

NALANDA OPEN UNIVERSITY

Course : M.A Psychology, Part-I

Paper : Paper-VI

**Prepared by : Dr. (Prof.) Prabha Shukla
Retd. Professor of Psychology, Patna University and
Chief Co-ordinator, School of Social Sciences,
Nalanda Open University**

Topic : प्रसरण-विश्लेषण (Analysis of Variance (ANOVA))

11.1 परिचय एवं परिभाषा (Introduction and Definition) :

सांख्यिकी, विज्ञान (Science) का शाखा है। यह शाखा शुद्ध विज्ञान (Pure Science) तथा व्यवहारिक विज्ञान (Applied Science) का समिश्रण है। सांख्यिकी आँकड़ों या संख्यात्मक प्रदत्तों (Numerical data) का व्यवस्थित

विश्लेषण करता है। उसकी सहायता से प्राप्त आँकड़ों (Data) का उपयोग वर्गीकरण, विश्लेषण, व्याख्या; आदि के लिए किया जाता है।

प्रसरण-विश्लेषण (Anova) सांख्यिकी की एक महत्वपूर्ण प्रविधि है। यह प्राप्त तुलनात्मक अध्ययन प्रस्तुत करता है। इसके प्रयोग से कई समूहों के प्राप्त आँकड़ों (Data) का तुलनात्मक अध्ययन किया जा सकता है। जब दो या इससे अधिक समूहों के अध्ययनों के अंतर की सार्थकता जाँचनी हो तब प्रसरण विश्लेषण बहुत उपयोगी सिद्ध होता है। इसके द्वारा विभिन्न समूहों के बीच के अंतर एक परीक्षण के रूप में किया जा सकता है। उससे पूरे निष्कर्ष को संक्षिप्त (Summary) रूप में प्रस्तुत किया जाता है। यह प्राचलिक सांख्यिकी (Parametric Statistics) की विधि है। इसे एफ-टेस्ट (F-Test) के नाम से भी जाना जाता है। प्राचलिक सांख्यिकी का अर्थ है कि इसके आँकड़े (Data) जनसंख्या का प्रतिनिधित्व करते हों। साथ ही प्रसामान्य वितरण (Normal Distribution) के अनुरूप हो। विभिन्न समूहों के मध्य के अंतर को अन्य विधियों द्वारा (T-Test) भी जाना जा सकता है। परन्तु इसमें अधिक कठिनाई होती है। जबकि एफ टेस्ट (F-Test) या प्रसरण विश्लेषण की मदद से इसे सरलता से किया जा सकता है। इससे स्पष्ट और क्रमबद्ध (Systematic) निष्कर्ष प्राप्त होता है।

यह विधि अनेक घरों के प्रभाव का अध्ययन एक साथ कर सकता है। उदाहरण के लिए बाजार सर्वेक्षण करना हो अथवा विज्ञापन के प्रभाव का अध्ययन करना है। तब यह विभिन्न चरों (Variables) के प्रभाव का व्यवस्थित निष्कर्ष प्रस्तुत करता है। विज्ञापन का बिक्री (Sales) ग्राहकों के मनोवृत्ति (Attitude), जानकारी (Awareness), फैशन आदि पर क्या प्रभाव पड़ता है। इस प्रकार के अध्ययन में प्रसरण विश्लेषण का प्रयोग करना महत्वपूर्ण होता है।

प्रसरण विश्लेषण के आरंभिक प्रयोग का श्रेय सांख्यिकी तथा अनुवाशिकता अध्ययन के विद्वान आर० ए० फिशर (R.A. Fisher) को जाता है। यह सांख्यिकी के क्षेत्र में फिशर का महत्वपूर्ण योगदान है। उन्होंने 1920 तथा 1930 के दशकों में इसे विकसित किया। इसी कारण से इसे फिशर के प्रसरण विश्लेषण (Fisher's Anova or Fisher's Analysis of Variance) के नाम से भी जाना जाता है। इसे संक्षिप्त रूप से एफ टेस्ट (F-Test) भी कहा जाता है। प्रसरण विश्लेषण सांख्यिकी मॉडलों (Statistical Models) और उनसे संबंधित विधियों को बेहतरीन तरीके से दर्शाता है।

एनोवा को उनके आँकड़ों (Data) के अनुसार पर निम्नलिखित मॉडलों (Models) में रखा जा सकता है।

◆ मॉडल-1 निश्चित प्रभाव मॉडल (Fixed Effect Model)—जब आँकड़े (Data) प्रसामान्य वितरण (Normal Distribution) से प्राप्त किया जाता है तथा वे मात्र मध्यमान (Mean) में अलग हो सकते हैं।

◆ मॉडल-2 यादृच्छिक प्रभाव मॉडल (Random Effect Model)—जब आँकड़े जनसंख्या के विभिन्न स्तरों (Hierarchy) को व्यक्त करते हैं। तब इसे मॉडल 2 माना जाता है।

◆ मॉडल-2 ऐसे मॉडल उपरोक्त दोनों प्रकार के परिस्थितियाँ का मिश्रण होता है।

प्रसरण-विश्लेषण की व्याख्या अनेक विद्वानों ने दिया है। पेटर स्ट्रेटोन एवं निवकी हस (Peter Stratton and Nicky Hayes, 1991) के कथन अनुसार, “प्रसरण-विश्लेषण एक सांख्यिकीय कार्य प्रणाली है, जिसके द्वारा यह निर्धारित किया जाता है कि प्राप्तियों के समूह एक दूसरे से भिन्न हैं या नहीं।” (Analysis of Variance is a statistical procedure to test whether groups of scores differ from each other.)

