

अर्थशास्त्र

(सांख्यिकी)

अध्याय-2: आँकड़ों का संगठन



आँकड़ों का संगठन

वर्गीकरण आँकड़ों को उनकी समानता और निकट सम्बन्ध के अनुसार वर्गों और समूहों में व्यवस्थित करने की एक प्रक्रिया है। बड़ी संख्या में आंकड़े अपने मूल रूप में शुद्ध आंकड़े कहलाते हैं। चर और गुण : जब आँकड़ों का वर्गीकरण समय या आकार की मात्रा के रूप में किया जा सकता है तो उसे चर कहते हैं।

आँकड़ों के वर्गीकरण

एकत्रित आँकड़ों को उनकी समानता और असमानताओं के आधार पर विभिन्न वर्गों व समूहों में विभाजित करना वर्गीकरण कहलाता है।

वर्गीकरण आँकड़ों को उनकी समानता और निकट सम्बन्ध के अनुसार वर्गों और समूहों में व्यवस्थित करने की एक प्रक्रिया है। बड़ी संख्या में आंकड़े अपने मूल रूप में शुद्ध आंकड़े कहलाते हैं। चर और गुण : जब आँकड़ों का वर्गीकरण समय या आकार की मात्रा के रूप में किया जा सकता है तो उसे चर कहते हैं।

वर्गीकरण की विशेषताएँ-

- **स्पष्टता (Clarity):** वर्गीकरण करते समय विभिन्न वर्ग इस प्रकार निर्धारित किये जाने चाहिए कि उनमें सरलता व स्पष्टता हो
- **व्यापकता (Comperhensiveness):** किसी भी समस्या से सम्बन्धित आँकड़ों का वर्गीकरण इतना व्यापक होना चाहिए।
- **सजातीयता (Homogeneity):** वर्गीकरण करते समय इस बात का ध्यान रखना चाहिए कि प्रत्येक वर्ग विशेष में रखे जाने वाले पदों के गुण एक समान ही हों। यानि कि भिन्न-भिन्न वर्गों के गुण एक दूसरे से भिन्न हों।
- अनुकूलता
- **लोचदार (Elastic):** वर्गीकरण लोचदार होना चाहिए। उसमें उद्देश्यों की आवश्यकता के अनुसार विभिन्न वर्गों में परिवर्तन करने की संभावना होनी चाहिए।

- **स्थिरता (Stability):** आँकड़ों को तुलना योग्य बनाने तथा परिणामों की अर्थपूर्ण तुलना करने के लिए आवश्यक है, स्थिरता हो।
- **उपयुक्तता (Suitability):** वर्गों की रचना उद्देश्यानुसार होनी चाहिए। जैसे-व्यक्तियों की आर्थिक स्थिति या बचत प्रवृत्ति जानने के लिए आय के आधार पर वर्गों की रचना करना उपयुक्त रहेगा।

वर्गीकरण का आधार-

कालानुक्रमिक वर्गीकरण:- जब आँकड़ों को समय के संदर्भ जैसे- वर्ष, तिमाही, मासिक या साप्ताहिक आदि रूप में आरोही या अवरोही क्रम में वर्गीकृत किया जा सकता है।

स्थानिक वर्गीकरण:- जब आँकड़ों को भौगोलिक स्थितियों जैसे देश, राज्य, शहर, जिला, कस्बा आदि में वर्गीकृत किया जाता है।

गुणात्मक वर्गीकरण:- विशेषताओं पर आधारित आँकड़ों के वर्गीकरण को गुणात्मक वर्गीकरण कहा जाता है। जैसे राष्ट्रीयता, साक्षरता, लिंग, वैवाहिक स्थिति आदि।

मात्रात्मक वर्गीकरण:- जब विशेषताओं की प्रकृति मात्रात्मक होती है। जैसे ऊँचाई, भ्रम, आयु, आय, छात्रों के अंक आदि।

चर

चर से अभिप्राय किसी तथ्य की वह विशेषता है जिसमें परिवर्तन होते रहते हैं तथा जिन्हें किसी इकाई द्वारा मापा जा सकता है।

चर की प्रकार

संतत चर- वे चर हैं जो मापदण्डों की इकाइयों में होते हैं और अपने वर्गों में विभक्त किए जा सकते हैं। इन्हें भिन्नात्मक रूप में लिखा जा सकता है।

