबंध रखती है। इसको सामाजिक विज्ञान की एक ऐसी शाखा के रूप में परिभाषित किया जा सकता है जो किसी भी आर्थिक दृश्यघटना का विश्लेषण करने के लिए गणित, सांख्यिकीय निष्कर्ष एवं आर्थिक सिद्धांत की तकनीकों का उपयोग करती हो।

यह आर्थिक आँकड़ों के लिए गणितीय आँकड़ों के अनुप्रयोगों से संबंध रखती है। इसका उद्देश्य गणितीय संबंध की सहायता से निर्मित आर्थिक मॉडलों को अनुभवजन्य समर्थन प्रदान करना और इस तरह से संख्यात्मक परिणाम प्राप्त करना होता है। इस प्रकार, अर्थमिति आर्थिक सिद्धांत, गणितीय अर्थशास्त्र और आर्थिक सांख्यिकी तीनों विधाओं का प्रयोग करती है।

यही कारण है कि अर्थमिति गणितीय समीकरणों एवं आर्थिक आँकड़ों के रूप में परिष्कृत गणितीय उपकरणों का प्रयोग कर आर्थिक सिद्धांत (व्यष्टि अर्थशास्त्र अथवा समष्टि अर्थशास्त्र) की अंतर्क्रिया हेतु एक मंच बन जाती है। यहाँ आर्थिक सांख्यिकी आँकड़ों के संग्रहण, प्रसंस्करण एवं प्रस्तुति द्वारा विकसित की जाती है।

गणितीय अर्थशास्त्र का मुख्य सरोकार आर्थिक सिद्धांत को गणितीय रूपों अथवा समीकरणों में व्यक्त किए जाने से होता है। इन समीकरणों को अंततः मॉडलों के रूप में व्यक्त कर दिया जाता है। ध्यान देने की बात है कि गणितीय अर्थशास्त्र किसी भी सिद्धांत की मापनीयता अथवा अनुभवजन्य सत्यापन का मूल्यांकन नहीं करता है।

आर्थिक सांख्यिकी मुख्य रूप से लेखाचित्रों, आरेखों और तालिकाओं के रूप में आर्थिक आँकड़ों के संग्रहण, प्रसंस्करण एवं प्रस्तुति से संबंध रखती है। ये आँकड़े परिवारों और फर्मों से संबंधित व्यष्टि–अर्थशास्त्रीय चर विषयक हो सकते हैं अथवा ये सकल घरेलू उत्पाद, रोजगार, कीमतों आदि समष्टि–अर्थशास्त्रीय चरों से संबंधित हो सकते हैं।

अर्थमितीय मॉडलों के लिए आँकड़े प्राथमिक आँकड़ों अथवा द्वितीयक आँकड़ों के रूप में हो सकते हैं। एक आर्थिक सांख्यिकीविद् आमतौर पर स्वयं को आँकड़ों के सारणीकरण एवं प्रसंस्करण तक ही सीमित रखता है। जबकि अर्थमिति मुख्य रूप से आर्थिक सिद्धांतों के अनुभवजन्य सत्यापन में रुचि रखती है। एक अर्थमितिविद् एकाधिक मॉडल तैयार करेगा और फिर आर्थिक सिद्धांतों का परीक्षण करेगा। गणितीय अर्थशास्त्र में संबंध नियतात्मक होता है। उदाहरण के लिए,

$$Y_i = a + bX_i \quad \ldots (1.1)$$

ऊपर दिए गए समीकरण (1.1) में,

Y व्याख्या किया जाने वाला चर है,

X व्याख्या करने वाला चर, तथा

a व b प्राचल हैं।

दूसरी ओर, अर्थमिति में संबंधों की प्रकृति प्रसंभाव्य अर्थात् अनेक संभावनाओं में से चयनित कोई एक होती है। हम समीकरण (1.1) में एक प्रसंभाव्य त्रुटि चर जोड़ते हैं। उदाहरण के लिए,

$$Y_i = a + bX_i + u_i \quad \ldots (1.2)$$

इन चरों के बीच प्रसंभाव्य संबंध पर विस्तृत चर्चा हम इकाई 4 में करेंगे। अर्थमिति में हमें आमतौर पर आर्थिक आँकड़ों की अनूठी प्रकृति के कारण विशेष तरीकों की आवश्यकता होती है क्योंकि ऐसे आँकड़े संगृहीत परीक्षणों के तहत उत्पन्न नहीं होते। अर्थमिति का उद्देश्य केवल सांख्यिकीय निष्कर्ष की तकनीक का प्रयोग कर आर्थिक सिद्धांत और वास्तविक माप के बीच की खाई को पाटना होता है। तदनुसार, अर्थमिति की तीन प्रमुख विशेषताओं पर ध्यान दिया जाना चाहिए, यथा –

- प्रथम, अर्थमिति आर्थिक संबंधों के मात्रात्मक विश्लेषण से सरोकार रखती है।
- दूसरे, यह आर्थिक सिद्धांत और तर्क पर आधारित होती है।
- तीसरे, इसमें निष्कर्ष ज्ञात करने के लिए उपयुक्त आकलन विधियों की आवश्यकता होती है।

इस प्रकार, यदि उक्त संबंध मात्रात्मक शब्दों में व्यक्त नहीं किया जाता है तो हम अर्थमितीय उपकरण प्रयोग नहीं कर सकते हैं। इसके अलावा, चर किसी सिद्धांत अथवा तर्क के अनुसार संबंध दर्शाते हैं; अन्यथा यह संबंध उस अवैध सहसंबंध के समान ही होगा जिसका हमने सांख्यिकी में अध्ययन किया था।

1.3 अर्थशास्त्र और अर्थमिति

आर्थिक सिद्धांत में विवरण अपनी प्रकृति में गुणात्मक हो सकते हैं। दूसरी ओर, जैसा कि ऊपर चर्चा की गई है, अर्थमिति गणितीय अर्थशास्त्र, आर्थिक सांख्यिकी और गणितीय सांख्यिकी का संयोजन होती है।

चलिए, एक उदाहरण लेते हैं। माँग के नियम में कहा गया है कि यदि अन्य बातें पूर्ववत् रहें तो किसी भी वस्तु की कीमत में वृद्धि से उस वस्तु की माँग की मात्रा में कमी आने की प्रत्याशा की जाती है। अतः आर्थिक सिद्धांत किसी भी जिंस की कीमत और माँग की गई मात्रा के बीच एक नकारात्मक अथवा व्युत्क्रम संबंध का पूर्वानुमान लेकर चलता है।

माँग का नियम दो चरों, यथा वस्तु की कीमत और माँग की गई मात्रा, के बीच संबंध की शक्ति की कोई संख्यात्मक माप प्रदान नहीं करता है। वह इस प्रश्न का उत्तर देने में भी विफल रहता है कि वस्तु की कीमत में एक निश्चित परिवर्तन के परिणामस्वरूप उसकी मात्रा कितनी ऊपर या नीचे जाएगी।

अर्थमिति अधिकांश आर्थिक सिद्धांतों को अनुभवजन्य विषय–वस्तु प्रदान करती है। व्यावहारिक दुनिया में अर्थशास्त्र के वास्तविक अनुप्रयोग में विभिन्न महत्वपूर्ण आर्थिक चरों का पूर्वानुमान शामिल होता है, यथा – बिक्री, ब्याज दर, मुद्रा आपूर्ति, मूल्य लोच, आदि।

जब यह समझने की बात आती है कि ज्ञात चर किसी समयावधि में कैसे व्यवहार करेंगे अथवा ये चर एक दूसरे से किस प्रकार जुड़े हैं, एक अर्थशास्त्री की भूमिका किसी भी अर्थव्यवस्था के लिए बहुत महत्वपूर्ण हो जाती है। दूसरी ओर, एक अर्थमितिविद् से उसके द्वारा माँग की गई मात्रा पर प्रस्तावित मूल्यवृद्धि के प्रभाव का आकलन किए जाने की अपेक्षा हो सकती है। उदाहरण के लिए, बिजली की कीमत में वृद्धि के प्रभाव का अनुमान किसी अर्थमितिविद् द्वारा लगाया जा सकता है और बिजली बोर्ड तदनुसार कीमत में वृद्धि कर सकता है।

बोध प्रश्न 1

1. अर्थमिति, गणितीय अर्थशास्त्र और सांख्यिकी के बीच अंतर स्पष्ट कीजिए।

   ……………………………………………………………………………

   ……………………………………………………………………………

   ……………………………………………………………………………

   ……………………………………………………………………………

2. अर्थमिति की प्रमुख विशेषताओं पर प्रकाश डालें।

   ……………………………………………………………………………

   ……………………………………………………………………………

   ……………………………………………………………………………

   ……………………………………………………………………………

1.4 अर्थमिति की कार्यप्रणाली

अर्थमिति में आमतौर पर हमारे समक्ष अनेक प्रकार के आर्थिक मुद्दे आते हैं। ये मुद्दे अर्थशास्त्र की किसी भी शाखा से हो सकते हैं, जैसे कि व्यष्टि अर्थशास्त्र, समष्टि अर्थशास्त्र, लोक अर्थशास्त्र, अंतर्राष्ट्रीय व्यापार, आदि। ये अर्थव्यवस्था के किसी भी अन्य क्षेत्र – जैसे कि कृषि, उद्योग एवं सेवाओं – से भी हो सकते हैं।

आपके समक्ष समस्या एकदम भिन्न भी हो सकती है। बहरहाल, अर्थमितीय अध्ययन में हमें कुछ सामान्य चरणों का अनुसरण अवश्य करना होता है। ये चरण निम्नवत् हैं –

1. सिद्धांत अथवा परिकल्पना कीसमुक्ति तैयार करना
2. उक्त सिद्धांत के गणितीय मॉडल की विस्तारपूर्वक व्याख्या करना
3. सांख्यिकीय अथवा अर्थमितीय मॉडल विस्तारपूर्वक व्याख्या करना
4. प्रयोजनीय आँकड़े प्राप्त करना
5. अर्थमितीय मॉडल के प्राचलों का आकलन करना
6. परिकल्पना का परीक्षण करना
7. भविष्यवाणी अथवा पूर्वानुमान व्यक्त करना
8. परिणामों की व्याख्या करना

इन आठ चरणों को कुछ और विस्तार से समझे जाने की आवश्यकता है।

चलिए, एक उदाहरण पर विचार करते हैं ताकि मुद्दों को भली भाँति समझा जा सके। जैसा कि आप प्रावेशिक समष्टि अर्थशास्त्र से जानते ही हैं, उपभोग व्यय परिवारों की आय पर निर्भर करता है।

आइए,देखें कि उपर्युक्त संबंध पर कोई अर्थमितीय अध्ययन किस प्रकार कार्यान्वित किया जा सकता है।

चरण 1: सिद्धांत अथवा परिकल्पना की समुक्ति तैयार करना

उपभोग और आय के बीच संबंध अपनी प्रकृति में जटिल होता है। ऐसे अनेक कारक हैं जो परिवार के उपभोग व्यय को प्रभावित करते हैं, यथा – परिवार का आकार, शिक्षा का स्तर, परिवार के सदस्यों की स्वास्थ्य स्थिति, रहने का स्थान (ग्रामीण/ शहरी), आदि।

एक सरल मॉडल में, तथापि, केन्जियन उपभोग फलन उपभोग व्यय और घरेलू आय के बीच संबंध स्थापित करके चलता है। कीन्स द्वारा दो अवधारणाओं का प्रयोग किया जाता है, यथा –

- औसत उपभोग प्रवृत्ति (APC),

  और

- सीमांत उपभोग प्रवृत्ति (MPC)।

कीन्स के अनुसार, आय स्तर बढ़ने पर औसत उपभोग प्रवृत्ति में गिरावट दिखाई पड़ने लगती है। उपर्युक्त समुक्तिको हम एक परिकल्पना के रूप में ले सकते हैं। याद रखें कि परिकल्पना किसी सिद्धांत अथवा तर्क पर ही आधारित होती है।

चरण 2: उक्त सिद्धांत के गणितीय मॉडल की विस्तारपूर्वक व्याख्या करना

उक्त उपभोग फलन अब निम्नलिखित रूप में दिखाई पड़ता है –

$$C_i = C_0 + cY_i \quad \dots (1.3)$$

चर C और Y क्रमशः उपभोग व्यय और आय को निर्दिष्ट करते हैं। ध्यान दें कि यहाँ C_0स्वायत्त उपभोग है, जो कि जीवित रहने के लिए आवश्यक न्यूनतम है। यदि किसी परिवार की आय शून्य हो फिर भी C_0उपभोग तो होगा ही। यह भी ध्यान देने की बात है कि औसत उपभोग प्रवृत्ति में गिरावट के लिए समीकरण (1.3) के प्राचलों को निम्नलिखित दो शर्तें पूरी करनी ही होंगी –

$C_0 > 0$ और $0 < c < 1$

उपर्युक्त दोनों शर्तें हमें गणितीय रूप में परिकल्पना की रचना करने में मदद करेंगी।

चरण 3: सांख्यिकीय अथवा अर्थमितीय मॉडल विस्तारपूर्वक व्याख्या करना

समीकरण (1.3) में निर्दिष्ट उपभोग आय संबंध अपनी प्रकृति में सटीक है। यदि हम समीकरण (1.4) के लिए आरेख आलेखित करते हैं तो हमें एक ऋजु रेखा प्राप्त होगी। जैसा कि पहले उल्लेख किया गया है, अर्थमिति में संबंधों की प्रकृति *प्रसंभाव्य* होती है।

आइए, एक समान आय स्तर वाले दो परिवारों पर विचार करते हैं। आय के अलावा कुछ अन्य कारकों – जैसे परिवार के सदस्यों की स्वास्थ्य स्थिति – के कारण उनका उपभोग व्यय एक दूसरे से भिन्न होगा। ऐसे कारकों को शामिल करने के लिए हम अपने मॉडल में एक अन्य चर, u_i, को शामिल करते हैं। इस चर को कुछ शर्तें पूरी करनी होंगी (जिन पर चर्चा इकाई 4 में की जाएगी)। इस प्रकार उपभोग फलन का अर्थमितीय विनिर्देशन निम्नवत् होगा –

$$C_i = C_0 + cY_i + u_i \quad \dots (1.4)$$

चरण 4: प्रयोजनीय आँकड़े प्राप्त करना

आँकड़े प्राथमिक स्रोतों अथवा द्वितीयक स्रोतों से प्राप्त किए जा सकते हैं। प्राथमिक आँकड़ों एवं द्वितीयक आँकड़ों संबंधी विस्तृत विवरण के लिए पाठ्यक्रम 'BECC 107: अर्थशास्त्र हेतु सांख्यिकीय विधियाँ' की इकाई 1 का संदर्भ लें। उक्त इकाई में नमूना सर्वेक्षण की प्रक्रिया और द्वितीयक आँकड़ों के महत्वपूर्ण स्रोतों पर चर्चा की गई है।

समीकरण (1.4) में दिए गए हमारे अर्थमितीय मॉडल के आकलन के लिए हमें दो चरों – यथा, आय (Y) और उपभोग व्यय (C) – पर आँकड़ों की आवश्यकता होती है। जैसा कि आप जानते हैं, आय और व्यय प्रवाह चर होते हैं। तदनुसार, हमें इन चरों के लिए एक समयावधि निर्दिष्ट करनी होगी। मापन में सुविधा की दृष्टि से हम मासिक आय और मासिक व्यय ले सकते हैं। दूसरे, हमें यह परिभाषित करना होगा कि किसी परिवार का गठन किन से होता है – घर के वे सभी जो सदस्य हैं और घर के वे सभी जो सदस्य नहीं हैं। तीसरे, हमें एकत्र किए गए आँकड़ों की प्रकृति के आधार पर निर्णय लेना होगा।

जैसा कि आप जानते हैं कि आँकड़े चार प्रकार के होते हैं, यथा –

1. समय–शृंखला आँकड़े (time series data),
2. प्रतिनिध्यात्मक आँकड़े (cross section data),
3. समूहीकृत आँकड (pooled data), और
4. पट्टिका आँकड़े (panel data)।