भारतीय विद्वानों द्वारा व्यक्त किए गए परिभाषाओं में डॉ० मोहसिन (Dr. Mohsin 1985) की परिभाषा उल्लेखनीय है। उन्होंने एनोवा को निम्नलिखित प्रकार से व्यक्त किया है—“प्रसरण विश्लेषण कई समूहों के बीच अन्तर का एक समग्र परीक्षण प्रस्तुत करता है जिसका व्यवहार साथ-साथ किया जा सकता है।” (Analysis of Variance provides a blanket test of difference among the several groups which can be applied simultaneously.)

उपरोक्त परिभाषाएँ प्रसरण-विश्लेषण की विशेषताएँ बताती हैं। प्रसरण-विश्लेषण की कुछ महत्वपूर्ण अभिधारणाएँ भी होती हैं। प्रसरण-विश्लेषण के उपयोग के लिए इन शर्तों का ध्यान रखना आवश्यक है। प्रसरण-विश्लेषण के प्रयोग के लिए महत्वपूर्ण शर्तें या अभिधारणाएँ (Assumptions) निम्नलिखित हैं—

अभिधारणाएँ (Assumptions)

एनोवा निकालते वक्त हम निम्नलिखित बातों को मान कर चलते हैं—

1. विषय की स्वतंत्रता (Independence of Case) यह शैली (Design) की आवश्यकता है।

- सामान्यता (Normality) हर प्रतिदर्श समूह का बँटवारा सामान्य है।
 - भिन्नता में समांगता (Homogeneity of Variance) हर समूह के जानकारी (Data) में अंतर एक समान है। अर्थात् विचलन समजातीय होना चाहिए।
- ये तीनों अभिधारणायें यह बताती हैं कि त्रुटि अवशिष्ट (Error Residuals) स्वतंत्र, समान एवं सामान्य तरीके से बँटे हैं।

11.2 महत्त्व (Importance of Anova) :

यह प्राचलिक सांख्यिकीय का महत्त्वपूर्ण विधि है। इसकी उपयोगता दो या दो से अधिक समूहों के लिए किया जाता है। इसके मदद में अंतर के साथ-साथ उसकी सार्थकता भी स्पष्ट होती है। इसका उपयोग मनोविज्ञान, विभिन्न शोध (रिसर्चों) गणित के क्षेत्र में सर्वेक्षण जैसे अनेक क्षेत्रों में होता है। इसमें प्रतिदर्श यादृच्छिक चयन विधि (Sample Selection by Random Sampling) द्वारा लिया जाता है। अतः जनसंख्या के सभी इकाईयों के समान रूप से लिया जाता है। अतः पक्षपात (Biases) की सम्भावना कम रहती है। एनोवा द्वारा विभिन्न समूहों के अंतर को व्यवस्थित रूप से संक्षेप (Summary) में स्पष्ट किया जा सकता है। जिसकी सहायता से एफ-परीक्षण कर एफ (F) निकला जाता है। फिर टेबल की सहायता से प्राप्त निष्कर्ष की सार्थकता जाँची जाती है।

11.3 उपयोग (Application of Anova) :

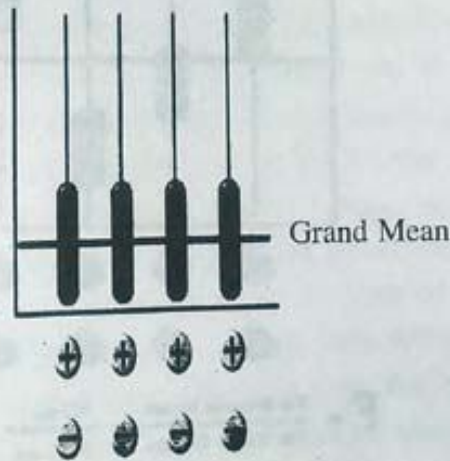
प्रसरण विश्लेषण का उपयोग चरों के नजर में रखते हुए किया जाता है। जो निम्नलिखित हैं—

- सरल प्रसरण विश्लेषण (Simple Analysis of Variance)—सरल एनोवा का अर्थ है एक समूह/चर पर आधारित प्रसरण-विश्लेषण। जब मात्र एक समूह/चर (Group Variable) के विभिन्न समूहों के बीच के अंतर की सार्थकता जाँचनी हो तब इसका उपयोग किया जाता है। इसे एकमार्गी प्रसरण-विश्लेषण (One Way Analysis of Variance) भी कहा जाता है।

चित्र-1 चित्रों के द्वारा एनोवा की व्याख्या

Understanding ANOVA Visually

MS_{Between} MS_{Within} Instructions

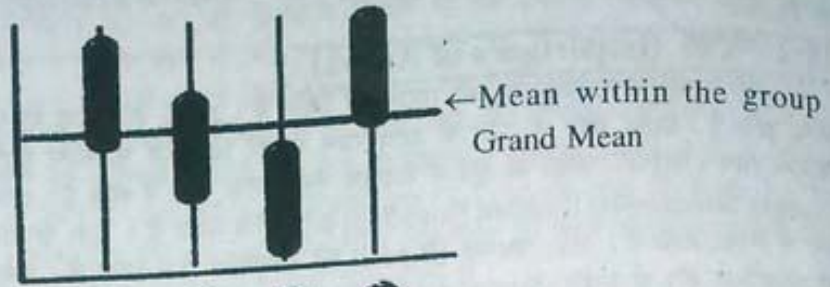


$$F = \frac{\text{Var Between Means}}{\text{Var Within Groups}} = \frac{MS_{\text{bet}}}{MS_{\text{w/in}}} = \text{---}$$