विविक्त चर- ये चर केवल निश्चित मान वाले हो सकते हैं। इसके मान केवल परिमित 'उछाल से बदलते हैं। यह उछाल एक मान से दूसरे मान के बीच होते हैं, परन्तु इसके बीच में कोई मान नहीं आता है।

बारम्बारता वितरण- यह अपरिष्कृत आँकड़ों को एक मात्रात्मक चर में वर्गीकृत करने का एक सामान्य तरीका है। यह दर्शाता है कि किसी चर के भिन्न मान विभिन्न वर्गों में अपने अनुरूप वर्गों में बारम्बारताओं के साथ कैसे वितरित किए जाते हैं।

- **वर्ग-** निश्चित सीमाओं के विस्तार को जिसमें मर्दें शामिल होती हैं, वर्ग कहा जाता है जैसे- 0-10, 10-20, 20-30 आदि।
- **वर्ग सीमाएँ-** प्रत्येक वर्ग की दो सीमाएँ होती हैं - निम्न सीमा तथा ऊपरी सीमा। उदाहरण के लिए 10-20 के वर्ग में 10 निम्न सीमा (L_1) तथा ऊपरी सीमा (L_2) है।
- **वर्गान्तर-** वर्ग की ऊपरी सीमा तथा निम्न सीमा के अन्तर को वर्गान्तर कहते हैं। उदाहरण के लिए 10-20 का वर्गान्तर 10 है।
- **वर्ग आवृत्ति-** किसी वर्ग में शामिल मर्दों की संख्या को उस वर्ग की आवृत्ति या बारम्बारता कहते हैं। इसे f द्वारा प्रदर्शित करते हैं।

मध्य मूल्य- किसी वर्ग के वर्गान्तर का मध्य बिन्दु ही मध्य मूल्य कहलाता है। इसे वर्ग की ऊपरी सीमा व निम्न सीमा के योग को 2 से भाग देकर प्राप्त किया जा सकता है। इसे वर्ग चिन्ह भी कहते हैं।

अपवर्जी विधि- इस पद्धति का उपयोग उन श्रृंखलाओं के लिए किया जाता है जिनमें एक वर्ग की ऊपरी सीमा अगले वर्ग की निचली सीमा बन जाती है। इसे अपवर्जी श्रेणी कहते हैं क्योंकि किसी वर्ग अंतराल की ऊपरी सीमा की बारम्बारताएं उस वर्ग विशेष में शामिल नहीं होती हैं। इस प्रकार की श्रृंखला में, एक वर्ग की ऊपरी सीमा अगले वर्ग की निचली सीमा बन जाती है, उदाहरण के लिए, 0-10, 10-20, 20-30, इत्यादि। ऊपरी सीमा को बाहर रखा गया है लेकिन निचली सीमा को वर्ग अंतराल में शामिल किया गया है। निरंतर चरों के डेटा के लिए यह विधि सबसे उपयुक्त है।

समावेशी विधि- डेटा के वर्गीकरण की इस पद्धति के तहत, वर्गों का निर्माण इस तरह से किया जाता है कि एक वर्ग अंतराल की ऊपरी सीमा अगले वर्ग अंतराल की निचली सीमा के रूप में खुद

को दोहराती नहीं है। ऐसी श्रृंखला में, ऊपरी सीमा और निचली सीमा दोनों को विशेष वर्ग अंतराल में शामिल किया जाता है, उदाहरण के लिए, 1-5, 6-10, 11-15, और इसी तरह। अंतराल 1-5 में दोनों सीमाएँ अर्थात् 1 और 5 शामिल हैं।

सूचना की हानि- बारम्बारता वितरण के रूप में आँकड़ों के वर्गीकरण में एक अन्तर्निहित दोष पाया जाता है। यह परिष्कृत आँकड़ों को सारांश में प्रस्तुत कर उन्हें संक्षिप्त एवं बोधगम्य तो बनाता है, परन्तु इसमें वे विस्तृत विवरण प्रकट नहीं हो पाते जो अपरिष्कृत आँकड़ों में पाए जाते हैं। अतः अपरिष्कृत आँकड़ों को वर्गीकृत करने में सूचना की हानि होती है।