1. **समय–शृंखला आँकड़े**

   किसी चर पर समय–शृंखला आँकड़े किसी समयावधि विशेष में नियमित रूप से एकत्र किए जाते हैं। कुछ ऐसे चर हैं जिन पर आँकड़े दैनिक आधार पर उपलब्ध होते हैं, जैसे सेंसेक्स और निफ्टी।

   कुछ अन्य चरों के मामले में ये मासिक आधार पर (जैसे, उपभोक्ता मूल्य सूचकांक), तिमाही आधार पर (जैसे, जीडीपी) अथवा वार्षिक आधार पर (जैसे, राजकोषीय घाटा) उपलब्ध होते हैं।

2. **प्रतिनिध्यात्मक आँकड़े**

   प्रतिनिध्यात्मक आँकड़ों का अर्थ होता है – एक समय–बिंदु पर अनेक चरों पर आँकड़े। उदाहरण के लिए, एक नमूना सर्वेक्षण के माध्यम से हम व्यय, आय, बचत, ऋण आदि पर घरेलू आँकड़े एकत्र कर सकते हैं।

   याद रखें कि समय–शृंखला आँकड़े किसी चर पर एक समयावधि विशेष में नियमित रूप से एकत्र किए जाते हैं, जबकि प्रतिनिध्यात्मक आँकड़े एक ही समय–बिंदु पर अनेक चरों पर केंद्रित होता है।

   जनगणना के आँकड़े प्रतिनिध्यात्मक आँकड़ों का ही एक उदाहरण है।

3. **समूहीकृत आँकड़े**

   एकत्र किए गए समूहीकृत आँकड़ों में हमारे पास समय–शृंखला आँकड़े और प्रतिनिध्यात्मक आँकड़े दोनों के तत्व होते हैं। वस्तुतः यह प्रतिनिध्यात्मक आँकड़ों की ही एक समय शृंखला होती है। प्रत्येक प्रतिनिधि में समुक्तियाँ संभवतः एक ही इकाई को इंगित न करती हों।

   चलिए, एक उदाहरण पर विचार करते हैं। भारत में जनगणना के आँकड़े दशवार्षिक रूप से एकत्र किए जाते हैं। फिर भी प्रत्येक जनगणना में परिवारों की संख्या अलग–अलग होती है। इस तरह के आँकड़ों को समय के साथ जनसंख्या के अभिलक्षणों में बदलाव का विश्लेषण करने के लिए एकत्र किया जा सकता है।

   आप समूहीकृत आँकड़ों के कई अन्य उदाहरणों के बारे में सोच सकते हैं। ये उदाहरण रोजगार और बेरोजगारी सर्वेक्षण, कार्यबल भागीदारी दरें, मानव विकास सूचकांक आदि हो सकते हैं।

4. **पट्टिका आँकड़े**

   ये एक विशेष प्रकार के समूहीकृत आँकड़े होते हैं। यहाँ अनेक समय–बिंदुओं पर एक ही नमूना इकाइयों पर समुक्तियाँ एकत्र की जाती हैं।

   मान लीजिए कि हम शेयर बाजार में शेयरों पर लाभ की परिवर्तनशीलता का विश्लेषण करना चाहते हैं। हम यहाँ 50 पब्लिक लिमिटेड कंपनियों का नमूना ले सकते हैं और अगस्त 2021 माह के लिए उनके निवल संपत्ति मूल्य (NAV) का अवलोकन प्रतिदिन कर सकते हैं।

   तदनुसार हमें 50 फर्मों के 31 प्रतिनिधि (अगस्त के महीने में 31 दिन) प्राप्त होते हैं। यही प्रतिनिधि पट्टिका आँकड़ों का गठन करते हैं।

यदि सभी समुक्तियाँ (समयावधि 1 से t तक के लिए और नमूना इकाइयों 1 से n तक के लिए) उपलब्ध हों तो हम इसे एक 'संतुलित पट्टिका' कहते हैं। यदि इनमें कुछ समुक्तियाँ गायब हों तो हम इसे एक 'असंतुलित पट्टिका' कहते हैं।

चरण 5: अर्थमितीय मॉडल के प्राचलों का आकलन करना

हमने BECC 107 के खंड 4 में नमूनाकरण प्रक्रिया, सांख्यिकीय अनुमान और परिकल्पना के परीक्षण के विषय में चर्चा की है। अब हमें इन अवधारणाओं को गहराई से समझने की आवश्यकता होगी। याद रखें कि अर्थमितीय आकलन में समीकरणों की संख्या प्राचलों की संख्या से कहीं अधिक होती है। ऐसे मॉडलों का आकलन करने के लिए हमें कुछ आकलन विधियों की आवश्यकता होती है।

जैसा कि आप इस पाठ्यक्रम की बाद की इकाइयों में पढ़ेंगे, आकलन करने की कुछ ही विधियाँ प्रचलित हैं। आपको पाठ्यक्रम 'BECC 107: अर्थशास्त्र हेतु सांख्यिकीय विधियाँ' की इकाई 5 में न्यूनतम वर्ग विधि से परिचित कराया गया है। आकलन के प्रयोजन से कुछ अर्थमितीय सॉफ्टवेयर उपलब्ध भी हैं। अर्थमितीय सॉफ्टवेयर के विषय में हम पाठ्यक्रम 'BECE 142: व्यावहारिक अर्थमिति' में पढ़ेंगे।

चरण 6: परिकल्पना का परीक्षण करना

एक बार जब हम प्राचलों के आकलन प्राप्त कर लेते हैं तो परिकल्पना के परीक्षण की आवश्यकता होती है। जैसा कि आप जानते हैं, किसी भी आकलक के नमूना वितरण में आकलन भिन्न–भिन्न नमूनों में भिन्न–भिन्न होता है। आपने जो आकलन किया हो वह संयोग की बात हो सकती है, और आपका प्राचल उस आकलन से काफी भिन्न भी हो सकता है। हमें यह पुष्टि करने की आवश्यकता होगी कि क्या प्राचल और आकलन के बीच का अंतर वास्तव में विद्यमान है अथवा यह नमूना चयन में उतार–चढ़ाव का मामला है।

उपभोग फलन (1.4) के लिए हमें स्थिति $C_0 > 0$ के परीक्षण के लिए एकपक्षीय t–परीक्षण प्रयोग करना होगा।

सीमांत उपभोग प्रवृत्ति के लिए हमें द्विपक्षीय t–परीक्षण $H_0 : c = 0$ प्रयोग करना होगा।

दोनों प्राचलों का एक साथ परीक्षण करने के लिए हमें F–परीक्षण प्रयोग करना होगा।

मॉडल के सही विनिर्देशन की जाँच करने की भी आवश्यकता होती है।

यहाँ दो प्रश्नों के उत्तर महत्वपूर्ण होते हैं, यथा –

1. समाश्रयण मॉडल में कितने व्याख्यात्मक चर होने चाहिए, और
2. इस मॉडल का फलनिक रूप क्या है?

यह उपभोग फलन (देखें समीकरण (1.4)) दो–चर समाश्रयण मॉडल का एक उदाहरण है। उक्त मॉडल में एक व्याख्या किया गया चर और एक व्याख्यात्मक चर है।

यदि हम व्याख्यात्मक चर (जैसे शिक्षा, आवासीय क्षेत्र का प्रकार, आदि) अपेक्षाकृत अधिक संख्या में शामिल करते हैं तो यह एक बहुरेखीय समाश्रयण मॉडल बन जाता है। इसका फलनिक रूप फिर से रेखीय या गैर–रेखीय हो सकता है।

चरण 7: भविष्यवाणी अथवा पूर्वानुमान व्यक्त करना

आकलित मॉडल का प्रयोग पूर्वानुमान लगाने अथवा भविष्यवाणी के लिए किया जा सकता है। हमारे पास आश्रित चर का वास्तविक मान है। आकलित समाश्रयण मॉडल के आधार पर हम आश्रित चर

का आकलित मूल्य प्राप्त करते हैं। इन दोनों के बीच कोई भी विसंगति पूर्वानुमान की त्रुटि कहलाती है। पूर्वानुमान की इस त्रुटि का यथासंभव छोटा होना अपेक्षित होता है।

चरण 8: परिणामों की व्याख्या करना

किसी भी आकलन की सही–सही व्याख्या किए जाने की आवश्यकता होती है। इस पाठ्यक्रम की बाद की इकाइयों में हम मॉडल विनिर्देशन एवं परिणाम की व्याख्या जैसे विषयों पर चर्चा करेंगे। अपने आकलित मॉडल का प्रयोग नीति अनुशंसाओं के लिए भी किया जा सकता है।

1.5 साहचर्य और कारण–कार्य संबंध

जैसा कि आप पाठ्यक्रम 'BECC 107: अर्थशास्त्र हेतु सांख्यिकीय विधियाँ' से जानते हैं, सहसंबंध (correlation) का तात्पर्य दो चरों के बीच संबंध से होता है। तकनीकी रूप से हम किन्हीं भी दो चरों के बीच साहचर्य गुणांक का पता लगा सकते हैं। उदाहरण के लिए, इग्नू पुस्तकालय में आने वाले छात्रों की संख्या और दिल्ली में सड़क दुर्घटनाओं की संख्या।

कुछ मामलों में हम सहसंबंध गुणांक भी उच्च पाते हैं। हालाँकि चरों के बीच इस तरह के संबंध नकली साहचर्य की ओर ले जाते हैं। यदि हम दो ऐसे चर लेते हैं जहाँ सहसंबंध गुणांक उच्च हो और फिर समाश्रयण (regression) विश्लेषण करते हैं तो हम अपने आकलन को सांख्यिकीय रूप से महत्वपूर्ण पाएँगे। ऐसी समाश्रयण पद्धतियाँ अर्थहीन होती हैं।

इस प्रकार समाश्रयण विश्लेषण एक चर के दूसरे चर से साहचर्य (association) अथवा एक चर की दूसरे चर पर निर्भरता से संबंध रखता है। फिर भी इसका अभिप्राय 'कारण–कार्य संबंध' (cause and effect relationship) से नहीं होता। कार्य–कारण की अवधारणा अर्थशास्त्र में विद्यमान सिद्धांतों से उत्पन्न होनी चाहिए। अतएव, कोई भी सांख्यिकीय संबंध केवल सांख्यिकीय रूप से सशक्त अथवा विचारोत्तेजक हो सकता है।

जब तक चरों के बीच कार्य–कारण सिद्धांत स्थापित नहीं किया जाता है, तब तक आर्थिक सिद्धांत के परीक्षण के प्रयोजन का कोई अर्थ नहीं होगा। अधिकांश आर्थिक सिद्धांत इस परिकल्पना का परीक्षण करते हैं कि क्या एक चर का दूसरे चर पर कोई कार्य–कारण प्रभाव पड़ता है। इस प्रकार समाश्रयण विश्लेषण में तर्क अथवा आर्थिक सिद्धांत बहुत महत्वपूर्ण होता है। हमें चरों के बीच संबंध के लिए तर्क स्थापित किए बिना समाश्रयण नहीं चलाना चाहिए।

आइए, अब हम माँग के नियम के विषय पर चर्चा करते हैं। उपभोक्ता माँग का विश्लेषण करते समय हमें अन्य कारकों जैसे कि आय, अन्य वस्तुओं की कीमत, व्यक्तियों की अभिरुचि एवं अधिमानों को अपरिवर्तित रखने के लिए माँग की गई मात्रा पर माल की कीमत में परिवर्तन के प्रभाव को समझने की आवश्यकता होती है। बहरहाल, यदि अन्य कारकों को तय नहीं किया जाता है तो माँग की गई मात्रा पर मूल्य परिवर्तन के कार्य–कारण प्रभाव को जानना असंभव होगा।

बोध प्रश्न 2

1. किसी अर्थमितीय अध्ययन में आप किन चरणों का अनुसरण करेंगे? स्पष्ट कीजिए।

   ....

   ....

   ....

   ....

2. मान लीजिए कि आपको केन्जियन उपभोग फलन पर एक अर्थमितीय अध्ययन करना है। उन चरणों को लिखिए जिनका आप अनुसरण करेंगे।

…………………………………………………………………………

…………………………………………………………………………

…………………………………………………………………………

…………………………………………………………………………

3. कारण और प्रभाव संबंध से आप क्या समझते हैं? यह साहचर्य से किस प्रकार भिन्न होता है?

…………………………………………………………………………

…………………………………………………………………………

…………………………………………………………………………

…………………………………………………………………………

1.6 सार–संक्षेप

इस इकाई में हमने अर्थशास्त्र के क्षेत्र में अर्थमिति के महत्व पर चर्चा की। अर्थमिति यथार्थ जगत को सिद्धांत से जोड़ती है। यह हमें सिद्धांत की वैधता का पता लगाने में भी मदद करती है।

प्रत्येक अर्थमितीय मॉडल के पीछे कोई तर्क होना चाहिए। चरों के बीच संबंध किसी आर्थिक सिद्धांत अथवा तर्क से ही उत्पन्न होना चाहिए। किसी समाश्रयण मॉडल का मात्र आकलन कर लेना अर्थहीन परिणाम ही दे सकता है। इस इकाई में हमने अर्थमितीय विश्लेषण करने के चरणों का भी वर्णन किया है। कोई भी अर्थमितीय अध्ययन करते समय हमें इन्हीं आठ चरणों का पालन करना चाहिए।

1.7 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1. पाठांश 1.2 में हमने दर्शाया है कि अर्थमिति और अर्थशास्त्र, सांख्यिकी एवं गणितीय अर्थशास्त्र के बीच एक अंतराफलक होता है। उसी को विस्तार से समझाएँ।
2. अर्थमिति की तीन प्रमुख विशेषताएँ हैं। सबसे पहले, अर्थमिति आर्थिक संबंधों के मात्रात्मक विश्लेषण से संबंधित है। दूसरे, यह आर्थिक सिद्धांत और तर्क पर आधारित है। तीसरे, इसमें निष्कर्ष निकालने के लिए उपयुक्त आकलन विधियों की आवश्यकता होती है।

बोध प्रश्न 2

1. आपको पाठांश 1.4 में उल्लिखित आठ चरणों की व्याख्या करनी चाहिए।
2. आपको पाठांश 1.4 में दिए गए आठ चरणों का पालन करना चाहिए।

   आपके उत्तर में निम्नलिखित बातें शामिल हो सकती हैं –

   - सिद्धांत की समुक्ति

     $0 < MPC < 1$
   - मॉडल का गणितीय विनिर्देशन

     $C = \beta_1 + \beta_2 Y, 0 < \beta_2 < 1$
   - अर्थमितीय विनिर्देशन मॉडल

     $C = \beta_1 + \beta_2 Y + u$

- आँकड़ों का संग्रह
  आरबीआई हैंडबुक ऑफ स्टैटि्स्टक्स से प्राप्त द्वितीयक आँकड़े
- प्राचल आकलन
  $$\widehat{C_i} = -184.08 + 0.7164Y_i$$
- परिकल्पना परीक्षण
  $\beta_1 > 0$ अथवा $\beta_2 > 0$
- पूर्वानुमान : Y का मान ज्ञात होने पर C का मान क्या होगा?