$$F = \frac{\text{Variation Between Means}}{\text{Variation Within Groups}}$$

Understanding ANOVA Visually

● MSBetween ● MSWithin ● Instructions

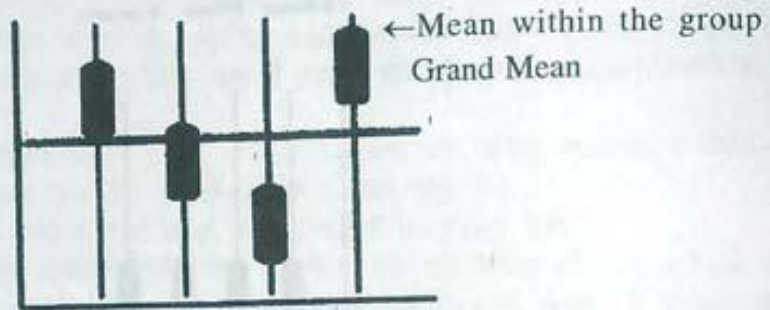


$$F = \frac{\text{Var Between Means}}{\text{Var Within Groups}} = \frac{MS_{\text{bet}}}{MS_{\text{within}}} = \frac{D}{B}$$

$$F = \frac{D}{B}$$

Understanding ANOVA Visually

● MSBetween ● MSWithin ● Instructions



$$F = \frac{\text{Var Between Means}}{\text{Var Within Groups}} = \frac{MS_{\text{bet}}}{MS_{\text{within}}} = \frac{D}{B}$$

$$F = \frac{D}{B}$$

चित्र-3

2. जटिल प्रसरण-विश्लेषण (Complex Analysis of Variance)-जब एक से अधिक स्वतंत्र चरों (Independent Variables) को परिचालित (Manipulate) कर उनके प्रभावों की सार्थकता का जाँच करने की

आवश्यकता पड़ती है। तब जटिल प्रसरण विश्लेषण (Complex Anova) का प्रयोग किया जाता है। इसे द्विमागी प्रसरण-विश्लेषण (Two way Analysis of Variance) भी कहा जाता है। इसमें विभिन्न कारक अभिकल्पना (Factorial Design) का उपयोग किया जाता है।

11.4 सरल प्रसरण-विश्लेषण (Simple Analysis of Variance) :

इसे एकमागी प्रसरण-विश्लेषण (One way Analysis of Variance) के नाम से भी जाना जाता है। इसकी उपयोगिता सिर्फ एक स्वतंत्र चर (Independend Variable) के दो या दो से अधिक समूहों के बीच के अन्तर की सार्थकता जानने के लिए किया जाता है। दूसरे शब्दों में, दो समूहों से प्राप्त मध्यमानों में पाए जानेवाले अन्तर का अध्ययन किया जाता है। जिससे यह स्पष्ट होता है कि प्राप्त अन्तर सार्थक है या नहीं।

प्रसरण विश्लेषण जानने के लिए सबसे पहले सभी समूहों का मध्यमान (Grand Mean) निकाला जाता है चित्र संख्या-1। समूहों के प्रसरण के दो प्रमुख श्रोत होते हैं। पहला प्रसरण (Variation) प्रत्येक समूह के अन्दर (Within each Group) पाया जाता है। जैसा चित्र संख्या-2 और उसे स्पष्ट होता है। प्रत्येक चित्र में लाल रंग से चार समूह दर्शाये गए हैं। जिनका अपना-अपना मध्यमान (Mean within the group) हरे लाईन से उनपर दिखलाया गया है। प्रत्येक समूह के अंदर पाया जानेवाला प्रसरण (Variance within each group) व्यक्ति भिन्नता (Individual Difference) जैसे कारणों से होते हैं।

दूसरे प्रकार के प्रसरण समूहों के बीच के प्रसरण (Variation) के कारण होता है। अर्थात् विभिन्न समूहों के बीच का प्रसरण (Variation between the groups) इसका स्रोत होता है। रोनाल्ड फिशर (Ronald Fisher) ने इन्हीं प्रसरण के विश्लेषण (Analysis of Variance) को एफ परीक्षण (F-Test) के रूप में विकसित किया।

$$F = \frac{\text{Variation between groups}}{\text{Variation within groups}} = \frac{MS_{\text{Between}}}{MS_{\text{Within}}}$$

चित्र संख्या (2) दो MS Between/MS Within को नारंगी रंग और बैंगनी रंग से दिखलाया गया है। F = नीला बुलेट है जो F का मान लगभग 1 दिखा रहा है बुलेट के ऊपर अंकित स्केल से यह स्पष्ट है। इसी प्रकार चित्र संख्या 3 में चार लाल रंग के बुलेट चार समूहों को दर्शा रहे हैं। जिसके मध्य की हरी रेखा उन समूहों का मध्यमान है। लम्बे हरे रेखा द्वारा सभी समूहों के मध्यमान (Grand Mean Between the groups) अंकित है। चित्र से स्पष्ट है कि कौन समूह कुल मध्यमान (Grand Mean) के ऊपर या नीचे है। समूह के प्रसरण के आधार पर समूह के अंदर (within the groups) के प्रसरण और समूहों के बीच (between the groups) के प्रसरण का अध्ययन किया जा सकता है। F = MS Between/MS Within को नारंगी रंग और बैंगनी रंग के बुलेट से स्पष्ट देखा जा सकता है। यहाँ F = नीला बुलेट एफ के (F Ratio) के मान को दिखला रहा है। जो 3 और 4 अंकों के मध्य है नीचे बुलेट के ऊपर अंकित स्केल इसे बतला रहा है। साधारण प्रसरण-विश्लेषण का प्रयोग अलग-अलग प्रकार से होता है। यह प्राप्त आँकड़ों (Data) के अनुसार निर्धारित होता है कि कौनसी परिस्थिति में कौन Anova उपयोगी होगा। सरल प्रसरण-विश्लेषण (Simple Anova) को निम्नलिखित स्थानों अवस्थाओं (Condition) प्रयुक्त किया जा सकता है—