श्रृंखला को प्रकार

1. **व्यक्तिगत श्रृंखला-** वह श्रृंखला है जिसमें प्रत्येक इकाई का अलग-अलग माप प्रकट किया जाता है, जो कि निम्न उदाहरण से स्पष्ट है-

अनुक्रमांक	अंक
1	18
2	95
3	82
4	59
5	92

2. **खण्डित या विविक्त श्रृंखला-** वह श्रृंखला है जिसमें आँकड़ों को इस प्रकार प्रस्तुत किया जाता है कि प्रत्येक मद का निश्चित माप स्पष्ट हो जाता है, जो कि निम्न उदाहरण से स्पष्ट है-

परिवार का आकार	परिवारों की संख्या
1	15
2	10
3	20
4	30
5	15
6	10

3. सतत या अविच्छिन्न श्रृंखला- वह श्रृंखला है जिसमें इकाईयों का निश्चित माप संभव नहीं होता इसलिए इन्हें वर्ग सीमाओं में प्रकट किया जाता है जो कि निम्न उदाहरण से स्पष्ट है-

प्राप्तांक	आवृत्ति
0-10	5
10-20	7
20-30	10
30-40	8

स्मरणीय बिन्दु-

अपरिष्कृत आँकड़ों को सरल, संक्षिप्त तथा व्यवस्थित ढंग से प्रस्तुत करने को आँकड़ों का व्यवस्थितिकरण कहा जाता है ताकि उन्हें आसानी से आगे के सांख्यिकीय विश्लेषण के योग्य बनाया जा सके।

NCERT SOLUTIONS

प्रश्न (पृष्ठ संख्या 37 - 38)

प्रश्न 1 निम्नलिखित में से कौन-सा विकल्प सही है-

- एक वर्ग मध्यबिन्दु बराबर है-
 - a) उच्च वर्ग सीमा तथा निम्न वर्ग सीमा के औसत के।
 - b) उच्च वर्ग सीमा तथा निम्न वर्ग सीमा के गुणनफल के।
 - c) उच्च वर्ग सीमा तथा निम्न वर्ग सीमा के अनुपात के।
 - d) उपर्युक्त में से कोई नहीं।
- दो चरों के बारम्बारता वितरण को इस नाम से जानते हैं-
 - a) एकविचर वितरण
 - b) द्विचर वितरण
 - c) बहुचर वितरण
 - d) इनमें से कोई नहीं।
- वर्गीकृत आँकड़ों में सांख्यिकीय परिकलन आधारित होता है-
 - a) प्रेक्षणों के वास्तविक मानों पर
 - b) उच्च वर्ग सीमाओं पर
 - c) निम्ने वर्ग सीमाओं पर
 - d) वर्ग के मध्य बिन्दुओं पर।
- अपवर्जी विधि के अन्तर्गत-
 - a) किसी वर्ग की उच्च वर्ग सीमा को वर्ग अन्तराल में समावेशित नहीं करते।
 - b) किसी वर्ग की उच्च वर्ग सीमा को वर्ग अन्तराल में समायोजित करते हैं।
 - c) किसी वर्ग की निम्न वर्ग सीमा को वर्ग अन्तराल में समावेशित नहीं करते हैं।
 - d) किसी वर्ग की निम्न वर्ग सीमा को वर्ग अन्तराल में समावेशित करते हैं।
- परास का अर्थ है-
 - a) अधिकतम एवं न्यूनतम प्रेक्षणों के बीच अन्तर

- b) न्यूनतम एवं अधिकतम प्रेक्षणों के बीच अन्तर
 c) अधिकतम एवं न्यूनतम प्रेक्षणों को औसत
 a) अधिकतम एवं न्यूनतम प्रेक्षणों का अनुपात।

उत्तर –

1. उच्च वर्ग सीमा तथा निम्न वर्ग सीमा के औसत के।
2. द्विचर वितरण।
3. वर्ग के मध्य बिन्दुओं पर।
4. किसी वर्ग की उच्च वर्ग सीमा को वर्ग अन्तराल में समावेशित नहीं करते।
5. अधिकतम एवं न्यूनतम प्रेक्षणों के बीच अन्तर।