3. समाश्रयण विश्लेषण एक चर के दूसरे चर से संबंध अथवा एक चर की दूसरे चर पर निर्भरता से संबंध रखता है। यह कार्य–कारण नहीं दर्शाता है।

   कार्य–कारण की अवधारणा बाहरी आँकड़ों से ही उत्पन्न होती है। यह अर्थशास्त्र में कोई विद्यमान सिद्धांत हो सकता है। अतएव, कोई सांख्यिकीय संबंध केवल सांख्यिकीय रूप से सशक्त अथवा विचारोत्तेजक हो सकता है।

   अधिकांश आर्थिक सिद्धांत इस परिकल्पना का परीक्षण करते हैं कि क्या एक चर का दूसरे चर पर कोई कार्य–कारण प्रभाव पड़ता है। वास्तव में, समाश्रयण दो या दो से अधिक चरों के बीच संबंध के विषय में ही बतलाता है; यह साहचर्य विचारोत्तेजक हो सकता है। जब तक चरों के बीच कोई कारणता स्थापित नहीं की जाती है, तब तक आर्थिक सिद्धांत के परीक्षण के प्रयोजन का कोई अर्थ नहीं होगा।

इकाई 2 सांख्यिकीय अवधारणाएँ : एक विहंगावलोकन *

इकाई की रूपरेखा

2.0 उद्देश्य

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

- प्रायिकता बंटन की अवधारणा और उसके महत्व की व्याख्या कर सकें;
- विभिन्न प्रकार के प्रायिकता बंटनों की पहचान कर सकें;
- सामान्य, t, F और काई–वर्ग जैसे विभिन्न प्रायिकता बंटनों के गुणधर्मों का वर्णन कर सकें;
- प्राचलों के आकलन (estimation) की प्रक्रिया को स्पष्ट कर सकें; तथा
- एक अच्छे आकलक (estimator) के गुणों का वर्णन कर सकें।

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

2.1 विषय–प्रवेश

अर्थमिति के उपकरणों को समझने में सांख्यिकीय अवधारणाएँ और आकलन विधियाँ अत्यधिक महत्व रखती हैं। अतः आपको विभिन्न अवधारणाओं को परिभाषित करने और उनके बीच अंतर करने में सक्षम होना चाहिए। अर्थमिति का सार अनुभवजन्य विश्लेषण पर आधारित है, जो कि आँकड़ों से संबंध रखता है। वास्तव में, अर्थमितीय विश्लेषण के उपकरण सांख्यिकीय विधियों से ही उद्भूत होते हैं।

सांख्यिकीय अवधारणाएँ अनिश्चितता की विद्यमानता में निर्णय लेने के लिए हमारा मार्गदर्शन करती हैं। सांख्यिकी आँकड़े संग्रह करने की विधियों के लिए मंच प्रदान करती है, जो कि अर्थमितीय विश्लेषण करने हेतु आधार प्रदान करता है। अर्थशास्त्रियों को विशाल जनसंख्या के साथ काम करने की आवश्यकता पड़ती है, जो कि एक चुनौती के रूप में सामने आता है। अतएव, उपयुक्त प्रतिदर्श (sample) अर्थात् नमूने का चयन करने और प्रायिकता बंटन के आधार पर उपयुक्त निष्कर्ष (inference) निकालने की आवश्यकता होती है।

अर्थमिति आपसे सांख्यिकीय अवधारणाओं की गहरी समझ रखे जाने की अपेक्षा करती है, ताकि अर्थशास्त्रियों को सही प्रतिदर्श चुनने और चुने हुए प्रतिदर्श से सही निष्कर्ष निकालने में मदद मिल सके। जनसंख्या वस्तुओं, घटनाओं अथवा लोगों के किसी भी संग्रह को कहा जा सकता है। जनसंख्या के प्रत्येक अवयव की जाँच करना कठिन होता है। अतः जनसंख्या (population) का कोई उपसमुच्चय लेना और फिर उसकी जाँच करना ही औचित्यपूर्ण लगता है। जनसंख्या के इस उपसमुच्चय को 'प्रतिदर्श' कहा जाता है, जिसका उपयोग आगे निष्कर्ष निकालने के लिए किया जाता है।

यदि प्रतिदर्श यादृच्छिक और पर्याप्त रूप से बड़ा हो तो प्रतिदर्श से एकत्र की गई जानकारी का उपयोग जनसंख्या के विषय में अनुमिति अथवा आकलन करने के लिए किया जा सकता है। ऐसा कोई भी परीक्षण जो यादृच्छिक परिणाम देता हो उसे 'यादृच्छिक परीक्षण' कहा जाता है। ऐसा कोई भी चर जो ऐसे मान लेता हो जो यादृच्छिक प्रक्रिया का परिणाम हों तो उसे 'यादृच्छिक चर' कहा जाता है। इस प्रकार, किसी भी यादृच्छिक चर के लिए प्रत्येक परिणाम घटना की निश्चित प्रायिकता से जुड़ा होता है।

जब हम यादृच्छिक चर परिमित मान (finite values) लेते हैं तो वे 'असतत यादृच्छिक चर' हो जाते हैं। यदि कोई यादृच्छिक चर किन्हीं दो बिंदुओं के बीच अनंत संख्या में मान ग्रहण करता है तो फिर इसे ''सतत यादृच्छिक चर' कहा जाता है।

यादृच्छिक चर एक प्रायिकता बंटन दर्शाते हैं। यदि यादृच्छिक चर असतत होता है तो इससे जुड़े प्रायिकता फलन को 'प्रायिकता बंटन फलन' कहा जाता है। दूसरी ओर, यदि यादृच्छिक चर सतत हो तो प्रायिकता फलन को 'प्रायिकता घनत्व फलन' कहा जाता है।

यादृच्छिक चरों में उनकी प्रायिकताओं के आधार पर बंटन फलनों की विविधता हो सकती है। इस इकाई में आमतौर पर प्रयोग किए जाने वाले कुछ बंटन फलनों का ही वर्णन किया गया है।

2.2 सांख्यिकीय अनुमिति

अपने पाठ्यक्रम BECC 107 में हमने सांख्यिकीय अनुमिति (statistical inference) की प्रक्रिया पर विस्तार से चर्चा की है (संदर्भ के लिए BECC 107 की इकाई 13 और 14 देखें)। सांख्यिकीय अनुमिति किसी जनसंख्या से लिए गए किसी नमूने में निहित जानकारी के आधार पर उस जनसंख्या के अभिलक्षणों के विषय में निष्कर्ष निकालने की विधि को कहा जाता है।

याद रखें कि जनसंख्या माध्य हमें भले ही ज्ञात न हो, प्रतिदर्श माध्य तो हम जानते ही हैं। सांख्यिकीय अनुमिति में हम दो प्रकार के प्रश्नों का उत्तर देने में रुचि रखते हैं। प्रथम, जनसंख्या

माध्य का मान क्या होगा? इसका उत्तर जनसंख्या माध्य के विषय में सूचित अनुमान लगाने में निहित होता है।

सांख्यिकीय अनुमिति के इस पहलू को 'आकलन' कहा जाता है। दूसरा प्रश्न जनसंख्या माध्य के विषय में प्रस्तुत किए गए कुछ निश्चित अभिकथनों से संबंधित होता है।

मान लीजिए कि किसी बिजली के बल्ब के निर्माता का दावा है कि उसके बिजली के बल्बों का औसत जीवन 2000 घंटों के बराबर है। प्रतिदर्श आँकड़ों के आधार पर क्या हम कह सकते हैं कि कथन सही नहीं है?

सांख्यिकीय अनुमिति के इस पहलू को 'परिकल्पना परीक्षण' कहा जाता है। इस प्रकार सांख्यिकीय अनुमिति दो बातों से संबंध रखती है, यथा –

1. आकलन (estimation), और
2. परिकल्पना परीक्षण (hypothesis testing)।

हम इकाई में हम प्राचलों के आकलन के विषय में ही चर्चा करेंगे। परिकल्पना परीक्षण पर चर्चा इकाई 3 में की जाएगी।

यदि 28 कंपनियों का प्रत्याशित मूल्य–अर्जन (price-earning) अनुपात 23 : 25 हो तो इस प्रतिदर्श औसत का प्रयोग पूँजी के जनसंख्या औसत के आकलन के रूप में किया जा सकता है। जैसा कि आप जानते हैं, नमूना औसत (या, नमूना माध्य) चर $\bar{X}$ से निरूपित किया जाता है।

इस प्रतिदर्श माध्य का आकलन X के प्रत्याशित मान के रूप में किया जा सकता है, जो कि जनसंख्या माध्य है। प्रतिदर्श मान $(\bar{X})$ से जनसंख्या मान E(X) तक सामान्यीकरण की यह प्रक्रिया ही सांख्यिकीय अनुमिति का सार है।

सांख्यिकीय अनुमिति का उद्देश्य नमूने से जनसंख्या के अभिलक्षणों को समझना होता है। ये जनसांख्यिकीय अभिलक्षण जनसंख्या के 'प्राचल' कहलाते हैं और प्रतिदर्श के अभिलक्षण 'आँकड़े' कहे जाते हैं।

प्रतिदर्श का प्रयोग कर जनसंख्या प्राचल निर्धारित करने और परिकलित करने की विधि को *आकलन* कहा जाता है।

2.3 केंद्रीय सीमा प्रमेय

जब यादृच्छिक चरों के फलन स्वतंत्र और अभिन्न रूप से बंटित होते हैं तो जैसे–जैसे प्रतिदर्श आकार बढ़ता है, प्रतिदर्श माध्य सामान्य रूप से जनसंख्या माध्य के आसपास बंटित किया जाता है और प्रतिदर्श आकार '*n*' बढ़ने पर मानक विचलन कम हो जाता है।

यदि X_1, X_2, X_3, ... और X_n स्वतंत्र हों और अभिन्न रूप से माध्य μ और मानक विचलन σ के साथ बंटित हों तो प्रतिदर्श माध्य $(\bar{X})$) निम्नलिखित समीकरण से दर्शाया जाता है –

$$\bar{X} = \frac{(X_1 + X_2 + ...+ X_n)}{n} \quad ... (2.1)$$

केंद्रीय सीमा प्रमेय का अर्थ होता है कि प्रत्याशित प्रतिदर्श माध्य और मानक विचलन (SD) निम्नानुसार अभिसरण करेंगे –

$E(\bar{X}) = \mu$

तथा

$$SD(\bar{X}) = \frac{\sigma}{\sqrt{n}} \quad ... (2.2)$$

केंद्रीय सीमा प्रमेय (CLT) में कहा गया है कि

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \rightarrow N\,(0, 1)$$

क्योंकि

$$n \rightarrow \infty \qquad \text{... (2.3)}$$

इसका अर्थ यह भी होता है कि

$$\bar{X} \rightarrow N\,(\mu, \frac{\sigma^2}{n})$$

दूसरे शब्दों में, प्रतिदर्श माध्य को माध्य μ और मानक विचलन $\frac{\sigma^2}{n}$ के साथ किसी प्रसामान्य यादृच्छिक चर के साथ सन्निकटित किया जा सकता है।

आगे हम कुछ महत्वपूर्ण प्रायिकता बंटन फलनों पर चर्चा करेंगे।

2.4 प्रसामान्य बंटन

प्रसामान्य बंटन (जिसे z–बंटन भी कहा जाता है) एक सतत प्रायिकता बंटन फलन होता है। केंद्रीय सीमा प्रमेय के कारण यह फलन बहुत उपयोगी है।

इसका अर्थ यह है कि स्वतंत्र बंटनों से स्वतंत्र रूप से लिए गए यादृच्छिक चरों की समुक्तियों के प्रतिदर्शों के मध्यमान सामान्य बंटन के लिए अभिसरण करते हैं। जब समुक्तियों की संख्या पर्याप्त रूप से बड़ी होती है तो यह सामान्य रूप से बंटित हो जाता है।

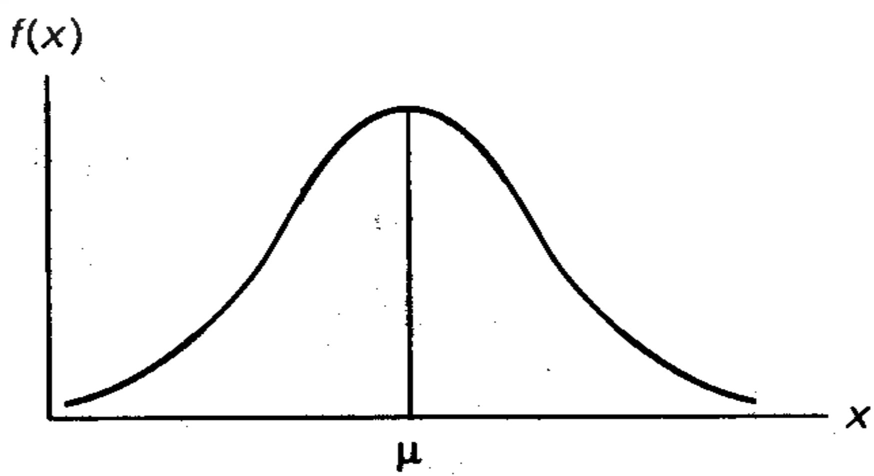
प्रसामान्य बंटन को 'घंटी वक्र' भी कहा जाता है (देखें चित्र 2.1)। प्रसामान्य बंटन का प्रायिकता घनत्व फलन (pdf) निम्नवत् होता है –

$$f(\,x\,|\,\mu, \sigma^2\,) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \text{... (2.4)}$$

जहाँ, μ बंटन अथवा माध्य की प्रत्याशा है,

σ मानक विचलन है, और

σ^2 विचरण है।



चित्र 2.1: प्रसामान्य प्रायिकता बंटन

प्रसामान्य बंटन के कुछ महत्वपूर्ण गुणधर्म निम्नवत् हैं –

1. प्रसामान्य बंटन वक्र घंटी के आकार का होता है।
2. प्रसामान्य वक्र माध्य μ के आसपास ही सममित होता है।
3. वक्र के नीचे का कुल क्षेत्रफल 1 के बराबर होता है।
4. वक्र का क्षेत्रफल पूर्णतः उसके माध्य एवं मानक विचलन से दर्शाया जाता है।

मानक प्रसामान्य बंटन : N ~ N (0,1)

यदि यह माध्य शून्य ($\mu = 0$) और इकाई विचरण ($\sigma^2 = 1$) के साथ एक प्रसामान्य बंटन होता है तो प्रायिकता बंटन फलन निम्नानुसार दर्शाया जाता है –

$$f(x \mid 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \text{... (2.5)}$$

ऊपर उल्लिखित प्रसामान्य बंटन के सभी गुणधर्म मानक प्रसामान्य बंटन होने की स्थिति में लागू होते हैं।

बोध प्रश्न 1

1. मान लीजिए कि X को सामान्य रूप से माध्य $\mu = 30$ मानक विचलन $\sigma = 4$ के साथ बंटित किया गया है। अब $P(X < 40)$ ज्ञात करें।

....

....

....

....

....

2. प्रसामान्य प्रायिकता बंटन के महत्वपूर्ण गुणधर्मों पर प्रकाश डालें।

....

....

....

....

....

3. किसी कंपनी द्वारा उत्पादित एक बिजली के बल्ब का जीवन 12 माह के औसत और 2 माह के मानक विचलन के साथ प्रसामान्य बंटन का अनुसरण करता है। कंपनी द्वारा उत्पादित बल्ब के चलने की प्रायिकता ज्ञात कीजिए –

   a) 7 माह से कम

   ....

   ....

   ....

   ....

   ....

b) 7 व 12 माह के बीच

....

....

....

....

....