अवस्थाएँ (Conditions)

1. मात्र एक समूह का प्रसरण-विश्लेषण (Anova in case of Single Group)
2. मात्र दो समूहों का प्रसरण-विश्लेषण (Anova in case of only two groups)
3. दो से अधिक समूहों या मध्यमानों का प्रसरण-विश्लेषण (Anova is case of more than two groups or means)

11-4.1 मात्र एक समूह का प्रसरण-विश्लेषण (Anova in case of Single Group) :

इस प्रकार के एनोवा का प्रयोग वैसी स्थिति में किया जाता है, जब समूह एक होता है। परन्तु उनका अध्ययन अलग-अलग परिस्थिति में किया जाता है। अथवा उनके अलग-अलग व्यवहारों का अध्ययन करना हो। उदाहरण के लिए कक्षा VIII के पढ़ाई का प्रभाव 10 विद्यार्थियों के एक समूह के (किन्हीं दो कारण के प्राप्ताकों) पर करना हो। इसी प्रकार एक अन्य उदाहरण है-8 कर्मचारियों (एक सकत) के कार्यक्षमता का अध्ययन तीन अलग-अलग कार्य-कार्यक्षमता का अध्ययन तीन अलग-अलग कार्य-वातावरण में करना। ऐसे में यह निष्कर्ष प्राप्त किया जा सकता है कि क्या अलग-अलग कार्य वातावरण का प्रभाव सार्थक है या नहीं। कर्मचारियों को तीन कार्य-वातावरण खराब (Poor), औसत (Average) तथा अच्छा (Good) में कार्य के अवसर दिए जाते हैं। उनके उत्पादन के आँकड़ों की सहायता से जमा किया जाता है। इन आँकड़ों की सहायता से सांख्यिकी के प्रसरण-विश्लेषण प्रविधि की गणना की जाती है। एक समूह के दो परिस्थितियों से प्राप्त आँकड़ों के आधार पर प्रसरण-विश्लेषण देखा जा सकता है। सूत्रों (Formula) की सहायता से प्रसरण सोव-विश्लेषण से प्राप्त एफ (F) की सार्थकता टेबल (F Distribution) से देखा जा सकता है। सार्थक एफ से स्पष्ट होता है कि कर्मचारियों के कार्य क्षमता पर वातावरण का प्रभाव पड़ता है। अगर एफ सार्थक नहीं आता है (Non Significant F) तब इसका अर्थ होता है कि वातावरण का कोई प्रभाव कर्मचारियों के कार्य क्षमता पर नहीं पड़ता है। अर्थात् शुन्य परिकल्पना स्वीकृत (Null Hypothesis is supported) हो जाती है।

इस एनेत्वा का एक उदाहरण निम्नलिखित है-

8 कर्मचारियों द्वारा तीन अलग-अलग कार्य-वातावरण-अच्छा, औसत तथा खराब में कार्य करवाया गया। ऐसे में उनके उत्पादकता के आँकड़े जमा किए गए। एनोवा की सहायता से यह देखने का प्रयास किया गया कि कार्य वातावरण का क्या प्रभाव उत्पादन पर पड़ता है या नहीं।

कार्य-वातावरण (Working Environment)

कर्मचारी Workers	अच्छा Good X	औसत Average X ₂	खराब Poor x ₃	(Σ X R)	(Σ X R) ²	x ₁ ²	x ₂ ²	x ₃ ²
1	7	6	5	18	324	49	36	25
2	10	7	5	22	484	100	49	25
3	10	8	6	24	576	100	64	36
4	8	7	5	20	400	64	49	25
5	10	6	4	20	400	100	36	16
6	9	7	6	22	484	81	49	36
7	9	7	5	21	441	81	49	25
8	10	6	5	21	441	100	36	25
	Σ73	Σ54	Σ41		3550	675	368	231
					(Σ X R) ²	ΣS ₁ ²	ΣS ₂ ²	ΣS ₃ ²

विभिन्न चरणों में गणना (Calculation)

1. सबसे पहले तीनों अवस्थाओं के आँकड़ों को जोड़ कर Σ X R (x₁ + x₂ + x₃ or 7 + 6 + 5) निकाला गया।

2. $\Sigma X R$ का वर्ग $08 \times 18 = 324$) $(\Sigma X R)^2$ कॉलम में लिखा गया।

3. अगले चरण में x_1, x_2, x_3 का वर्ग कॉलमों में $x_1^2 \times 7 = 49$, $(x_2^2 \times 6 = 36)$; $(x_3^2 = 5 \times 5 = 25)$ लिखा गया। सभी कॉलम में उनका जोड़ $(x_1^2 = 5 \times 5 = 25)$ लिखा गया। सभी कॉलम में उनका जोड़ $(x_1^2 = 675$; $x_2^2 = 368$; $x_3^2 = 213)$ लिखा गया।

4. $(\Sigma X R)$ का जोड़ $(\Sigma x \Sigma x_1 + \Sigma x_2 + \Sigma x_3)$ निकाला गया। $(7 + 3 + 54 + 4 + 168)$

5. $\Sigma x^2 = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 = 675 + 368 + 213 = 1256$

6. Correction = $C = (\Sigma X)^2 / n = (168)^2 / 24 = 28224 / 24 = 1176.00$

7. Total Sum of square TSS = $\Sigma x^2 - C = 1256 - 1176 = 80$

8. Between sum of squares of Column SS = BSS

$$BSS = \left\{ \frac{(\Sigma x_1)^2}{N} + \frac{(\Sigma x_2)^2}{N_2} + \frac{(\Sigma x_3)^2}{N_3} \right\} - C = \left\{ \frac{(73)^2}{8} + \frac{(54)^2}{8} + \frac{(41)^2}{8} \right\} - 1176$$

$$= \{666.13 + 364.5 + 210.13\} - 1176 = 1240.76 - 1176 = 64.76$$

9. Within sum of squares or Row SS = WSS

$$\frac{(\Sigma \times R)^2}{K} - C = \frac{3550}{3} - 1176 = 1183.33 - 1176 = 7.33$$