प्रश्न 2 वस्तुओं को वर्गीकृत करने में क्या कोई लाभ हो सकता है? अपनी दैनिक जीवन से एक उदाहरण देकर व्याख्या कीजिए।

उत्तर – हाँ वस्तुओं को वर्गीकृत करने का बहुत लाभ है-

- यह अपरिष्कृत आँकड़ों को सांख्यिकीय विश्लेषण के लिए एक सही रूप में संक्षिप्त करता है।
- यह जटिलताओं को दूर करता है तथा आँकड़ों की विशेषताओं को उजागर करता है।
- यह तुलना करने तथा निष्कर्ष निकालने में सहायता करता है। उदाहरण के लिए यदि एक विश्वविद्यालय के विद्यार्थियों को उनके विषय तथा लिंग के आधार पर वर्गीकृत किया जाए तो तुलना करना अति सरल होगा।
- यह दिए गए आँकड़ों के तत्वों के अंतर संबंध के बारे में जानकारी प्रदान करता है। उदाहरण के लिए साक्षरता तथा अपराध दरों के आँकड़ों से हम यह सहसंबंध स्थापित कर सकते हैं कि क्या ये एक दूसरे से संबंधित हैं।
- यह समान तत्वों को एक समान करके आँकड़ों को समरूप समूहों में परिवर्तित करता है तथा उनमें समान व असमानताएँ ज्ञात करता है।

प्रश्न 3 चर क्या है? एक संतत तथा विविक्त चर के बीच भेद कीजिए।

उत्तर –

- **चर-** वे मूल्य जिनका मान एक मद से दूसरे मद में बदलता रहता है और जो संख्यात्मक रूप में मापे जा सकते हैं उन्हें चर कहते हैं।
- **संतत तथा विविक्त चर में भेद-** संतत चर का कोई भी संख्यात्मक मान हो सकता है, जैसे

$$1, 2, \frac{1}{2}, \frac{3}{4}, \sqrt{2}, 1.732$$

आदि। जबकि विविक्त चर केवल निश्चित मान वाले हो सकते हैं।

- **जैसे-** छात्रों की संख्या, परिवार के सदस्यों की संख्या।

प्रश्न 4 आँकड़ों के वर्गीकरण में प्रयुक्त अपवर्जी तथा समावेशी विधियों की व्याख्या कीजिए।

उत्तर – आँकड़ों को संतत श्रृंखला में वर्गीकृत करने की दो विधियाँ हैं:

- अपवर्जी श्रृंखला
- समावेशी श्रृंखला

अपवर्जी श्रृंखला- इस विधि में एक वर्ग की निचली सीमा अगले वर्ग की ऊपरी सीमा होती है। इसमें ऊपरी सीमा वर्ग अन्तराल में शामिल नहीं होती।

उदाहरण के लिए-

X	F
0-10	10
10-20	5
20-30	4
30-40	8
40-50	6

समावेशी श्रृंखला- इस विधि में एक वर्ग की निचली सीमा अगले वर्ग की ऊपरी सीमा नहीं होती। इसमें निम्न तथा उच्च दोनों सीमाएँ वर्ग अंतराल में शामिल होती हैं।

उदाहरण के लिए-

अंक (X)	विद्यार्थियों की संख्या
0-9	10
10-19	5
20-29	4
30-39	8
40-49	9

प्रश्न 5 सारणी के आँकड़ों का प्रयोग करें, जो 50 परिवारों के भोजन पर मासिक व्यय (रु. में) को दिखलाती है-

1. भोजन पर मासिक परिवारिक व्यय का प्रसार ज्ञात कीजिए।
2. परास को वर्ग अंतराल की उचित संख्याओं में विभाजित करें तथा व्यय का बारंबारता वितरण प्राप्त करें।

उन परिवारों की संख्या पता कीजिए जिनका भोजन पर मासिक व्यय-

1. 2000/- रु. से कम है।
2. 3000/- रु. में अधिक है।
3. 1500/-रु. और 2500 रु के बीच है।

उत्तर -

$$\text{प्रसाल} = \text{अधिकतम मान} - \text{न्यूनतम मान} = 5090 - 1007 = 4082$$

क्रम.	वर्ग अंतराल	बारम्बारता	'से अधिक'	संचयी बारम्बारता
(i)	100-2000	33	50	
(ii)	2000-3000	11	17	(i) 33
(iii)	3000-4000	3	6	(ii) 6
(iv)	4000-5000	2	3	(iii) 19