2.5 काई–वर्ग बंटन

मान लीजिए कि X माध्य μ और मानक विचलन σ के साथ एक प्रसामान्य चर है। तब

$$z = \frac{X - \mu}{\sigma}$$

एक मानक प्रसामान्य चर, यथा $z \sim \mathrm{N}(0,1)$ होगा।

यदि हम z का वर्ग लेते हैं, यथा

$$z^2 = \left(\frac{\mathrm{X} - \mu}{\sigma}\right)^2$$

तो z^2 को एक डिग्री की स्वतंत्रता के साथ एक चर χ^2 के रूप में बंटित माना जाता है और पद χ_1^2 के रूप में व्यक्त किया जाता है।

यह स्पष्ट है कि चूँकि χ_1^2 एक वर्गित पद है; चर z को $-\infty$ और $+\infty$ के बीच रखने के लिए, पद χ_1^2 संख्या 0 और ∞ के मध्य स्थान लेगा (क्योंकि कोई भी वर्गित पद ऋणात्मक मान नहीं ले सकता)।

पुनः, चूँकि z का माध्य शून्य के बराबर है, चर z द्वारा लिए गए अधिकांश मान शून्य के निकट ही होंगे। परिणामतः, पद χ_1^2 का प्रायिकता घनत्व भी अधिकतम शून्य के निकट ही होगा।

ऊपर उल्लिखित परिणाम को सामान्यीकृत करते हुए, यदि $z_1, z_2, \ldots, z_k$ स्वतंत्र मानक प्रसामान्य चर हों तो पद

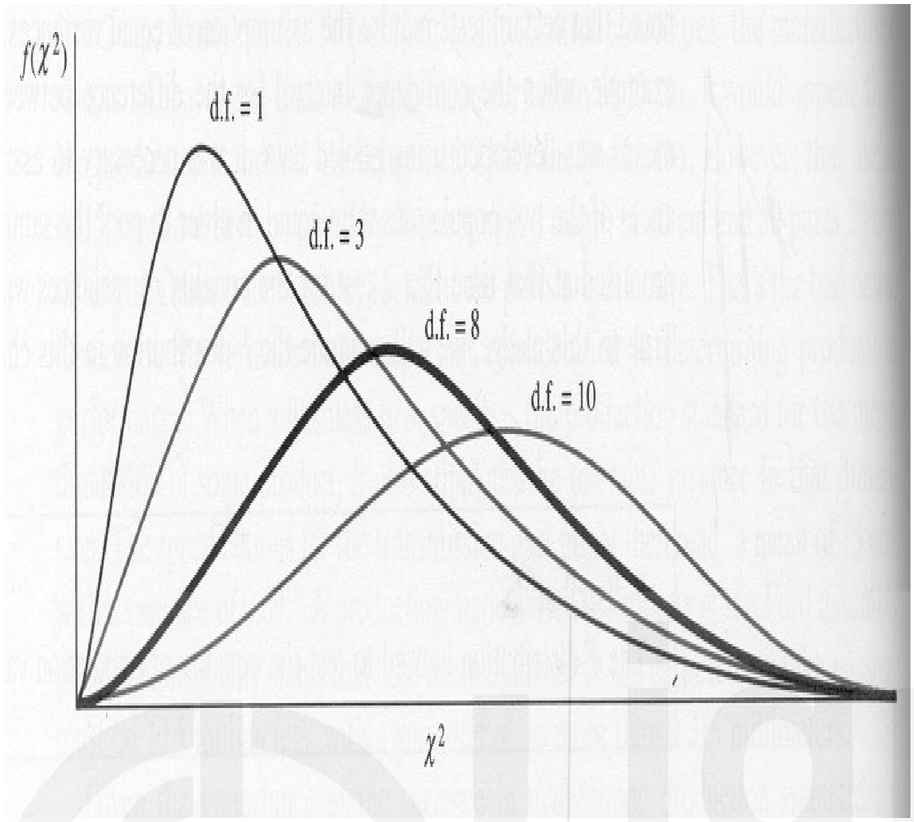
$$z = \sum_{i=1}^{ki} z_1^2$$

को एक स्वतंत्रता की k डिग्री वाला χ^2 चर कहा जाता है और पद χ_k^2 से निरूपित किया जाता है। नीचे दिया गया चित्र 2.2 स्वतंत्रता की विभिन्न कोटि वाले χ^2 चरों के लिए प्रायिकता वक्रों को दर्शाता है।

काई–वर्ग बंटन सर्वाधिक व्यापक रूप से प्रयोग किए जाने वाले प्रायिकता बंटनों में से एक है। काई–वर्ग प्रायिकता वक्र के नीचे का क्षेत्रफल 1 के बराबर होता है।

मानक प्रसामान्य बंटन से भिन्न, काई–वर्ग का बंटन प्रतिदर्श के आकार के अनुसार अपनी आकृति बदल लेता है।

प्रतिदर्श छोटे होने की स्थिति बंटन दाईं ओर तिरछा होता है, परंतु प्रतिदर्श का आकार बढ़ने पर यह सममित हो जाता है। काई–वर्ग बंटन के सभी मान धनात्मक ही होते हैं।



चित्र 2.2: काई–वर्ग प्रायिकता वक्र

काई–वर्ग बंटन के गुणधर्म

1. काई–वर्ग बंटन का माध्य स्वतंत्रता की कोटि संख्या (k) के बराबर होता है।
2. काई–वर्ग बंटन का विचरण स्वतंत्रता की कोटि संख्या के दुगने के बराबर होता है, यथा –

   $\sigma^2 = 2k$
3. जब स्वतंत्रता की कोटि संख्या 2 से अधिक या उसके बराबर होती है तो Y का अधिकतम मान तब होता है जब समीकरण निम्नवत् हो –

   $\chi^2 = k - 2$
4. जैसे–जैसे स्वतंत्रता की कोटि संख्या बढ़ती है, काई–वर्ग वक्र एक प्रसामान्य बंटन के निकट पहुँचता जाता है।

2.6 *t*–बंटन

आमतौर पर 'स्टूडेंट्स *t*–बंटन' को उसके सरल रूप '*t*–बंटन' के नाम से ही जाना जाता है।

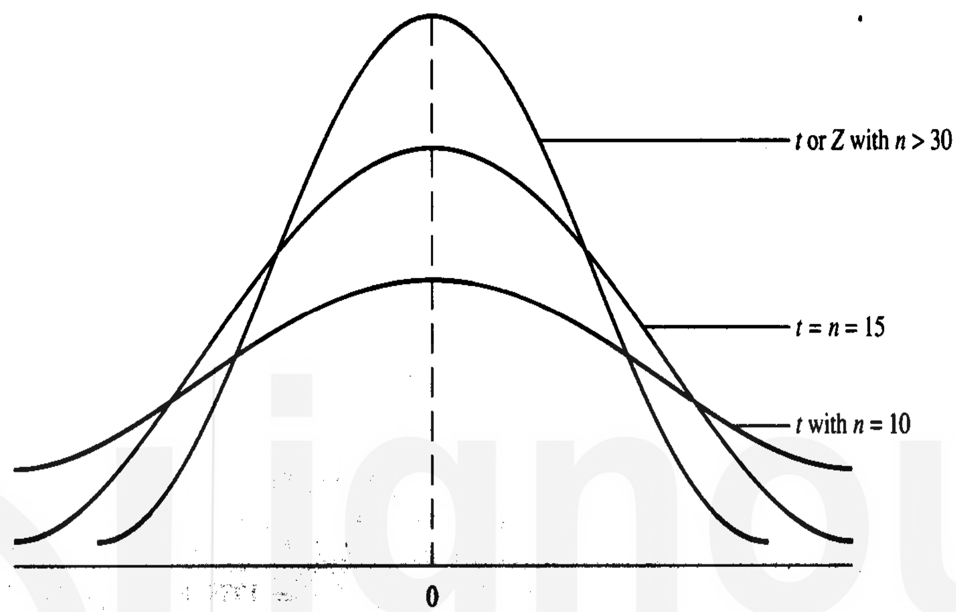
इसे इंग्लैण्ड के सांख्यिकीविद् डब्ल्यू. एस. गोसे द्वारा 'स्टूडेंट' उपनाम के तहत प्रस्तुत किया गया था। यह सतत प्रायिकता बंटनों के परिवार से संबंध रखता है।

यह *t*–बंटन वहाँ प्रयोज्य होता है जहाँ प्रतिदर्श आकार छोटा हो और जनसंख्या मानक विचलन अज्ञात हो।

यह ऐसे उदाहरणों में भी प्रयोज्य होता है जिनमें जनसंख्या प्राचल, यथा μ और σ ज्ञात न हों और प्रतिदर्श आँकड़ों का प्रयोग करके अनुमान लगाया जाता हो।

मानक प्रसामान्य बंटन (z) के उदाहरण में t–बंटन सममित होता है। किसी भी t–बंटन की ऊँचाई प्रतिदर्श के आकार पर निर्भर करती है (देखें चित्र 2.3)।

जैसे–जैसे चर n चर ∞ के निकट आता है, t–बंटन मानक प्रसामान्य बंटन के निकट पहुँचता जाता है।



चित्र 2.3: स्टूडेंट्स t–प्रायिकता वक्र

यदि z_1 एक मानक प्रसामान्य विचर हो, यथा $z_1 \sim N(0,1)$, और z_2 एक अन्य स्वतंत्र चर हो, जो कि स्वतंत्रता की k कोटियों के साथ काई–वर्ग बंटन का अनुसरण करता हो, यथा $z_2 \sim \chi_k^2$ तो कहा जाता है कि चर

$$t = \frac{z_1}{\sqrt{(z_2/k)}} = \frac{z_1\sqrt{k}}{\sqrt{z_2}} \qquad \text{... (2.6)}$$

स्वतंत्रता की k कोटियों के साथ स्टूडेंट्स k–बंटन का अनुसरण करता है।

उक्त k–बंटन का मान निम्नानुसार ज्ञात किया जा सकता है –

$$t = \frac{[\bar{X} - \mu]}{[\frac{s}{\sqrt{n}}]} \qquad \text{... (2.7)}$$

जहाँ, $\bar{X}$ प्रतिदर्श माध्य है, μ जनसंख्या माध्य है, s प्रतिदर्श का मानक विचलन है और n प्रतिदर्श का आकार है।

t–बंटन के गुणधर्म

1. इस बंटन का माध्य 0 के बराबर होता है।
2. इसमें विचरण $[k / (k - 2)]$ के बराबर होता है, जहाँ k स्वतंत्रता की कोटि बतलाता है और $k \geq 2$ होता है।
3. विचरण सदैव 1 से अधिक होता है, हालाँकि स्वतंत्रता की कोटि बड़ी होने पर यह 1 के निकट हो जाता है। स्वतंत्रता की अनंत कोटि के लिए t–बंटन मानक प्रसामान्य बंटन के समान ही होता है।

उक्त t–बंटन का प्रयोग निम्नलिखित परिस्थितियों में किया जा सकता है –

1. जब जनसंख्या बंटन प्रसामान्य हो।
2. जब जनसंख्या बंटन सममित हो, बिना किसी बहिर्वर्ती (मुख्य बिंदु से दूर या अलग) कारक के एकबहुलकी अर्थात् एकरूपात्मक हो, और प्रतिदर्श का आकार कम से कम 30 हो।
3. जब जनसंख्या बंटन मध्यम रूप से तिरछा हो, बिना किसी बहिर्वर्ती कारक के एकबहुलकी हो, और प्रतिदर्श का आकार कम से कम 40 हो।
4. जब प्रतिदर्श का आकार बिना किसी बहिर्वर्ती कारक के 40 से अधिक हो।

उपर्युक्त शर्तों को ध्यान से देखें। यदि मूल जनसंख्या (जिससे प्रतिदर्श लिया गया है) प्रसामान्य हो तो हम किसी भी प्रतिदर्श के आकार के लिए t–बंटन लागू कर सकते हैं।

यदि जनसंख्या प्रसामान्य न हो तो प्रतिदर्श का आकार बड़ा होना चाहिए।

इस t–बंटन का प्रयोग ऐसी जनसंख्या से लिए गए छोटे प्रतिदर्शों के साथ नहीं किया जाना चाहिए जो लगभग प्रसामान्य न हों।

2.7 F–बंटन

एक अन्य सतत प्रायिकता बंटन जिस पर अब हम चर्चा करेंगे, F–बंटन कहलाता है।

यदि z_1 और z_2 ऐसे दो काई–वर्गित चर हों जो स्वतंत्र रूप से क्रमशः स्वतंत्रता की कोटि k_1 और कोटि k_2 के साथ बंटित हों तो चर

$$F = \frac{z_1/k_1}{z_2/k_2} \qquad \text{... (2.8)}$$

क्रमशः स्वतंत्रता की कोटि k_1 और कोटि k_2 के साथ F–बंटन का अनुसरण करता है। यह चर F_{k_1,k_2} से इंगित किया जाता है, जहाँ पादांक k_1 और k_2 काई–वर्गित चरों से सहबद्ध स्वतंत्रता की कोटि दर्शाते हैं।

यहाँ ध्यान देने की बात है कि चर k_1 को स्वतंत्रता की 'अंश कोटि' कहा जाता है, और इसी भाँति, चर k_2 को स्वतंत्रता की 'हर कोटि' कहा जाता है।

उक्त F–बंटन के कुछ महत्वपूर्ण गुणधर्मों का उल्लेख नीचे किया गया है –

1. काई–वर्गित बंटन की भाँति, F–बंटन भी दाईं ओर तिरछा होता है। किंतु जैसे–जैसे k_1 और k_2 में वृद्धि होती है, F–बंटन प्रसामान्य बंटन के निकट पहुँच जाता है।

2. इस F–बंटन का माध्य $k_1 / (k_2 - 2)$ होता है, जो कि $k_2 > 2$ के लिए परिभाषित किया जाता है, और इसका विचरण निम्नवत् होता है –

$$\frac{2k_2^2(k_1 + k_2 - 2)}{k_1(k_2 - 2)^2(k_2 - 4)}$$

   जो कि $k_2 > 4$ के लिए परिभाषित किया जाता है।

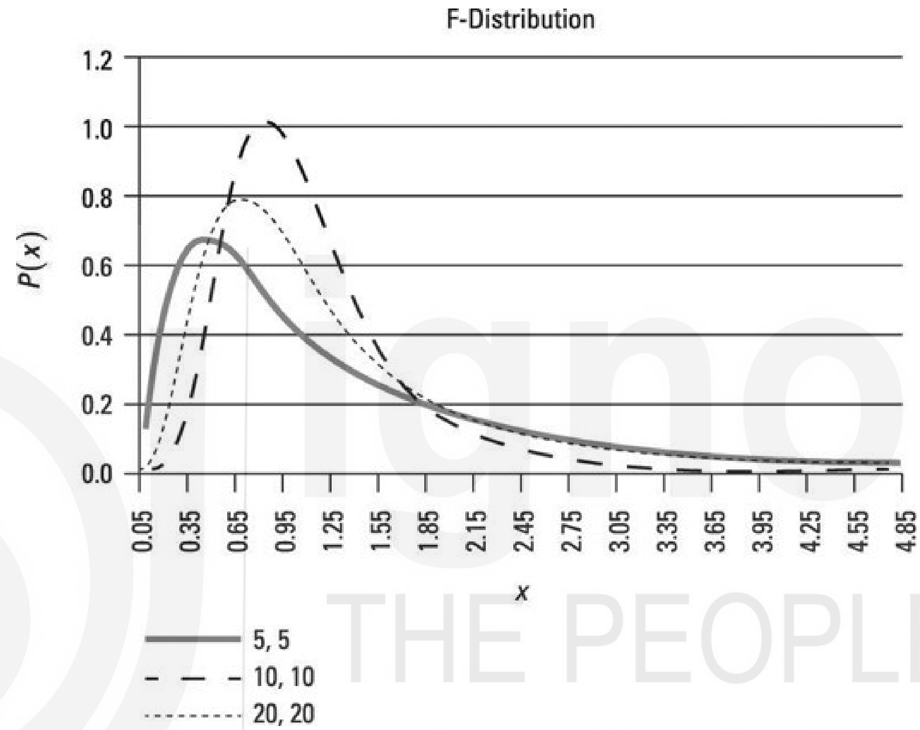
3. स्वतंत्रता की अंश कोटि और हर कोटि के रूप में क्रमशः 1 और k वाला कोई भी F–बंटन स्वतंत्रता की k कोटि वाले किसी स्टूडेंट्स t–बंटन का वर्ग होता है। प्रतीकात्मक रूप से,

$$F_{1,k} = t_k^2$$

4. स्वतंत्रता की काफी बड़ी हर कोटि k_2 के लिए स्वतंत्रता की अंश कोटि k_1 और F मान का गुणनफल स्वतंत्रता की काई–वर्गित मान के लगभग बराबर होता है, यथा –

$$k_1 F = \chi^2_{k_1}$$

उक्त F–बंटन का व्यापक रूप से सांख्यिकीय अनुमिति और परिकल्पनाओं के परीक्षण में प्रयोग किया जाता है। पुनः, इस प्रकार के प्रयोगों में F–प्रायिकता वक्र के अंतर्गत क्षेत्रों को प्राप्त करने और इसके परिणामस्वरूप F–घनत्व फलन को एकीकृत करने की भी अपेक्षा होती है। बहरहाल, इस उदाहरण में भी हमारे कार्य को F–तालिका के प्रावधान द्वारा सुगम बनाया जाता है।