10. Interaction SS or Residual SS = TSS - (Columnss + Row SS) = $80 - (64.76 + 7.33) = 80 - 72.09 = 7.91$

11. Degree of Freedom = df -

(a) Between SS (Columnss) के लिए $df = K - 1 = 3 - 2 = 2$

(b) WSS (Row SS) के लिए $df = r - 1 = 8 - 1 = 7$

(c) Interaction SS (Residual SS) के लिए $df = (K-1) \times (r-1) = (3-1) \times (8-1) = 2 \times 7 = 14$

(d) Total SS के लिए $df = N - 1 = 24 - 1 = 23$

प्रसरण-विश्लेषण का सारांश (Summary of Analysis of Variance)

Sources of Variance	Sum of Square	df	Mean Square of Variance	F	P
BSS (Columnss)	64.76	2	32.38	56.81	<.01
WSS (RowSS)	7.33	7	1.05	1.84	>.05
Interaction SS (Error)	7.91	14	.57		
Total	80.00	23			

$$F \text{ for Column} = \frac{\text{Column Variance}}{\text{Error Variance}} = \frac{1.05}{.57} = 1.84$$

$$F \text{ for Rows} = \frac{\text{Row Variance}}{\text{Error Variance}} = \frac{1.05}{.57} = 1.84$$

सार्थकता का स्तर (Level of Significance) प्राप्त एफ की सार्थकता जाँचने के लिए एफ टेबल (F Distribution Total) की मदद ली जाती है (इकाई के अंत में एक टेबल दिया गया है)। यहाँ F.01 स्तर पर सार्थक है (Significant at / <.01 Level)। इससे स्पष्ट होता है कि कर्मचारियों के कार्य क्षमता पर कार्य वातावरण (Working Environment) का प्रभाव पड़ रहा है।

11-1-2 मात्र दो समूहों का प्रसरण-विश्लेषण (Anova in case of Only two Groups) :

दूसरे प्रकार का सरल प्रसरण-विश्लेषण (Simple Anova) उसे कहते हैं। जिसमें सिर्फ किन्हीं दो समूहों के प्रसरण का विश्लेषण किया जाता है। इससे दो समूहों की सार्थकता जाँची जाती है। इससे स्पष्ट हो जाता है कि दो समूहों के बीच पाया गया अंतर संयोगवश है या सार्थक अंतर है। दो समूहों के बीच के अंतर के अध्ययन में एफ-परीक्षण (F-Test) के अलावा टी-परीक्षण (T-Test) भी प्रयुक्त हो सकता है। वास्तव में टी-परीक्षण तथा एफ परीक्षण (T & F Test) प्राचलित सांख्यिकी की प्रविधियाँ (Parametric Statistics Techniques) हैं। इन दोनों परीक्षणों का संबंध निम्नलिखित फार्मूला से स्पष्ट होता है।

$$t = \sqrt{F} \text{ और } F = t^2$$

दो समूहों के प्रसरण का विश्लेषण जाँचने के लिए दोनों समूह के एक चर को लिया जाता है। उदाहरण के लिए स्त्री और पुरुषों के 10-10 के दो समूहों के कार्यक्षमता के अंतर का अध्ययन करना हो। अथवा कक्षा VIII के 10 बालक और बालिकाओं के गणित के प्राप्तांकों का अध्ययन करना हो। जिससे यह समझा जा सके कि दोनों समूहों के बीच गणित के उपलब्धियों में अंतर है या नहीं। साथ ही यह भी स्पष्ट हो सके कि अंतर वास्तविक या सार्थक है या फिर यह अंतर संयोगवश पाया जा रहा है। इस प्रकार के अवस्था में प्रसरण-विश्लेषण प्रविधि का उपयोग किया जाता है। इससे कम समय और श्रम से व्यवस्थित (Systematic) निष्कर्ष प्राप्त किया जा सकता है। इससे परीक्षण का नतीजा सारांश (Summary) के रूप में पाया जाता है। प्राप्त एफ परीक्षण के निष्कर्ष को सार्थकता स्तर (Level of significance) के टेबल (टेबल संख्या 1) के अध्ययन से स्पष्ट हो जाता है कि अंतर सार्थक है या संयोगवश। इससे यह स्पष्ट पता चल जाता है कि शून्य परिकल्पना (Null Hypothesis) स्वीकृत या अस्वीकृत हो गया है। शून्य चरिकल्पना में कल्पना किया जाता है कि दो समूहों बालक-बालिकाओं के उपलब्धियों में कोई अंतर नहीं होना चाहिए इस अनुमान को मान कर परीक्षण किया जाता है। अर्थात् यौन भिन्नता के कारण दोनों चाहिए। इस प्रकार के पूर्व परिकल्पना को शून्य परिकल्पना (Null Hypothesis) कहा जाता है। जब दिए गए समस्या (Problems) का एफ मान (F Value) निकाला जाता है। तब एक वितरण टेबल (F Distribution Total) के आधार पर उसकी सार्थकता जाँची जाती है। अगर एफ सार्थक (Signifibution) पाया जाता है। तब इसका अर्थ है कि दोनों समूह बालक-बालिकाओं के उपलब्धियों में अंतर है। ऐसे में शून्य परिकल्पना अस्वीकृत (Reject) हो जाती है। पर अगर दो समूहों में सार्थक कोई अंतर नहीं पाया जाता है। तब यहाँ शून्य परिकल्पना (Null Hypothesis) स्वीकृत (Sustained) माना जाएगा।