(v)	5000-6000	1	1	
-----	-----------	---	---	--

प्रश्न 6 एक शहर में, यह जानने हेतु 45 परिवारों का सर्वेक्षण किया गया कि वे अपने घरों में कितनी संख्या में सेल फोनों का इस्तेमाल करते हैं। नीचे दिए गए उनके उत्तरों के आधार पर एक बारंबारता सारणी तैयार कीजिए।

1	3	2	2	2	2	1	2	1	2	2	3	3	3	3
3	3	2	3	2	2	6	1	6	2	1	5	1	5	3
2	4	2	7	4	2	4	3	4	2	0	3	1	4	3

उत्तर -

सेल फोन की संख्या	मिलान चिन्ह	बारंबारता
0		1
1		7
2		15
3		12
4		5
5		2
6		2
7		1

प्रश्न 7 वर्गीकृत आँकड़ों में 'सूचना की क्षति' का क्या अर्थ है?

उत्तर - बारम्बारता वितरण के रूप में आँकड़ों के वर्गीकरण में एक अन्तर्निहित दोष पाया जाता है। यह अपरिष्कृत आँकड़ों का सारांश प्रस्तुत कर उन्हें संक्षिप्त एवं बोधगम्य तो बनाता है, परन्तु इसमें वे विस्तृत विवरण नहीं हो पाते जो अपेक्षाकृत आँकड़ों में पाए जाते हैं। आवृत्ति वितरण के द्वारा आँकड़े संक्षिप्त हो जाते हैं वर्गीकृत होने से आँकड़ों से सूचना की क्षति होती है। एक बार आँकड़ों का वर्गीकरण हो जाने पर व्यक्तिगत आँकड़ों का अस्तित्व खत्म हो जाता है सांख्यिकीय गणनाएँ वास्तविक मूल्य पर आधारित नहीं होती हैं।

प्रश्न 8 क्या आप इस बात से सहमत हैं कि अपरिष्कृत आँकड़ों की अपेक्षा वर्गीकृत आँकड़े बेहतर होते हैं?

उत्तर – हाँ, हम इस बात से सहमत हैं कि अपरिष्कृत आँकड़ों की अपेक्षा वर्गीकृत आँकड़े बेहतर होते हैं। यह अपरिष्कृत आँकड़ों को सांख्यिकीय विश्लेषण के लिए एक सही रूप में संक्षिप्त करता है। यह जटिलताओं को दूर करता है तथा आँकड़ों की विशेषताओं को उजागर करता है। यह तुलना करने तथा निष्कर्ष निकालने में सहायता करता है। यह दिए गए आँकड़ों के तत्वों के अंतरसंबंध के बारे में जानकारी प्रदान करता है। यह समान तत्वों को एक समान करके आँकड़ों को समरूप समूहों में परिवर्तित करता है तथा उनमें समान व अमानताएँ ज्ञात करता है।

प्रश्न 9 एक-विचर एवं द्विचर बारम्बारता वितरण के बीच अन्तर बताइए।

उत्तर – एकल चर के बारम्बारता वितरण को एकविचर वितरण कहा जाता है जैसे किसी छात्र के प्राप्तांक एकल चर के एकविचर वितरण को प्रदर्शित करते हैं। जबकि एक द्विचर बारम्बारता वितरण दो चरों का बारम्बारता वितरण है जैसे किसी क्षेत्र में रहने वाले लोगों का वजन।

प्रश्न 10 निम्नलिखित आँकड़ों के आधार पर 7 का वर्ग अंतराल लेकर समावेशी विधि द्वारा एक बारम्बारता वितरण तैयार कीजिए।

28	17	15	22	29	21	23	27	18	12	7	2	9	4
1	8	3	10	5	20	16	12	8	4	33	27	21	15
3	36	27	18	9	2	4	6	32	31	29	18	14	13
15	11	9	7	1	5	37	32	28	26	24	20	19	25
19	20	6	9										

उत्तर –

वर्ग अंतराल	मिलान चिन्ह	बारंबारता
0-7		15
8-15		15
16-21		12
22-28		10
29-35		6
36-42		2
कुल		60

SHIVOM CLASSES
8696608541