चित्र 2.4: F–बंटन के प्रायिकता वक्र

उक्त F–बंटन का प्रयोग जनसंख्या विचरण का परीक्षण करने के लिए किया जाता है। हम यह जाँच भी कर सकते हैं कि क्या दो प्रसामान्य जनसंख्याओं में एक समान विचरण है। यहाँ शून्य परिकल्पना यह है कि विचरण एक समान होते हैं जबकि वैकल्पिक परिकल्पना यह है कि कोई एक विचरण दूसरे विचरणों से बड़ा होता है, यथा –

$$H_o: \sigma_1^2 = \sigma_2^2$$

$$H_A: \sigma_1^2 > \sigma_2^2$$

उक्त वैकल्पिक परिकल्पना में कहा गया है कि पहली जनसंख्या में अपेक्षाकृत बड़ा विचरण होता है। जबकि शून्य परिकल्पना का परीक्षण प्रत्येक जनसंख्या से एक प्रतिदर्श लेकर तथा आकलनों s_1^2 और s_2^2 की गणना करके किया जा सकता है। ये प्रतिदर्श स्वतंत्र रूप से क्रमशः n_1 और n_2 आकार के साथ लिए गए माने जाते हैं। हम निम्नलिखित अनुपात का परीक्षण करते हैं –

$$F = \frac{s_1^2}{s_2^2} \sim F_{n_{1-1,}, n_2 - 1} \qquad \text{... (2.9)}$$

यदि शून्य परिकल्पना सत्य न हो तो अनुपात इकाई से सांख्यिकीय रूप से भिन्न होगा। हमें F के परिकलित मान (समीकरण (2.9) से प्राप्त) की तुलना F के तालिकाबद्ध मान (पुस्तक के अंत में

परिशिष्ट तालिका में दिए गए) से करनी चाहिए। यदि परिकलित मान तालिकाबद्ध मान से अधिक हो तो शून्य परिकल्पना निरस्त हो जाती है।

बोध प्रश्न 2

1. एक नई विकसित बैटरी एक बार चार्ज करने पर 60 मिनट तक चलती है। मानक विचलन 4 मिनट है। गुणवत्ता नियंत्रण परीक्षण के उद्देश्य से विभाग यादृच्छिक रूप से 7 बैटरियों का चयन करता है। चयनित बैटरियों का मानक विचलन 6 मिनट है। इसकी क्या प्रायिकता है कि नए परीक्षण में मानक विचलन 6 मिनट से अधिक होगा?

   ....

   ....

   ....

   ....

2. काई–वर्ग बंटन को परिभाषित कीजिए। इसके महत्वपूर्ण गुणधर्मों पर प्रकाश डालें।

   ....

   ....

   ....

   ....

   ....

3. मान लीजिए कि स्नातक रिकॉर्ड परीक्षा (GRE) में अंक सामान्य रूप से 100 के जनसंख्या माध्य के साथ बंटित किए जाते हैं। माना कि 20 लोगों को यादृच्छिक रूप से चुनकर उनका परीक्षण किया जाता है। नमूना मानक विचलन 15 है। क्या संभावना है कि औसत परीक्षण अंक अधिकतम 110 होंगे?

   ....

   ....

   ....

   ....

   ....

4. परीक्षण करें कि क्या निजी हाई स्कूलों के छात्र अपने विज्ञान परीक्षण अंकों के संबंध में सरकारी हाई स्कूलों के छात्रों की तुलना में अधिक सजातीय होते हैं। यह ज्ञात है कि सरकारी और निजी स्कूलों के लिए प्रतिदर्श विचलन क्रमशः 91.74 और 67.16 है। छात्रों का प्रतिदर्श आकार सरकारी स्कूलों के लिए 506 और निजी स्कूलों के लिए 94 है।

   ....

   ....

   ....

   ....

   ....

2.8 प्राचलों का आकलन

आकलन दो प्रकार का हो सकता है, यथा –

1. बिंदु आकलन (point estimation), और
2. अंतराल आकलन (interval estimation)

बिंदु आकलन में हम एक बिंदु के रूप में जनसंख्या प्राचल के मान का अनुमान लगाते हैं। दूसरी ओर, अंतराल आकलन के मामले में हम प्रतिदर्श माध्य के आसपास की निचली और ऊपरी उन सीमाओं का अनुमान लगाते हैं जिनके भीतर जनसंख्या माध्य रहने की संभावना होती है।

2.8.1 बिंदु आकलन

चलिए, मान लेते हैं कि एक यादृच्छिक चर X सामान्य बंटन का अनुसरण करता है। जैसा कि आप जानते हैं, सामान्य बंटन को दो मापदंडों से दर्शाया जाता है, यथा माध्य और मानक विचलन। चूँकि हमारे पास समग्र जनसंख्या के लिए आँकड़े नहीं हैं (हमारे पास केवल एक प्रतिदर्श के लिए आँकड़े हैं), हमें केवल एक प्रतिदर्श के आधार पर ही माध्य $E(X) = \mu_X$ और विचलन σ_X^2 ज्ञात करने होंगे।

हम यह भी मान लेते हैं कि हमारे पास एक ज्ञात प्रायिकता बंटन (मान लीजिए, प्रसामान्य बंटन) से आकार n (मान लीजिए, प्रतिदर्श आकार $n = 50$) के यादृच्छिक प्रतिदर्श से आँकड़े हैं।

हम इस प्रतिदर्शज (statistic) का प्रयोग अज्ञात प्राचल (parameter) का आकलन करने के लिए करेंगे।

मान लीजिए कि हम प्रतिदर्श माध्य $\bar{X}$ 23.28 पाते हैं। इस एकल संख्यात्मक मान को उस प्राचल का 'बिंदु आकलन' कहा जाता है जहाँ

$$\bar{X} = \frac{\Sigma X_i}{n}$$

इस सूत्र को 'बिंदु आकलक' कहा जाता है। यहाँ ध्यान देने की बात है कि बिंदु आकलक एक यादृच्छिक चर होता है क्योंकि इसका मान प्रतिदर्श से प्रतिदर्श में भिन्न–भिन्न होता है।

2.8.2 अंतराल आकलन

बिंदु आकलन में हम प्राचल का आकलन एकल मान से करते हैं, जो कि प्रायः संबंधित प्रतिदर्श आँकड़े से प्राप्त होता है। बिंदु अनुमान इस अर्थ में यथार्थपरक नहीं हो सकता है कि प्राचल मान पूर्णतः उसके बराबर नहीं हो सकता है।

किसी वैकल्पिक प्रक्रिया का अर्थ कोई ऐसा अंतराल दे देना होता है जो निश्चित प्रायिकता के साथ प्राचल को धारण कर ले। यहाँ हम एक निचली सीमा और एक ऊपरी सीमा निर्दिष्ट कर देते हैं; प्राचल मान इसी के भीतर रहने की संभावना होती है।

इसके अलावा, हम अंतराल में शेष प्राचल की संभावना निर्दिष्ट कर देते हैं। हम अंतराल को 'आत्मविश्वास अंतराल' और इस अंतराल के भीतर शेष प्राचल की प्रायिकता को 'आत्मविश्वास स्तर' या 'आत्मविश्वास गुणांक' कहते हैं।

आत्मविश्वास अंतराल की अवधारणा कुछ जटिल है। हम इसको पाठ्यक्रम BECC 107, इकाई 13 में पहले ही स्पष्ट कर चुके हैं। आइए, इसका पुनरवलोकन करें।

हमने एक प्रसामान्य जनसंख्या से आकार n का एक प्रतिदर्श लिया है। हमें जनसंख्या माध्य μ_X और जनसंख्या विचरण σ_X^2 ज्ञात नहीं हैं। हमें केवल प्रतिदर्श माध्य $\bar{X}$ और प्रतिदर्श विचरण S_X^2 ज्ञात

हैं। चूँकि $\bar{X}$ प्रतिदर्शों में भिन्न–भिन्न होता है, चर μ_X के विषय में निष्कर्ष निकालने के लिए हम $\bar{X}$ के प्रतिदर्शन बंटन संबंधी गुणधर्मों का प्रयोग करते हैं।

यदि X को सामान्य रूप से बंटित किया जाता है, यथा हम जानते हैं कि

$$\bar{X} \sim \left(\mu_X, \frac{\sigma_x^2}{n} \right) \qquad \text{... (2.10)}$$

तो समीकरण (2.10) से हम कह सकते हैं कि प्रतिदर्शन माध्य $\bar{X}$ का प्रतिदर्श बंटन माध्य μ_X और मानक विचलन σ_X^2/n के साथ प्रसामान्य बंटन का अनुसरण करता है।

चलिए, उपर्युक्त समीकरण को एक मानक प्रसामान्य चर के रूप में परिवर्तित करते हैं, यथा –

$$Z = \frac{\bar{X} - \mu_x}{\frac{\sigma_X}{\sqrt{n}}} \sim N(0, 1) \qquad \text{... (2.11)}$$

अब हमारे समक्ष समस्या यह है कि हमें जनसंख्या विचरण σ_X^2 ज्ञात नहीं है। अतएव हम इसका अनुमानक लेते हैं, यथा –

$$S_X^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1}$$

ऐसी स्थिति में उपयुक्त परीक्षण आँकड़ा निम्नवत् होगा –

$$t = \frac{(\bar{X} - \mu_X)}{S_x / \sqrt{n}} \qquad \text{... (2.11)}$$

समीकरण (2.11) स्वतंत्रता की कोटि (n –1) के साथ t–बंटन का अनुसरण करता है।

समीकरण (2.11) में पदों को पुनर्व्यवस्थित करने से हमें μ_X का आत्मविश्वास अंतराल प्राप्त होता है। यह हमें μ_X का अंतराल आकलन प्राप्त करने में भी मदद करता है।

स्वतंत्रता की कोटि (d.f.) 27 के लिए महत्व के 5 प्रतिशत स्तर पर तालिकाबद्ध मान 2.052 होगा (देखें परिशिष्ट तालिका)। तदनुसार,

$$P(-2.052 \leq t \leq 2.052) = 0.95 \qquad \text{... (2.12)}$$

उक्त महत्वपूर्ण t मान उस t–बंटन वक्र के अंतर्गत क्षेत्र की प्रतिशतता दर्शाते हैं जो उन मानों के बीच रहता है। मान t = –2.052 को निम्न महत्वपूर्ण मान कहा जाता है, और मान t = 2.052 को उच्च महत्वपूर्ण मान कहा जाता है।

समीकरण (2.12) का अर्थ है कि स्वतंत्रता की कोटि 27 के लिए प्रायिकता 0.95 अथवा 95% है, यथा अंतराल (–2.052, 2.052) में μ_X शामिल है।

$$\therefore \qquad P\left(-2.052 \leq t = \frac{\bar{X} - \mu_X}{S_x / \sqrt{n}} \leq 2.852 \right)$$

$$\Rightarrow \qquad P\left(\bar{X} - 2.052 \frac{S_X}{\sqrt{n}} \leq \mu_X \leq \bar{X} + 2.052 \frac{S_X}{\sqrt{n}} \right) = 0.95 \qquad \text{... (2.13)}$$

समीकरण (2.13) का अंतराल हमें μ_X का अनुमानक प्रदान करता है। इसे ही वास्तविक परंतु अज्ञात जनसंख्या माध्य μ_X के लिए 95% आत्मविश्वास अंतराल (CI) कहा जाता है। मान 0.95 को आत्मविश्वास गुणांक कहा जाता है। इसका अर्थ है कि प्रायिकता 0.95 है, यथा यादृच्छिक अंतराल

$$\bar{X} \pm 2.052 \frac{S_X}{\sqrt{n}}$$

में वास्तविक μ_X विद्यमान है।

$\bar{X} - 2.052 \frac{S_X}{\sqrt{n}}$ अंतराल की निम्न सीमा कहलाती है।

$\bar{X} + 2.052 \frac{S_X}{\sqrt{n}}$ अंतराल की उच्च सीमा कहलाती है।

यह एक यादृच्छिक अंतराल कहलाता है क्योंकि ज्ञात मान $\bar{X}$ और $\frac{S_X}{\sqrt{n}}$ पर आधारित हैं जो कि प्रतिदर्श से प्रतिदर्श में भिन्न–भिन्न होंगे।

यह भी ध्यान देने की बात है कि μ_X कोई यादृच्छिक चर नहीं है; बल्कि यह एक निश्चित संख्या है। अतः हम कह सकते हैं कि "प्रायिकता 0.95 है, यथा इस अंतराल में μ_X निहित है।"

2.9 आकलकों के गुणधर्म

किसी आकलक को सर्वोत्तम रैखिक निष्पक्ष आकलक (BLUE) माना जाता है यदि वह रेखाकार, अपक्षपाती, कार्यक्षम (न्यूनतम विचरण के साथ) हो, और वह सुसंगत भी हो।

इसका अर्थ है कि प्रतिदर्श आकार बढ़ने पर आकलक का मान अपने वास्तविक जनसंख्या मान में समाभिरूप हो जाता हो।

एक अच्छे आकलक के सभी गुणधर्मों पर विस्तृत चर्चा नीचे की गई है।

2.9.1 रैखिकता

किसी आकलक को एक रैखिक आकलक कहा जाता है यदि वह निम्नलिखित प्रतिदर्श समुक्ति का एक रैखिक फलन हो –

$$\bar{X} = \sum_{i=1}^{n} \frac{Xi}{n}$$

$$= \frac{1}{n}\left(X_1 + X_2 + \ldots.. + X_n\right) \quad \ldots (2.14)$$

प्रतिदर्श माध्य रैखिक आकलक होता है क्योंकि वह समुक्तियों का एक रैखिक फलन होता है।

2.9.2 निष्पक्षता

प्रतिदर्शन में उतार–चढ़ाव के कारण किसी भी आँकड़े का मान प्रतिदर्शों में भिन्न–भिन्न होता है। यद्यपि किसी आँकड़े के विशिष्ट मान अज्ञात जनसंख्या प्राचल से भिन्न हो सकते हैं, औसतन, किसी भी आँकड़े का मान जनसंख्या प्राचल के बराबर ही होना चाहिए।

दूसरे शब्दों में, चर $\bar{X}$ का प्रतिदर्शन बंटन मुख्य रूप से चर μ_X की ओर ही प्रवृत्त होना चाहिए। इसे ही किसी आकलक के निष्पक्षता नामक गुणधर्म के रूप में जाना जाता है।

इसका अर्थ यह है कि यद्यपि किसी ज्ञात आकलक का कोई विशिष्ट मान जनसंख्या प्राचल के अज्ञात मान से अधिक या कम हो सकता है, आकलक की ओर से कोई पूर्वाग्रह नहीं होगा कि वे ऐसे मान लें जो अज्ञात जनसंख्या प्राचल से सदैव अधिक अथवा न्यून हों।

यदि हम यह स्वीकार कर लें कि माध्य (यहाँ, प्रत्याशा) सबसे महत्वपूर्ण प्रवृत्ति के लिए एक उचित मापदंड होता है तो $\bar{X}$ चर μ_X के लिए एक *निष्पक्ष आकलक* होगा, बशर्ते

$$E(\bar{X}) = \mu_X$$

2.9.3 न्यूनतम विचरण

चर μ_X के किसी भी आकलक को न्यूनतम विचरण आकलक कहा जाता है यदि उसका विचरण चर μ_X के किसी भी अन्य आकलक के विचरण से छोटा हो।

मान लीजिए कि चर μ_X के तीन आकलक हैं। इन तीन आकलकों में चर $\hat{\mu}_3$ का विचरण सबसे छोटा है। इस कारण यह न्यूनतम विचरण आकलक कहलाएगा।

2.9.4 कार्यक्षमता

निष्पक्षता का गुणधर्म अपने आप में पर्याप्त नहीं होता। किसी भी प्राचल के दो या दो से अधिक आकलक निष्पक्ष रूप से प्राप्त करना संभव होता है। अतः हमें सर्वाधिक कार्यक्षम आकलक ही चुनना चाहिए।