11-1-3 दो से अधिक समूहों या मध्यमानों का प्रसरण विश्लेषण (Anova in case of more than two groups or means) :

सरल प्रसरण-विश्लेषण का एक अन्य प्रकार है दो या अधिक समूहों या मध्यमनों के बीच के प्रसरण का विश्लेषण करना। उदाहरण के लिए-किसी व्यवसाय द्वारा दिए गए विज्ञापन के प्रभाव का अध्ययन विभिन्न समूहों पर करना हो। ऐसे में इस प्रविधि की सहायता ली जा सकती है। किसी व्यवसाय के विज्ञापन का प्रभाव विभिन्न सामाजिक-आर्थिक समूहों पर जानना है जैसे उच्च मध्य, निम्न सामाजिक आर्थिक लोगों के समूहों (High, Middle, Low SES Groups) पर जाँचना हो और देखना हो कि विज्ञापन का प्रभाव वास्तविक है या संयोगवश अंतर है।

तब इस प्रविधि का प्रयोग किया जा सकता है। इसी प्रकार विभिन्न सामाजिक आर्थिक समूह के बच्चों के गणितय क्षमता जाँचनी हो और निष्कर्ष निकालनता हो कि क्या सामाजिक आर्थिक स्थिति (Socio-Economic Status or SES) का प्रभाव बच्चों के गणितय उपलब्धि पर पड़ता है। यह यह प्रभाव संयोगवश है या वास्तविक या सार्थक है। ऐसे स्थिति में इस प्रविधि को प्रयोग में लाया जा सकता है। ऐसे में उच्च, मध्यम, निम्न, अति निम्न जैसे आय वर्ग के बच्चों को चुन कर उनके गणित के अंक के मध्यमान निकाला जाता है। फिर इस प्रविधि के मदद से विभिन्न समूहों के आन्तरिक प्रसरण और विभिन्न समूहों के मध्य के प्रसरण (Between Group variance and within the group variance) की सहायता से एफ-परीक्षण (F-Test) निकाला जाता है। एफ (F) निकालने के बाद टेबुल की सहायता से इसकी सार्थकता जाँची जाती है। जिससे स्पष्ट होता है कि विभिन्न समूहों के मध्य का अंतर सार्थक है या संयोगवश (by chance) इसके अतिरिक्त यह भी स्पष्ट होता है तब कि हम समझ सकते हैं कि सामाजिक आर्थिक स्थिति (Socio Economic Status SES) का प्रभाव बच्चों के गणित के उपलब्धि पर पड़ता है। यह सार्थकता किस स्तर का है यह भी टेबुल की मदद से जाना जा सकता है। दूसरे स्थिति में पाया जा सकता है कि विभिन्न समूहों के मध्य अंतर सार्थक नहीं वरन संयोगवश है। ऐसे में यह स्पष्ट होता है कि बच्चों के गणित के उपलब्धि पर उनके सामाजिक आर्थिक स्थिति (Socio-Economic Status or SES) का प्रभाव नहीं पड़ता है। ऐसे में इस परिकल्पना को बल (Support) मिलता है कि सामाजिक-आर्थिक स्थिति (SES) का प्रभाव बच्चों के गणित उपलब्धि (Null Hypothesis) के स्वीकृत किया जाता है कि बच्चों के उपलब्धियों पर सामाजिक आर्थिक स्थिति (SES) का शून्य प्रभाव पड़ता है।

इस प्रकार की समस्या और उसका हल आगे दिया है।

Problem: Susan predicts that students will learn most effectively with a constant background music, as opposed to an unpredictable sound or no sound at all. She randomly divides twenty-four students into three groups of eight. All students study a passage of text for 30 minutes. Those in group 1 study with background sound at a constant volume in the background. Those in group 2 study with noise that changes volume periodically. Those in group 3 study with no sound at all. After studying, all students take a 10 point multiple choice test over the material. Their scores follow:

group	test Scores							
1) Constant Sound	7	4	6	8	6	6	2	9
2) Random Sound	5	5	3	4	4	7	2	2
3) No Sound	2	4	7	1	2	1	5	5

गाना निश्चित शोर

कम ज्यादाशोर

बिना शोर

x_1	x_1^2	x_2	x_2^2	x_3	x_3^2
7	49	5	25	2	4
4	16	5	25	4	16
6	36	3	9	7	49
8	64	4	16	1	1
6	36	4	16	2	4
6	36	7	49	1	1
2	4	2	4	5	25
9	81	2	4	5	25

समस्या-सुजने ने बच्चों के शिक्षण के दौरान गाना सुनने के प्रभाव का अध्ययन करना चाहा। उसने यादृच्छिक चयन (Random Selection) के द्वारा 24 बच्चों के तीन समूह बनाए। प्रत्येक समूह में 8-8 बच्चे रखे गए। तीनों समूहों को 30 मिनट अध्ययन करने कहा गया। पहले समूह को निश्चित शोर (Volume) में गाने के आवाज के साथ अध्ययन करने दिया गया। दूसरे समूह के अध्ययन के दौरान शोर कम ज्यादा किया जाता रहा और तीसरे समूह को बिना शोर के अध्ययन करने दिया गया। इसके बाद सभी बच्चों का 10 प्वाई टेस्ट लिया गया। उससे उपरोक्त आँकड़े (Data) प्राप्त हुए।

$$\begin{array}{llllll} \Sigma x_1 = 48 & \Sigma x_1^2 = 322 & \Sigma x_2 = 32 & x_2^2 = 148 & x_3 = 28 & \Sigma x_3^2 = 125 \\ (\Sigma x_1)^2 = 2304 & & (\Sigma x_2)^2 = 1024 & & (\Sigma x_3)^2 = 576 & \\ M_1 = 6 & & M_2 = 4 & & M_3 = 3.375 & \end{array}$$

$$\text{Total SS} = (322 + 148 + 125) - \frac{(48 + 32 + 27)^2}{24}$$

$$= 595 - 477.04$$

$$\text{SS Total} = 117.96$$

$$\text{SS among} = \left[\frac{2304}{8} + \frac{1024}{8} + \frac{729}{8} \right] - 477.04$$

$$= 507.13 - 477.04$$

$$\text{SS among} = 30.08$$

$$\text{SS Within} = 117.96 - 30.08 = 87.88$$

Source	SS	df	MS	F
Among	30.08	2	15.04	3.59
Within	87.88	21	4.18	

*(according to the F sig/probability table with df = (2,21) F must be at least 3.4668 to reach p < .05, so F score is statistically significant)