मान लीजिए कि चर μ_X के दो आकलक निम्नानुसार दिए गए हैं –

$$\bar{X} \sim N\left(\mu_X, \frac{\sigma^2}{n}\right) \qquad \text{... (2.15)}$$

$$X_{med} \sim N\left(\mu_X, \left(\frac{\pi}{2}\right)\frac{\sigma^2}{n}\right), \qquad \pi = 3.142 \quad \text{(लगभग)} \qquad \text{... (2.16)}$$

बड़े प्रतिदर्शों के मामले में प्रसामान्य जनसंख्या के यादृच्छिक प्रतिदर्श से गणना की गई माध्यिका भी उसी μ_X के साथ प्रसामान्य बंटन का अनुसरण करती है। हालाँकि, इसमें एक बड़ा विचरण देखा जाता है, यथा –

$$\frac{Var(\bar{X}_{med})}{Var(\bar{X})} = \frac{\pi}{2}\frac{\frac{\sigma^2}{n}}{\frac{\sigma^2}{n}} - \frac{\pi}{2} = 1.571 \quad \text{(लगभग)} \qquad \text{... (2.17)}$$

समीकरण (2.17) का अर्थ है कि प्रतिदर्श माध्य का विचरण प्रतिदर्श माध्य के विचरण से 57% बड़ा है। अतः प्रतिदर्श माध्य माध्यिका (X_{med}) की तुलना में जनसंख्या माध्य का अपेक्षाकृत अधिक सटीक आकलन प्रस्तुत करता है। तदनुसार, चर $\bar{X}$ चर μ_X का एक *कार्यक्षम आकलक* है।

2.9.5 सर्वोत्तम रैखिक निष्पक्ष आकलक (BLUE)

मान लीजिए कि हम आकलकों के किसी वर्ग पर विचार करते हैं। इन आकलकों के बीच एक आकलक किन्हीं तीन गुणधर्मों को संतुष्ट करता है, यथा –

1. वह रैखिक है,
2. वह निष्पक्ष है, और
3. उसका विचरण न्यूनतम है।

ऐसी स्थिति में उसे 'सर्वोत्तम रैखिक निष्पक्ष आकलक' (Best Linear Unbiased Estimator - BLUE) कहा जाएगा।

2.9.6 सुसंगतता

सुसंगतता एक व्यापक प्रतिदर्श गुणधर्म है। यदि हम प्रतिदर्श आकार बढ़ाते हैं तो आकलक में प्राचल के मान तक पहुँचने की प्रवृत्ति होनी चाहिए। तदनुसार, किसी भी आकलक को सुसंगत कहा जाता है यदि वह $n \rightarrow \infty$ की भाँति प्राचल के रूप में अभिसरित हो जाता है।

मान लीजिए कि $X \sim N(\mu_X, \sigma_X^2)$ ज्ञात है।

अब हम जनसंख्या से आकार n का एक यादृच्छिक प्रतिदर्श लेंगे।

चर μ_X के दो आकलक निम्नवत् हैं –

$$\bar{X} = \Sigma \frac{X_i}{n} \quad \text{... (2.18)}$$

$$X^* = \Sigma \frac{X_i}{n+1} \quad \text{... (2.19)}$$

जैसा कि हमें ज्ञात है, प्रथम आकलक (2.18) ही प्रतिदर्श माध्य है और वह निष्पक्ष है क्योंकि $E(\bar{X}) = \mu_X$ होता है।

यहाँ दूसरा आकलक (2.19) पक्षपाती है क्योंकि

$$E(X^*) = \left(\frac{n}{n+1}\right)\mu_X$$

तदनुसार,

$$E(X^*) \neq \mu_X$$

जैसे–जैसे प्रतिदर्श आकार बढ़ता है, हमें दो आकलकों के बीच अधिक अंतर नहीं खोजना चाहिए। जैसे–जैसे n बढ़ेगा, चर X^* चर μ_X के निकट आता जाएगा। इस प्रकार के आकलक को *सुसंगत आकलक* के रूप में जाना जाता है।

कोई भी आकलक एक सुसंगत आकलक कहलाता है यदि वह प्राचल के सही मान तक पहुँचता है क्योंकि प्रतिदर्श का आकार बड़ा और बड़ा होता जाता है।

बोध प्रश्न 3

1. एक आदर्श आकलक के वांछनीय गुणधर्मों का वर्णन कीजिए।

....

....

....

....

....

2. आकार 30 के प्रतिदर्श के लिए, प्रतिदर्श माध्य और मानक विचलन क्रमशः 15 और 10 हैं। महत्व के 5 प्रतिशत स्तर पर जनसंख्या माध्य (μ_X) के आत्मविश्वास अंतराल की रचना करें।

....

....

....

....

....

2.10 सार–संक्षेप

सांख्यिकीय अवधारणाएँ अनिश्चितता की विद्यमानता में निर्णय लेने के लिए हमारा मार्गदर्शन करती हैं। इस इकाई में हमने कुछ आधारभूत सांख्यिकीय अवधारणाओं के बारे में चर्चा की।

हमने प्रसामान्य, मानक प्रसामान्य, काई–वर्ग, t और F जैसे विशिष्ट प्रकार के सतत प्रायिकता बंटन के बारे में भी चर्चा की। हमने इन वक्रों के प्रायिकता बंटन वक्रों को दर्शाया।

इस पुस्तक के अंत में दिए गए परिशिष्ट में हमने विभिन्न प्रकार की तालिकाओं का लाभ उठाया, यथा – प्रसामान्य क्षेत्र तालिका, और t, काई–वर्ग एवं F बंटन के महत्वपूर्ण मान।

उपर्युक्त के अतिरिक्त हमने एक अच्छे आकलक के गुणों का वर्णन भी किया है।

हमने एक आकलक के संदर्भ में निष्पक्षता, कार्यक्षमता और सुसंगतता जैसी अवधारणाओं की व्याख्या भी की है।

2.11 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1. आपको मानक प्रसामान्य वक्र के अंतर्गत क्षेत्रफल ज्ञात करना होगा।

   यदि X = 40, $z = \frac{(40-30)}{4} = 2.5$ P (X < 40) = P (z < 2.5) = [2.5 के बाएँ स्थित क्षेत्र]

   = 0.9938

2. पाठांश 2.4 पढ़ें और उत्तर दें।

3. a) P (X < 7) = P (Z < –2.5) = 0.0062

   b) P (20 < X < 22) = P (–2.5 < Z < 0) = 0.4938

बोध प्रश्न 2

1. जनसंख्या का मानक विचलन 4 मिनट है।

   प्रतिदर्श का मानक विचलन 6 मिनट है।
   प्रतिदर्श समुक्तियों की संख्या 7 है।

   $$X^2 = [(n-1)*s^2] / \sigma^2$$
   $$X^2 = [(7-1)*6^2] / 4^2 = 13.5$$

   स्वतंत्रता की कोटि $(n-1) = (7-1) = 6$ है।
   एक मानक विचलन के 6 मिनट से कम अथवा उसके बराबर होने की प्रायिकता 0.96 है।

इसका तात्पर्य यह है कि मानक विचलन के 6 मिनट से अधिक होने की प्रायिकता (1 – 0.96) = 0.04 है।

2. पाठांश 2.5 पढ़ें और उत्तर दें।

3. जनसंख्या माध्य $\mu = 100$ है। प्रतिदर्श आकार n = 20 है।

   स्वतंत्रता की कोटि (20 – 1) = 19 है। प्रतिदर्श माध्य $\bar{X}$ अधिकतम 110 होना चाहिए। प्रतिदर्श मानक विचलन s = 15 है।
   चूँकि हमें जनसंख्या मानक विचलन ज्ञात नहीं है, हम t–बंटन प्रयोग करते हैं। ज्ञात सूत्र को प्रयोग करने पर हमें प्राप्त होता है–

$$t = \frac{[\bar{X} - \mu]}{[\frac{s}{\sqrt{n}}]}$$

   तदनुसार,

$$t = \frac{110 - 10}{\frac{15}{\sqrt{20}}} = 0.996$$

   इसका अर्थ है कि 99.6% संभावना है कि प्रतिदर्श औसत 110 से अधिक नहीं होगा।
   स्वतंत्रता की कोटि (n_1–1) और (n_2–1) क्रमशः 505 और 93 हैं। हमारी शून्य परिकल्पना H_0 यह है कि विज्ञान के अंकों के संबंध में दोनों प्रकार के स्कूल समान रूप से सजातीय हैं। हम विचरणों की तुलना कर रहे हैं। अतः हम F–परीक्षण प्रयोग करेंगे, यथा –

$$F = {s_1^2}/{s_2^2} = \frac{91.74}{67.16} = 1.366$$

4. स्वतंत्रता की कोटि 505 और 93 के लिए F का तालिकाबद्ध मान 1.27 है। चूँकि परिकलित मान तालिकाबद्ध मान से अधिक है, हम H_0 को निरस्त करते हैं।हम यह निष्कर्ष निकालते हैं कि निजी स्कूलों के छात्र विज्ञान के अंकों के संबंध में अधिक सजातीय हैं।

बोध प्रश्न 3

1. पाठांश 2.9 पढ़ें और उत्तर दें।

2. चूँकि जनसंख्या मानक विचलन ज्ञात नहीं है, आप t–बंटन प्रयोग करें।

   स्वतंत्रता की कोटि 29 और 5 प्रतिशत महत्व के स्तर के लिए परिशिष्ट में दिए गए चर t के तालिकाबद्ध मान की जाँच करें।
   समीकरण (2.12) में दिए गए विश्वास्यता अंतराल (confidence interval) की रचना कीजिए।

इकाई 3 परिकल्पना परीक्षण : विहंगावलोकन*

इकाई की रूपरेखा

3.0 उद्देश्य

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

- परिकल्पना परीक्षण की संकल्पना और सार्थकता पर प्रकाश डाल सकें;
- किसी परीक्षण प्रतिदर्शज के अनुप्रयोगों का वर्णन कर सकें;
- समष्टि प्राचल संबंधी परिकल्पना के परीक्षण की प्रक्रिया स्पष्ट कर सकें;
- टाइप–I और टाइप–II त्रुटियों के बीच अंतर स्पष्ट कर सकें; तथा
- दो भिन्न–भिन्न प्रतिदर्शों से प्राचलों की तुलना करने के लिए परीक्षण व्यवहार में ला सकें।

3.1 विषय–प्रवेश

सांख्यिकीय अनुमिति का मूल उद्देश्य समष्टि प्राचलों के विषय में निर्णय लेने के लिए किसी प्रतिदर्श का प्रयोग करना ही होता है। परिकल्पना परीक्षण की संकल्पना किसी प्रतिदर्श का प्रयोग कर समष्टि प्राचलों के मान का पूर्वानुमान किए जाने के लिए अत्यंत महत्वपूर्ण मानी जाती है। समष्टि माध्य और विचरण से संबंधित परिकल्पनाओं का परीक्षण करने के लिए विभिन्न परीक्षण प्रतिदर्शजों का प्रयोग किया जाता है। परिकल्पना परीक्षण का प्रयोग कर दो भिन्न–भिन्न प्रतिदर्शों के विचरण की तुलना भी की जा सकती है।

परिकल्पना के परीक्षण के लिए दो उपागम प्रचलन में हैं –

(i) सार्थकता का परीक्षण उपागम, और

(ii) विश्वस्तता अंतराल उपागम।

* डॉ. पूजा शर्मा, असिस्टेंट प्रोफेसर, दौलत राम कॉलेज, दिल्ली विश्वविद्यालय।

किसी भी परिकल्पना का परीक्षण करते समय दो प्रकार की त्रुटियाँ होने की संभावना होती है – (i) टाइप–I त्रुटि, और (ii) टाइप–II त्रुटि।

इस इकाई में हम परिकल्पना परीक्षण की प्रक्रिया के बारे में विस्तार से जानेंगे, और समुचित परीक्षण प्रतिदर्शज के आधार पर शून्य–स्तरीय परिकल्पना को अस्वीकार करने की विधि की व्याख्या करेंगे।

3.2 परिकल्पना परीक्षण की प्रक्रिया

हम आर्थिक सिद्धांत अथवा तर्क के आधार पर ही किसी परिकल्पना का निर्माण करते हैं। कोई भी परिकल्पना किसी समष्टि के कुछ विशिष्ट अभिलक्षणों के विषय में कोई अस्थायी समुक्ति होता है। जैसा कि हम जानते हैं, किसी भी समष्टि का वर्णन उसके प्राचलों (जैसे माध्य, मानक विचलन, आदि) द्वारा किया जाता है। तदनुसार, कोई भी परिकल्पना किसी समष्टि प्राचल के विषय में कोई अवधारणा होती है। कोई परिकल्पना सत्य हो भी सकती है और नहीं भी। इसी बात का पता लगाने के लिए कि हम एक विशिष्ट अर्थमितीय विधि से किसी परिकल्पना का परीक्षण करते हैं।

किसी भी परिकल्पना के निर्माण में इस विषय में कोई पूर्व निर्णय अथवा अपेक्षा शामिल होती है कि कोई विशिष्ट प्राचल क्या मान ले सकता है। उदाहरण के लिए, पूर्व ज्ञान अथवा किसी विशेषज्ञ की राय ही हमें बताती है कि स्थानीय शेयर बाजार में वास्तविक औसत कीमत और कमाई (P/E) का अनुपात 20 है। तदनुसार हमारी परिकल्पना यह होगी है कि P/E अनुपात 20 के बराबर होता है।

इस परिकल्पना का परीक्षण करने के लिए मान लीजिए कि हम शेयरों का एक यादृच्छिक प्रतिदर्श एकत्र करते हैं और पाते हैं कि औसत P/E अनुपात 23 है। क्या यह 23 का आँकड़ा सांख्यिकीय रूप से 20 से भिन्न है?

प्रतिदर्शन विचरण के कारण प्रतिदर्श आकलन और उसके समष्टिक मान के बीच अंतर होने की संभावना है। यह संभव है कि सांख्यिकीय रूप से संख्या 23 संख्या 20 से बहुत भिन्न न हो। यदि ऐसा ही है तो हमें इस परिकल्पना को अस्वीकार नहीं करना चाहिए कि औसत P/E अनुपात 20 होता है।

परिकल्पना परीक्षण में चार महत्वपूर्ण घटक होते हैं –

i) शून्य–स्तरीय परिकल्पना,

ii) वैकल्पिक परिकल्पना,

iii) परीक्षण प्रतिदर्शज, और

iv) परिणामों की व्याख्या।

आइए, अब हम इन घटकों के विषय में विस्तार से जानते हैं।

(i) **शून्य–स्तरीय और वैकल्पिक परिकल्पनाओं का निर्माण** : परिकल्पनाएँ दो प्रकार की होती हैं, यथा शून्य–स्तरीय परिकल्पना और वैकल्पिक परिकल्पना। इनमें 'शून्य–स्तरीय परिकल्पना' वह समुक्ति होती है जिसे हम समष्टि के विषय में सत्य मानते हैं। इसे 'शून्य–स्तरीय' कहा जाता है, जिसका अर्थ होता है – निरर्थक या रिक्त। उदाहरण के लिए, एक शून्य–स्तरीय परिकल्पना यह हो सकती है कि रोजगार और शिक्षा के बीच कोई संबंध नहीं होता। अतएव, यदि हम शिक्षा पर रोजगार का समाश्रयण करते हैं तो समाश्रयण गुणांक शून्य होना चाहिए। आमतौर पर हम शून्य–स्तरीय परिकल्पना को H_0 से निरूपित करते हैं। वैकल्पिक परिकल्पना शून्य–स्तरीय परिकल्पना के विपरीत होती है।

वैकल्पिक परिकल्पना को आमतौर पर H_1 से निरूपित किया जाता है। यह ध्यान देने की बात है कि H_0 और H_1 'परस्पर अनन्य' होते हैं; वे संपाती रूप से विद्यमान नहीं हो सकते।

(ii) **परीक्षण प्रतिदर्शज की पहचान** : शून्य–स्तरीय परिकल्पना का परीक्षण किसी परीक्षण प्रतिदर्शज से किया जाता है। अर्थमिति में अनेक परीक्षण प्रतिदर्शज (यथा t, F, काई–वर्ग, आदि) उपलब्ध हैं। हमें अपने लिए उपयुक्त परीक्षण प्रतिदर्शज की पहचान करनी होती है।

(iii) **परीक्षण प्रतिदर्शज के मान के आधार पर परिणामों की व्याख्या** : अपना परीक्षण पूरा करने के बाद हम परिणामों की व्याख्या करते हैं। जब हम परीक्षण प्रतिदर्शज को हमारे पास उपलब्ध प्रतिदर्श आँकड़ों पर प्रयोग करते हैं तो हमें परीक्षण प्रतिदर्शज का एक निश्चित मान ज्ञात होता है (उदाहरण के लिए, 2.535 का t–अनुपात)। परिणामों की व्याख्या में परीक्षण प्रतिदर्शजों के इन दो मानों की तुलना शामिल होती है – तालिकाबद्ध मान और संगणित मान। यदि हमारी तुलना में संगणित मान तालिकाबद्ध मान से अधिक हो तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर देते हैं।

शून्य–स्तरीय परिकल्पना के तहत किसी भी परीक्षण प्रतिदर्शज का प्रतिदर्शन बंटन 'शून्य–स्तरीय बंटन' कहलाता है। जब आँकड़े शून्य–स्तरीय परिकल्पना के विरुद्ध प्रबल साक्ष्य दर्शाते हों तो परीक्षण प्रतिदर्शज का मान बहुत अधिक हो जाता है। परीक्षण प्रतिदर्शज के संगणित मान को देखकर ही हम अपने निष्कर्ष निकालते हैं।

परीक्षण प्रतिदर्शज के अलावा अर्थमितीय सॉफ्टवेयर एक p–मान भी प्रदान करता है। यह प्रायिकता p–मान ही दर्शाता है कि शून्य–स्तरीय परिकल्पना सत्य है। इस प्रकार, यदि हम 0.04 का p–मान ज्ञात करते हैं तो इसके अनुसार शून्य–स्तरीय परिकल्पना के सत्य होने की संभावना 0.04 या 4 प्रतिशत होगी। अतएव, यदि हम सार्थकता का 5 प्रतिशत स्तर लेते हैं तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर देते हैं।

3.3 आकलन विधियाँ

इस पाठ्यक्रम की इकाई 2 में हमने दो अवधारणाओं का वर्णन किया है, यथा बिंदु आकलन और अंतराल आकलन। वहाँ हमने कुछ प्रायिकता बंटन फलनों जैसे कि सामान्य, t, F और काई–वर्ग के विषय में भी चर्चा की है।

आकलन की मूल रूप से तीन विधियाँ प्रचलन में हैं –

(i) न्यूनतम वर्ग,

(ii) अधिकतम संभावना, और

(iii) आघूर्ण विधि।

इस पाठ्यक्रम में हम न्यूनतम–वर्ग आकलन विधि का ही व्यापक रूप से प्रयोग करेंगे। आगे इकाई 7 में हम अधिकतम–संभावना विधि से परिचित होंगे। इस पाठ्यक्रम में, बहरहाल, हमारा परिचय 'आघूर्ण विधि' से नहीं कराया जाएगा। इससे इतर, पाठ्यक्रम 'BECC 107: इकाई 5' में हमने समाश्रयण की अवधारणा पर चर्चा की है। उस इकाई के पाठांश 5.9 में हमने उल्लेख किया है कि समाश्रयण में त्रुटि चर को न्यूनतम किया जाना चाहिए। इस उद्देश्य से ही हमने त्रुटि पदों के वर्गों के योगफल ($\sum u_i^2$) को न्यूनतम किया।

अब आप अनुमान लगा सकते हैं कि इसे न्यूनतम वर्ग विधि क्यों कहा जाता है। इस पाठ्यक्रम में हम साधारण न्यूनतम वर्ग (OLS) विधि तक ही सीमित रहेंगे। इस विधि (OLS) में सबसे पहले हम दो–चर उदाहरण लेंगे। तदंतर हम इसे दो से अधिक चरों तक बढ़ाएँगे। यह हमें बहु समाश्रयण मॉडल की ओर ले जाएगा।

'साधारण न्यूनतम वर्ग' (OLS) नाम से ही ज्ञात होता है कि यह न्यूनतम वर्ग विधियों में सबसे सरल है। इसका तात्पर्य यह है कि आगे की जटिलताओं को भी इस (OLS) विधि के अंतर्गत लाया जा

सकता है। यथातथ्य रूप से ऐसा ही है; सामान्यीकृत न्यूनतम वर्ग (GLS), दो–चरण न्यूनतम वर्ग (2SLS), तीन–चरण न्यूनतम वर्ग (3SLS), आदि विधियाँ उपलब्ध हैं। इसीलिए, जब कभी हम न्यूनतम वर्ग विधि के विषय में पढ़ें तो सावधान रहें – इस बात पर ध्यान दें कि पाठ किस विधि का जिक्र कर रहा है।

जब कभी आप किसी प्रसंग में सामान्यीकृत न्यूनतम वर्ग (GLS) पदबंध से परिचित हों तो इसे साधारण न्यूनतम वर्ग (OLS) समझकर भ्रमित न हों – ये दोनों विधियाँ परस्पर भिन्न हैं।

उक्त दोनों (OLS और GLS) विधियों में त्रुटि पदों के वर्गों का योगफल न्यूनतम होता है (इसीलिए दोनों का न्यूनतम वर्ग विधि के रूप में संदर्भ दिया जाता है) परंतु परवर्ती (GLS) के उदाहरण में

समाश्रयण मॉडल का कुछ रूपांतरण होता है।

इस पाठ्यक्रम में न्यूनतम वर्गों की उन्नत विधियों पर विचार नहीं किया गया है।

याद रखें कि न्यूनतम वर्ग विधि का क्रियान्वयन करने के लिए आपको चरों के विषय में कोई समाश्रयण बंटन फलन मानकर चलने की आवश्यकता नहीं पड़ती है।

अधिकतम संभावना (ML) विधि चरों के विषय में कोई संभाव्यता बंटन मानकर चलती है। प्रसामान्य बंटन अधिकतम संभावना आकलन में सर्वाधिक प्रयुक्त प्रायिकता बंटन फलन है। अधिकतम संभावना (ML) विधि में हम एक संभावना फलन बनाते हैं, जो कि प्रायिकता बंटन फलन से अवकलित होता है।

यह ध्यान देने की बात है कि अर्थमिति में हमें आँकड़े दिए जाते हैं – ये आँकड़े किसी प्रतिदर्श सर्वेक्षण से प्राप्त किए जाते हैं। हम समाश्रयण मॉडल के प्राचलों का आकलन इस अवधारणा के तहत करते हैं कि आँकड़े किसी प्रायिकता बंटन फलन (उदाहरण के लिए, प्रसामान्य बंटन) का अनुसरण करते हैं।

प्रायिकता फलन किसी भी प्रायिकता बंटन फलन का अनुसरण कर सकता है; सिर्फ प्रसामान्य बंटन का ही नहीं। अपने सांख्यिकी पाठ्यक्रम से याद करें कि प्रायिकता बंटन फलन में हमें प्राचल दिए जाते हैं और हम किसी विशिष्ट आँकड़ा समुच्चय की विद्यमानता की संभावना का पता लगाते हैं। अधिकतम संभावना (ML) विधि में हम इसके विपरीत करते हैं, यथा हमें आँकड़े दिए जाते हैं और हम किन्हीं प्राचलों का आकलन कर रहे होते हैं।

आघूर्ण विधि (MOM) आघूर्ण जनक फलन (MGF) गुणधर्मों का प्रयोग करती है। 'आघूर्ण' की अवधारणा से परिचय BECC 107 की इकाई 4 में कराया गया है। कुछ प्रायिकता बंटनों के आघूर्ण जनक फलन का प्रयोग प्राचलों के आकलन के लिए किया जाता है। आघूर्ण विधि काफी उन्नत है और इस पाठ्यक्रम के विषय–क्षेत्र से बाहर है।

3.4 अस्वीकृति क्षेत्र और त्रुटियों के प्रकार

पिछली इकाई में हमने बिंदु आकलन और अंतराल आकलन के विषय में चर्चा की थी। परिकल्पना परीक्षण और अंतराल आकलन के पीछे अंतर्निहित विचार एक ही है। स्मरण करें कि किसी भी

विश्वास्यता अंतराल का निर्माण किसी नियत विश्वास्यता स्तर वाले प्रतिदर्श माध्य के इर्दगिर्द ही किया जाता है। तदनुसार, 95 प्रतिशत के विश्वास्यता स्तर का तात्पर्य यह होगा है कि 95 प्रतिशत उदाहरणों में समष्टि माध्य प्रतिदर्श माध्य से आकलित विश्वास्यता अंतराल में ही रहेगा। इसमें यह निहित है कि 5 प्रतिशत मामलों में समष्टि माध्य विश्वास्यता अंतराल के भीतर नहीं रहेगा।

यहाँ ध्यान दें कि जब समष्टि माध्य विश्वास्यता अंतराल के भीतर नहीं रहता है तो हमारे परीक्षण प्रतिदर्शज द्वारा शून्य–स्तरीय परिकल्पना को अस्वीकार कर दिया जाना चाहिए।

3.4.1 वृहद प्रतिदर्शों के लिए अस्वीकृति क्षेत्र

आइए, अब हम क्रांतिक क्षेत्र की अवधारणा की व्याख्या करते हैं। प्रतिदर्श माध्य ($\bar{x}$) का प्रतिदर्शन बंटन माध्य μ और मानक विचलन $\frac{\sigma}{\sqrt{n}}$ के साथ प्रसामान्य बंटन का अनुसरण करता है। प्रतिदर्शन बंटन के मानक विचलन को ही 'मानक त्रुटि' के रूप में जाना जाता है।

तदनुसार, $\bar{x}$ को एक मानक सामान्य चर, z, में परिवर्तित किया जा सकता है, ताकि यह माध्य 0 और मानक विचलन 1 के साथ प्रसामान्य बंटन का अनुसरण करे।

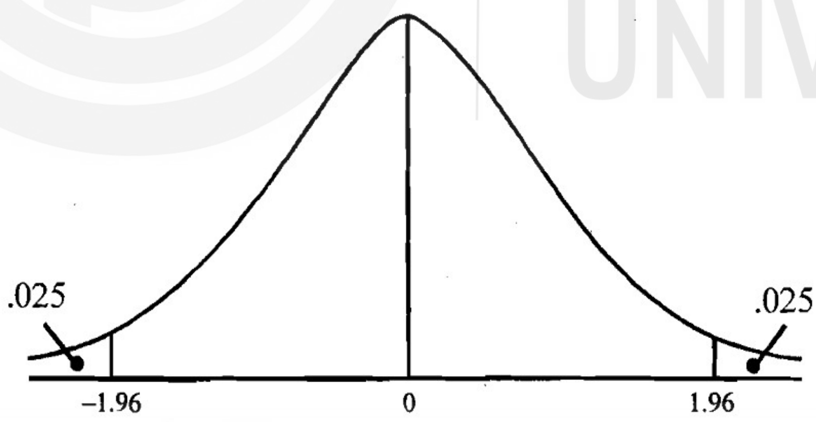
संकेताक्षरों में,

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \text{ और}$$

$$z \sim N(0,1)$$

याद करें कि मानक सामान्य वक्र के तहत क्षेत्र चर z द्वारा ग्रहण किए गए मानों की भिन्न श्रेणी के लिए प्रायिकता दर्शाता है। इन प्रायिकताओं को मानक सामान्य वक्र के तहत क्षेत्र के रूप में प्रस्तुत किया जाता है।

आइए, नीचे चित्र 3.1 में दिए गए मानक सामान्य वक्र के माध्यम से क्रांतिक क्षेत्र या अस्वीकृति क्षेत्र की अवधारणा की व्याख्या करें। जब हमारे पास 95 प्रतिशत का कोई विश्वास्यता गुणांक होता है तो मानक सामान्य वक्र के अंतर्गत आने वाला क्षेत्र भी 95 प्रतिशत ही होता है। तदनुसार, वक्र के नीचे का 95 प्रतिशत क्षेत्र $-1.96 \le z \le 1.96$ से घिरा है। शेष 5 प्रतिशत क्षेत्र $z \le -1.96$ और $z \ge 1.96$ द्वारा आच्छादित है। इस प्रकार मानक सामान्य वक्र के दोनों ओर 2.5 प्रतिशत क्षेत्र अस्वीकृति क्षेत्र का निर्माण करता है। यह क्षेत्र चित्र 3.1 में स्पष्ट दर्शाया गया है। यदि प्रतिदर्श माध्य अस्वीकृति क्षेत्र में आता है तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर देते हैं।



चित्र 3.1: क्रांतिक क्षेत्र

3.4.2 एक–पुच्छ और दो–पुच्छ परीक्षण

चित्र 3.1 में हमने मानक सामान्य वक्र के दोनों ओर अस्वीकृति क्षेत्र दर्शाया है। तथापि, अनेक उदाहरणों में हम मानक सामान्य वक्र के एक ओर (या तो बाएँ या फिर दाएँ) अस्वीकृति क्षेत्र रख सकते हैं। याद रखें कि यदि सार्थकता का स्तर α हो तो दो–पुच्छ परीक्षण के लिए मानक

सामान्य वक्र के दोनों ओर $\frac{\alpha}{2}$ क्षेत्र रखा जाता है। दूसरी ओर, यदि हमारा परीक्षण एक–पुच्छ परीक्षण हो तो α क्षेत्र को मानक सामान्य वक्र के एक ओर ही रखा जाता है। इस प्रकार एक–पुच्छ और दो–पुच्छ परीक्षण के लिए क्रांतिक मान भिन्न–भिन्न होते हैं।

एक–पुच्छ अथवा दो–पुच्छ परीक्षण का चयन वैकल्पिक परिकल्पना के निर्माण पर निर्भर करता है। जब वैकल्पिक परिकल्पना $H_A : \bar{x} \neq \mu$ प्रकार की होती है तो हमारे पास दो–पुच्छ परीक्षण होता है क्योंकि $\bar{x}$ का मान चर μ से अधिक अथवा कम हो सकता है। दूसरी ओर, यदि वैकल्पिक परिकल्पना $H_A : \bar{x} < \mu$ प्रकार की हो तो संपूर्ण अस्वीकृति मानक सामान्य वक्र के बाईं ओर होती है। इसी प्रकार, यदि वैकल्पिक परिकल्पना $H_A : \bar{x} > \mu$ प्रकार की हो तो संपूर्ण अस्वीकृति मानक सामान्य वक्र के दाईं ओर होती है।

चर z के लिए क्रांतिक मान सार्थकता के स्तर पर निर्भर करते हैं। इस पुस्तक के अंत में परिशिष्ट तालिका 14.1 में सार्थकता के कुछ निर्दिष्ट स्तरों (α) के लिए ये क्रांतिक मान दिए गए हैं।

3.4.3 लघु प्रतिदर्शों के लिए अस्वीकृति क्षेत्र

लघु प्रतिदर्शों ($n \leq 30$) के उदाहरण में यदि समष्टि मानक विचलन ज्ञात हो तो हम परिकल्पना परीक्षण के लिए z–प्रतिदर्शज प्रयोग करते हैं। दूसरी ओर, यदि समष्टि मानक विचलन ज्ञात न हो तो हम t–प्रतिदर्शज प्रयोग करते हैं। यही मानदंड परिकल्पना परीक्षण पर भी लागू होते हैं।