Formula

$$\text{SSTotal} (\Sigma x_1^2 + \Sigma x_2^2 + \dots \Sigma x_3^2) - \frac{(\Sigma x_1 + \Sigma x_2 + \dots \Sigma x_3)^2}{N}$$

$$\text{SSTotal} = \left[\frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} \right] - \frac{(\Sigma x_1 + \Sigma x_2 + \dots \Sigma x_3)^2}{N}$$

$$\text{SS within} = \text{SS total} - \text{SS among}$$

$$\text{df among} = r - 1$$

$$\text{df within} = N - r$$

$$\text{MS among} = \frac{\text{SS}_{\text{among}}}{\text{df}_{\text{among}}}$$

$$\text{MS within} = \frac{\text{SS}_{\text{within}}}{\text{df}_{\text{within}}}$$

$$F = \frac{\text{MS}_{\text{among}}}{\text{MS}_{\text{within}}}$$

x = individual observation

r = number of groups

N = total number of observation (all groups)

n = number of observation in group

Steps (assuming three groups)

छ: कॉलम बनाए-

Create si columns: " x_1 ", " x_1^2 ", " x_2 ", " x_2^2 ", and " x_3^2 "

1. Put the raw data, according to group, in " x_1 ", " x_2 ", and " x_3 "

प्राप्तकों को x_1, x_2, x_3 समूहों में लिखें।

2. Calculate the sum for group 1.

सभी को अलग-अलग समूह में जोड़े

3. Calculate the $(\sum x)^2$ for group 1.

सभी कॉलम का वर्ग निकालें x_1^2, x_2^2, x_3^2

4. Calculate $\sum x^2$ for group 1.

समूह 1 का मध्यमान निकाले

5. Calculate $\sum x^2$ for group 1.

समूह 1 का $\sum x_1^2$ निकालें।

6. Repeat steps 2-5 for groups 2 and 3

समूह 2 और 3 का मध्यमान तथा $\sum x_2^2, \sum x_3^2$

7. Set up SS total SS among formulas and calculate

7. सूत्रों का प्रयोग करें

8. Calculate SS within

SSwithin का गणना करें।

9. Enter soms of squares into the ANOVA table, and complete the table by calculating: df among, dfwithin, MSamong, and MSwithin, and F

अनोवा टेबल बनाएँ

10. Check to see if F is statistically significant on probility table with appropriate defree of freedom and $p < .05$.

F के मान का df एनोवा टेबल में देखें। (टेबुल-1)

Interpretation: Susan can conclude that her hypothesis may be supported. The means are as the predicted, in that the constant music group has the highest score. However, the significant F only indicates that at least two means are significantly different from one another, but she can't know which specific mean pairs significantly differ until she conducts a post-hoc anlysis.

निष्कर्ष-सुजेन के पूर्वकल्पना F (F Ratio) के आधार पर स्वीकृत हो गई। उन ने अनुमान लगाया था कि गानों सुनने के साथ बच्चे अध्ययन अच्छा कर सकते हैं। इसकी पुष्टि हुई। गाना या सुनने के दौरान पढ़ने वाले समूह ने सर्वाधिक अंक प्राप्त किए। परन्तु कौन अध्ययन समूह वास्तव में अलग है। यह जानने के लिए अन्य प्रविधियों का प्रयोग आवश्यक है।

उपरोक्त समस्या में प्रस्तुत हुए गणीत के चरणों को पुनः निम्नलिखित प्रकार से स्पष्ट किया जा रहा है-

Culcation (one way):/simple analysis of variace.

Computation by hand of a One-way ANOVA is tedious; computation of a two-way ANOVA is nearly impossible. Here is a general overview of how to do it:

सरल प्रसरण-विश्लेषण/एकमात्री प्रसरण विश्लेषण गणना की प्रक्रिया कठिन और जटिल है। नीचे दिए गए चरणों के अनुसार इसकी गणना की जा सकती है।

Set up data table to help keep track of the data

Col 1	Col 2	Col 3	Col 4	Col 5	Col 6
Group A data	Group B data	Group C data	A ²	B ²	C ²
Sum A	Sum B	Sum C	(sum A) ²	(sum B) ²	(sum C) ²
Mean A	Mean B	Mean C			

$$SS_T = \sum X^2 - \frac{(\sum X_T)^2}{N}$$

$$SS_b = \sum \frac{(\sum X)^2}{N} - \frac{(\sum X_T)^2}{N}$$

$$SS_w = SS_T - SS_b$$

$$df_b = (\text{number of groups} - 1)$$

$$df_T = (\text{number of subjects} - 1)$$

$$df_w = df_T - df_b$$

$$MS_b = \frac{SS_b}{df_b}$$

$$MS_w = \frac{SS_w}{df_w}$$

$$F_w = \frac{MS_b}{MS_w}$$

1. Compute the total sum of squares (SST) SST निकालें।
2. Compute the between groups sum of squares (SSB) SSB निकालें।
3. Compute the within groups sum of squares (SSW) SSW निकालें।
4. Calculate the degree of freedom for between groups: Df= number of groups - 1 = (N) Df निकाले (समूहों के बीच)
5. Calculate the degrees of freedom for the total: Df= total participants - 1 = (D-1) = O Df निकाले (पूरे-समूह के लिए)
6. Calculate the degree of freedom for the within groups: Df= (df total - df between groups) df निकाले (समूह के अंदर)