लघु प्रतिदर्शों के उदाहरण में यदि समष्टि मानक विचलन ज्ञात हो तो परीक्षण प्रतिदर्शज निम्नवत् होगा –

$$z = \frac{|\bar{x} - \mu|}{\sigma / \sqrt{n}} \qquad \text{... (3.1)}$$

दूसरी ओर, यदि समष्टि मानक विचलन ज्ञात न हो तो परीक्षण प्रतिदर्शज निम्नवत् होगा –

$$t = \frac{|\bar{x} - \mu|}{s / \sqrt{n}} \qquad \text{... (3.2)}$$

बहरहाल, t–बंटन के उदाहरण में वक्र के नीचे का क्षेत्र (जिसका अर्थ प्रायिकता होता है) स्वतंत्रता की कोटि के अनुसार बदलता रहता है। तदनुसार, चर t का क्रांतिक मान ज्ञात करते समय हमें स्वतंत्रता की कोटि को ध्यान में रखना चाहिए। निष्कर्षतः चर t का क्रांतिक मान ज्ञात करते समय हमें दो बातें याद रखनी चाहिए, यथा – i) सार्थकता का स्तर, और ii) स्वतंत्रता की कोटि।

3.5 त्रुटियों के प्रकार

परिकल्पना परीक्षण में हम विश्वास्यता की कोटि विशेष तक ही किसी परिकल्पना को अस्वीकार करते हैं अथवा अस्वीकार नहीं करते हैं। जैसा कि हम जानते हैं, 0.95 के विश्वास्यता गुणांक का अर्थ होता है कि 100 प्रतिदर्शों में से 95 प्रतिदर्शों में प्राचल स्वीकृति क्षेत्र के भीतर रहता है और 5 प्रतिशत प्रतिदर्शों में प्राचल अस्वीकृति क्षेत्र में रहता है। इस प्रकार 5 प्रतिशत उदाहरणों में प्रतिदर्श समष्टि से लिया जाता है परंतु प्रतिदर्श माध्य समष्टि माध्य से बहुत भिन्नता दर्शाता है। ऐसे उदाहरणों में प्रतिदर्श समष्टि का होता है परंतु हमारी परीक्षण प्रक्रिया इसे निरस्त कर देती है।

स्पष्टतः हम एक ऐसी त्रुटि करते हैं कि H_0 सत्य सिद्ध होता है परंतु फिर भी उसे अस्वीकार कर दिया जाता है। इसे ही 'टाइप–I त्रुटि' कहा जाता है। इसी प्रकार, ऐसी परिस्थितियाँ भी हो सकती

हैं जब H_0 सत्य सिद्ध न हो, परंतु प्रतिदर्श जानकारी के आधार पर हम इसे अस्वीकार नहीं करते हैं। इस प्रकार की त्रुटि को 'टाइप–II त्रुटि' कहा जाता है (देखें तालिका 3.1)।

यहाँ ध्यान दें कि टाइप–I त्रुटि यह निर्दिष्ट करती है कि हम किस सीमा तक त्रुटि को सहन करने की स्थिति में हैं। टाइप–I त्रुटि सार्थकता के स्तर के बराबर होती है, और इसे चर α से दर्शाया जाता है। याद रखें कि विश्वास्यता गुणांक $1-\alpha$ के बराबर होता है।

तालिका 3.1: त्रुटियों के प्रकार

	H_0 सत्य	H_0 सत्य नहीं
H_0 निरस्त करें	टाइप–I त्रुटि	सही निर्णय
H_0 निरस्त न करें	सही निर्णय	टाइप–II त्रुटि

किसी भी टाइप–I त्रुटि करने की प्रायिकता को चर α से निरूपित किया जाता है और इसे 'सार्थकता का स्तर' कहा जाता है। दूसरी ओर, टाइप–II त्रुटि करने की प्रायिकता को बीटा (β) कहा जाता है। तदनुसार,

टाइप–I त्रुटि = α = प्रायिकता (H_0 | H_0 का अस्वीकरण सत्य है)

टाइप–II त्रुटि = β = प्रायिकता (H_0 | H_0 का स्वीकरण असत्य है)

बोध प्रश्न 1

1) एक–पुच्छ परीक्षण और दो–पुच्छ परीक्षण के बीच अंतर स्पष्ट करें।

....

....

....

....

2) टाइप–I त्रुटि और टाइप–II त्रुटि के बीच अंतर स्पष्ट करें।

....

....

....

....

3) मान लीजिए कि किसी व्यक्ति का कोलेस्ट्रॉल स्तर सामान्य रूप से 180 के माध्य और 20 के मानक विचलन के साथ बंटित किया जाता है। मान 225 से अधिक के कोलेस्ट्रॉल स्तर को 'स्वस्थ नहीं' मानकर रोग निर्णय किया जाता है।

   a) यहाँ टाइप–I त्रुटि होने की प्रायिकता क्या है?

   ....

   ....

   ....

   ....

b) यदि हम चाहते हों कि टाइप–I त्रुटि की प्रायिकता 2% हो तो लोगों को किस स्तर पर 'स्वस्थ नहीं' माना जाना चाहिए?

....

....

....

....

3.6 परीक्षण की शक्ति

जैसा कि ऊपर बताया गया है, परिकल्पना परीक्षण में टाइप–I और टाइप–II त्रुटियाँ पाई जाती हैं। तदनुसार, दो प्रकार के जोखिम सामने आते हैं, यथा –

(i) हानिभय α जो कि यह प्रायिकता निरूपित करता है कि जब शून्य–स्तरीय परिकल्पना सत्य हो और निरस्त नहीं की जानी चाहिए तो भी उसे निरस्त कर दिया जाता है।

(ii) हानिभय β जो कि यह प्रायिकता निरूपित करता है कि जब शून्य–स्तरीय परिकल्पना वास्तव में असत्य हो तो भी उसे निरस्त नहीं किया जाता है।

परीक्षण की शक्ति का संदर्भ $(1 - \beta)$ के रूप में दिया जाता है, जो कि β का संपूरक होता है। यह मूल रूप से टाइप–II त्रुटि न करने की प्रायिकता होती है।

किसी भी 95% विश्वास्यता गुणांक का अर्थ होता है कि हम टाइप–I त्रुटि करने की अधिकतम 5% संभावना को ही स्वीकार करने के लिए तैयार होंगे। किसी भी सत्य परिकल्पना को हम 100 में से 5 से अधिक बार अस्वीकार नहीं करना चाहेंगे। इसे ही 'सार्थकता का 5 प्रतिशत स्तर' कहा जाता है।

परीक्षण की शक्ति वास्तविक समष्टि माध्य और परिकल्पित माध्य के बीच अंतर की सीमा पर निर्भर करती है। यदि अंतर बड़ा है तो परीक्षण की शक्ति अंतर के छोटे होने की तुलना में बहुत अधिक होगी।

इसीलिए, सार्थकता के स्तर α का चयन बहुत महत्वपूर्ण होता है। चर α के बड़े मान का चयन करना शून्य–स्तरीय परिकल्पना को अस्वीकार करना सरल बना देता है, जिससे परीक्षण की शक्ति $(1 - \beta)$ बढ़ जाती है।

इसके साथ ही प्रतिदर्श आमाप बढ़ने से आकलों में सटीकता बढ़ जाती है और समष्टि प्राचल और प्रतिदर्श के बीच अंतर का पता लगाने की क्षमता बढ़ जाती है, जिससे परीक्षण की शक्ति बढ़ जाती है।

3.7 प्राचल आकलन के उपागम

सांख्यिकीय परिकल्पना परीक्षण में आकलन सिद्धांत ऐसे अनुभवजन्य आँकड़ों के मापन पर आधारित प्राचलों के मानों का आकलन करने से संबंध रखता है जिसमें कोई यादृच्छिक घटक विद्यमान होता है। आकलन की विधि के लिए किसी शून्य–स्तरीय परिकल्पना और किसी अनुरूप वैकल्पिक परिकल्पना को सूत्र के रूप में व्यक्त किए जाने की आवश्यकता होती है, जिसे शून्य–स्तरीय परिकल्पना के संबंध में निर्णय लेने के लिए प्रयोग किए जाने वाले दो उपागमों के आधार पर आगे निरस्त कर दिया जाता है अथवा अस्वीकार नहीं किया जाता है। अगले पाठांश में ऐसी ही दो विधियों का वर्णन किया गया है।

3.7.1 सार्थकता का परीक्षण उपागम

परिकल्पना परीक्षण हेतु 'सार्थकता का परीक्षण' उपागम के लिए किसी भी परीक्षण प्रतिदर्शज का प्रयोग किया जा सकता है। चलिए, यहाँ हम t–प्रतिदर्शज पर विचार करते हैं।

$$t = \frac{\acute{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} \qquad \text{... (3.3)}$$

यदि $\acute{X}$ और μ_X के बीच का अंतर कम है तो पद | t | का मान भी कम ही होगा, जहाँ | t | ही t–प्रतिदर्शज का निरपेक्ष मान होगा। यह ध्यान देने की बात है कि यदि $\acute{X} = \mu_X$ हो तो $t = 0$ होगा। इस उदाहरण में हम शून्य–स्तरीय परिकल्पना को अस्वीकार नहीं करते हैं। जब | t | बड़ा हो जाता है तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार करने के लिए अधिक इच्छुक हो सकते हैं।

उदाहरण : मान लीजिए कि किसी आँकड़ा समुच्चय $\acute{X} = 23.25$, $S_X = 4.49$ और $n = 28$ के लिए हमारी शून्य–स्तरीय और वैकल्पिक परिकल्पनाएँ निम्नवत् हैं –

H_0: $\mu_X = 18.5$ और H_A: $\mu_X \neq 18.5$

$$\therefore \qquad t = \frac{23.25-1\ \ .5}{9.49/\sqrt{28}} = 2.6486 \qquad \text{... (3.4)}$$

अब हमें शून्य–स्तरीय परिकल्पना को अस्वीकार करने की प्रायिकता, यथा α (टाइप–I त्रुटि किए जाने की प्रायिकता) को निर्दिष्ट करने की आवश्यकता है। चलिए, चर α को 5% पर तय कर लेते हैं।

H_0: $\mu_X = 18.5$

H_A: $\mu_X \neq 18.5$ (दो–पुच्छ परीक्षण)

चूँकि परिकलित t–मान 2.6486 है, यह मान t–बंटन के दाईं पुच्छ वाले क्रांतिक क्षेत्र में स्थित है। अतएव हम इस शून्य–स्तरीय परिकल्पना (H_0) को अस्वीकार कर देते हैं कि वास्तविक समष्टि माध्य 18.5 है।

किसी परीक्षण के सांख्यिकीय रूप से सार्थक होने का अर्थ होगा कि हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर सकते हैं। इसमें यह अर्थ निहित है कि प्रतिदर्श मान और सार्थकता मान (जिसे तालिकाबद्ध मान भी कहा जाता है) के बीच देखे गए अंतर की प्रायिकता कम नहीं है और ऐसा संयोगवश नहीं है। दूसरी ओर, किसी परीक्षण के सांख्यिकीय रूप से सार्थक न होने का अर्थ होगा कि हम शून्य–स्तरीय परिकल्पना को अस्वीकार नहीं करते हैं। यहाँ प्रतिदर्श मान और सार्थकता मान के बीच अंतर प्रतिदर्शन विचरण अथवा संयोगवश हो सकता है।

3.7.2 विश्वस्तता अंतराल उपागम

चलिए, मान लेते हैं कि सार्थकता का स्तर या टाइप–I त्रुटि आने की संभावना $\alpha = 5\%$ पर तय की गई है। यह भी मान लीजिए कि वैकल्पिक परिकल्पना द्विपक्षीय है। अब मान लें कि हम t–बंटन प्रयोग करते हैं क्योंकि विचरण ज्ञात नहीं है। दी गई t–तालिका से हम t का क्रांतिक मान स्वतंत्रता की 8 कोटि के साथ $\alpha = 5\%$ पर $(n-K) = (10-2)$ पाते हैं। हम यह मान 2. 360 पाते हैं। इस प्रकार हम विश्वास्यता अंतराल का निर्माण निम्नवत् करते हैं –

$$P(-2.360 \leq t \leq 2.306) = 0.95 \qquad \text{... (3.5)}$$

यह प्रायिकता कि t–मान सीमाओं ($-2.360 \leq t \leq 2.360$) के बीच अवस्थित है, 0.95 अथवा 95% है। मान –2.360 और 2.360 ही क्रांतिक t–मान हैं।

यदि हम समीकरण (3.2) से t को प्रतिस्थापित करते हैं तो हमें प्राप्त होता है –

$$P\left(-2.306 \leq \frac{b_2-\beta_2}{\mathrm{SE}(\mathrm{b}_2)} \leq 2.306\right) = 0.95 \qquad \text{... (3.6)}$$

जैसा कि हम इकाई 4 में देखेंगे, मानक त्रुटि (SE) (b_2) का मान $\frac{\hat{\sigma}}{\sqrt{\Sigma x_i^2}}$ होता है।

यदि हम उपर्युक्त मान को समीकरण (3.6) में प्रतिस्थापित करते हैं और हमें प्राप्त होने वाले पदों को फिर से व्यवस्थित करते हैं तो हमें यह समीकरण प्राप्त होता है –

$$P\left(b_2 - 2.306\frac{\hat{\sigma}}{\sqrt{\Sigma x_i^2}} \leq \beta_2 \leq b_2 + 2.306\frac{\hat{\sigma}}{\sqrt{\Sigma x_i^2}}\right) = 0.95 \qquad \text{... (3.7)}$$

समीकरण (3.7) प्राचल β_2 के लिए एक 95:% विश्वास्यता अंतराल प्रदान करता है। इस प्रकार के विश्वास्यता अंतराल को स्वीकृति के क्षेत्र (H_0) के रूप में जाना जाता है। विश्वास्यता अंतराल के बाहर के क्षेत्र को अस्वीकृति क्षेत्र (H_A) के रूप में जाना जाता है।

यदि विश्वास्यता अंतराल में प्राचल β_2 का मान शामिल हो तो हम परिकल्पना को अस्वीकार नहीं करते हैं। किंतु यदि प्राचल विश्वास्यता अंतराल के बाहर अवस्थित हो तो हम शून्य–स्तरीय परिकल्पना को अस्वीकार कर देते हैं।

बोध प्रश्न 2

1) परीक्षण की शक्ति से आप क्या समझते हैं?

....

....

....

....

2) विश्वास्यता अंतराल का निर्माण कैसे होता है? स्पष्ट करें।

....

....

....

....

3.8 सार–संक्षेप

इस इकाई में समष्टि प्राचलों के संबंध में सांख्यिकीय अनुमिति की प्रक्रिया का विस्तृत वर्णन किया गया। समष्टि प्राचलों के परिकल्पना परीक्षण के लिए दो उपागम प्रचलन में हैं – सार्थकता का परीक्षण उपागम और विश्वास्यता अंतराल उपागम।

इकाई में यह भी बताया गया कि परिकल्पना के परीक्षण में त्रुटियाँ भी शामिल होती हैं। किसी भी परिकल्पना की स्वीकृति अथवा अस्वीकृति के संबंध में निर्णय लेते समय दो प्रकार की त्रुटियाँ हो सकती हैं – टाइप–I त्रुटि और टाइप–II त्रुटि।

परीक्षण की शक्ति कोई भी टाइप–II त्रुटि न करने की प्रायिकता $(1 - \beta)$ होती है।

3.9 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1. पाठांश 3.4.2 पढ़ें और उत्तर दें।
2. त्रुटियों के प्रकार हमने तालिका 3.1 में दर्शाए हैं। आप उन्हीं को विस्तार से समझाएँ।
3. a) इसका परीक्षण करने के लिए हम z–प्रतिदर्शज का प्रयोग करते हैं, यथा

   $z = (X - \mu) / \sigma$,
   $z = (225 - 180) / 20 = 2.25$
   b) मान 2.25 के z–मान के अनुरूप क्षेत्र 0.0122 है, जो कि टाइप–I त्रुटि होने की प्रायिकता दर्शाता है। मान 2% के रूप में पुच्छ का क्षेत्र $Z = 2.05$ से मेल खाता है।

   $Z = (X - \mu)/\sigma$

   $2.05 = (X - \mu)/ 20$,

   यथा $(X - \mu) = 2.05 * 20 = 41$

   $X = 41 + 180 = 221$

बोध प्रश्न 2

1. पाठांश 3.6 पढ़ें और उत्तर दें।
2. पाठांश 3.7.2 पढ़ें और उत्तर दें।