7. Compute the mean squares for each (MSB, MSW) MSB, MSW का वर्ग निकालें।
8. Compute the value for F by dividing the mean square for between groups by the mean square for within groups इसके लिए MSM/MSW सूत्र का प्रयोग करें।
9. And finally: Determine whether F is significant: प्राप्त F की सार्थकता जाँच करें।
 - o The columns on the critical values table are labeled "degrees of freedom associated with the between groups degrees of freedom," which, in this case, would be the N variable.
 - o The rows are labeled "degrees of freedom associated with the within groups degrees of freedom."
 - o Find the point at which these two values interest. This is the critical value of F. If your value is greater than the critical value, reject the null hypothesis. If it is not, do not reject the null hypothesis

सरल एनोवा की सीमायें (Limitation of Simple Anova)

इस प्रविधि में जब तीन या अधिक समूहों का अध्ययन किया जाता है। तब सार्थक अन्तर मिलने पर यह समझना है। इसे सुजेन के प्रयोग की सहायता से समझा जा सकता है। सुजेन ने तीन समूहों पर अध्ययन किया। उसने अध्ययन के आरम्भ में पूर्वकल्पना की कि गाना सुनकर अध्ययन करना बच्चों के लिए लाभदायक है। प्राप्त एफ (F) द्वारा उसके पूर्वकल्पना की पुष्टि भी हो गई। पर यह स्पष्ट नहीं हो सका कि कौन सा अन्तर सार्थक है—समूह A और B समूह A और C या समूह B और C के मध्य। इसके लिए अन्य उपयुक्त सांख्यिकी प्रविधियों की आवश्यकता पड़ती है। जिसे पोस्ट होक (Post hoc) प्रविधि कहा जाता है।

एनोवा एक जटिल सांख्यिकी प्रविधि है। यह शोध और अनुसंधानों के लिए महत्वपूर्ण प्रविधि है। यह परिकल्पनाओं के जाँच में अहम भूमिका निभाती है। शून्य परिकल्पना (Null Hypothesis) के परिक्षण के लिए सर्वाधिक उपयुक्त प्रविधि है। इसकी जटिलता के सरल करने के लिए एनोवा हल करने के लिए अनेक कम्प्यूटर साफ्टवेयर (Software) भी बनाए गए हैं। इसे हल करने और एनोवा को ग्राफ रूप में प्रस्तुत करने के लिए SPSS साफ्टवेयर (SPSS Software) सहायक है।

A Hand Calculation Problem

Suppose a teacher is interested in determining whether there is a difference in reading achievement results using three different reading approaches. She randomly assigns 5 children to each of the three reading programs. We use our six-step approach.

एक शिक्षिका बच्चों के पाठन का अध्ययन करना चाहती है। वह पाठन (Reading) के विभिन्न प्रकारों (approaches) का प्रभाव जाँचना चाहती है। इसके लिए उसने बच्चों के तीन समूह बनाए। प्रत्येक समूह में 5 बच्चे रखे गए। सभी समूहों के प्राप्तांक निम्नलिखित हैं। इसे नीचे दिये गए छः चरणों में हल करें।

Step 1

$$H_0: u_1 = u_2 = u_3 = u_4 \quad H_1: \text{not } H_0$$

परिकल्पना Step 1 w- पाठक के विभिन्न प्रकार शून्य परिकल्पना (HO): $u_1 = u_2 = u_3 = u_4$ परिकल्पना (H1): शून्य परिकल्पना अस्वीकृत होगा।

Step 2

$$\alpha = .05$$

$$df_{BG} = (nT - k) = 12$$

$$df_{WG} = (nT - k) = 12$$

$$F_{cv} = 3.88$$

परिकल्पना Step 2 सार्थकता स्तर Alpha (Significance level = .05) df_{BG} = समूहों के अंदर का $df = (3-1) = 2$ df_{WG} = समूह के बीच का $df = (15-3) = 12$

Step 3

Group 1 48 50 53 52 50

Group 2 51 50 52 50 51

Group 3 61 62 60 63 60

Step 4

Try to calculate the F value following these steps. It is important that you are able to do these calculations by hand as problems similar to this

निम्नलिखित चरणों के अनुसार गणना कर एफ का मान (Value of F) ज्ञात करें।

- Calculate SXT and SXT²
- Calculate SSTOT
- Calculate SSBG
- Calculate SSWG
- Calculate both mean squares
- Calculate the F statistic

Step 5

Reject H₀

Step 6

The Means of the three groups are significantly different. There are differences in reading depending upon which reading treatment a child was assigned ($F(2,12) = 88.935, p < .05$).

11.5 मॉडल प्रश्न (Model Questions) :

1. प्रसरण-विश्लेषण का परिचय देते हुए परिभाषित करें।

Explain and define Analysis of Variance.

2. प्रसरण-विश्लेषण प्राचलिक या अप्राचलिक सांख्यिकी है? क्यों प्रसरण विश्लेषण की अभिधारणायें और उपयोग का वर्णन करें।

Analysis of Variance is a Parametric or non parametric statistics? Why? Explain the assumptions and applications of ANOVA.

3. सरण प्रसरण विश्लेषण का वर्णन करें। इसका प्रयोग किन-किन अवस्थाओं में किया जा सकता है? वर्णन करें।

Explain simple Analysis of Variance. In what conditions it may be used? Explain.