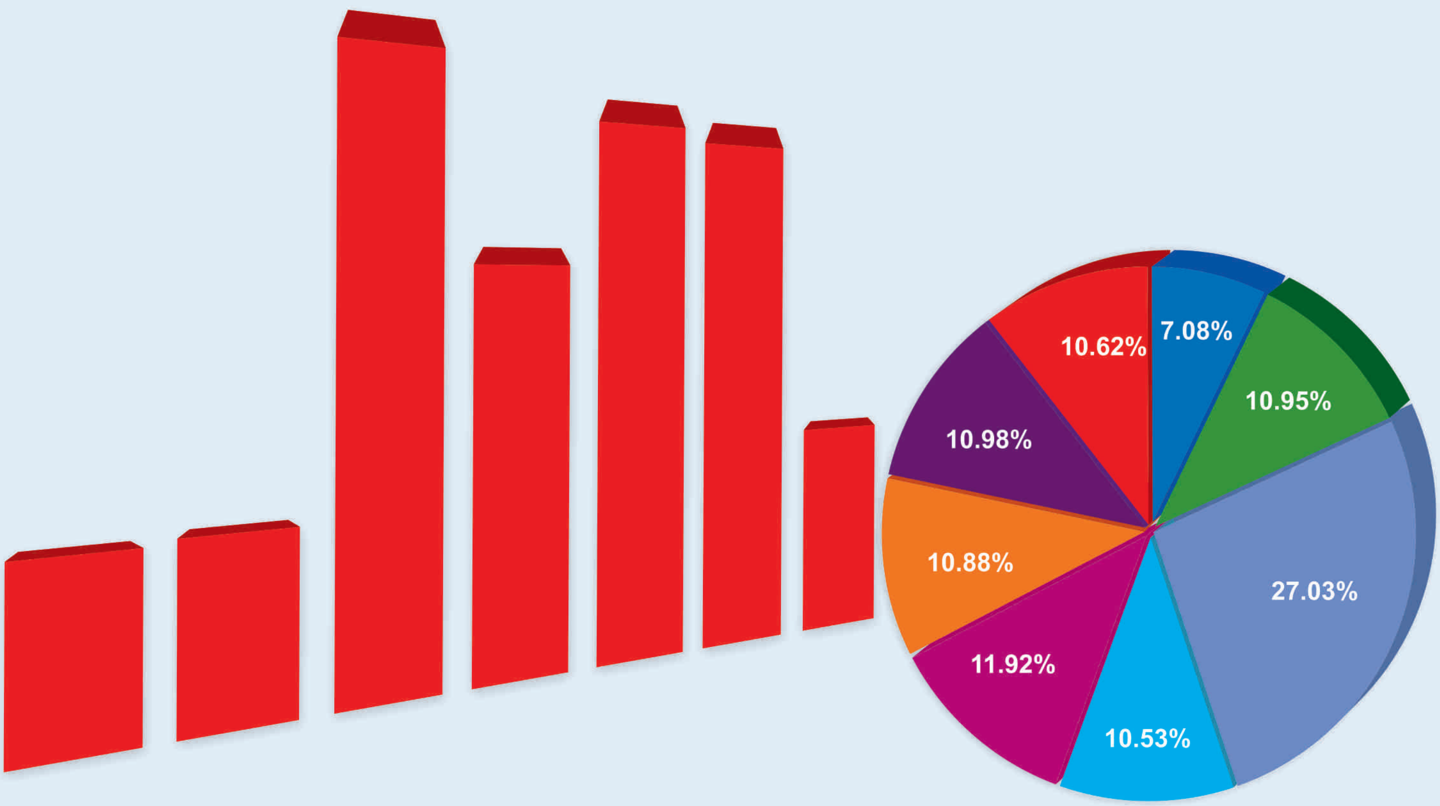
Think IAS...

दृष्टि The Vision

Think Drishti

बिहार लोक सेवा आयोग (BPSC)

सांख्यिकीय विश्लेषण, आरेखन और चित्रण



दूरस्थ शिक्षा कार्यक्रम *(Distance Learning Programme)*

Code: BRPM24

बिहार लोक सेवा आयोग (BPSC)

सांख्यिकीय विश्लेषण, आरेखन और चित्रण

दृष्टि The Vision

641, प्रथम तल, डॉ. मुखर्जी नगर, दिल्ली-110009
दूरभाष: 011-47532596, 8750187501
टोल फ्री : **1800-121-6260**
Web: www.drishtiIAS.com
E-mail : online@groupdrishti.com

पाठ्यक्रम, नोट्स तथा बैच संबंधी updates निरंतर पाने के लिये निम्नलिखित पेज को "like" करें
www.facebook.com/drishtithevisionfoundation
www.twitter.com/drishtiias

विषय सूची (Contents)

1.	सांख्यिकी : एक परिचय	5 – 6
2.	आँकड़ों का संकलन	7 – 9
3.	आँकड़ों का संगठन	10 – 19
4.	आँकड़ों का प्रस्तुतीकरण	20 – 74
5.	केंद्रीय प्रवृत्ति की माप	75 – 116
6.	विक्षेपण की माप	117 – 138
7.	सह-संबंध	139 – 149
8.	प्रायिकता	150 – 158

अध्याय 1

सांख्यिकी : एक परिचय (Statistics : An Introduction)

परिचय *(Introduction)*

सांख्यिकी गणित की वह शाखा है जिसमें आँकड़ों का संग्रहण, प्रदर्शन, वर्गीकरण और उसके गुणों के आकलन का अध्ययन किया जाता है।

दूसरे शब्दों में कहें तो आँकड़ों के संकलन, संघटन, प्रस्तुतीकरण, विश्लेषण तथा अर्थनिर्णयन में जो विधियाँ प्रयोग की जाती हैं उनका अध्ययन ही सांख्यिकी कहलाता है।

सांख्यिकी अर्थात् 'स्टेटिस्टिक्स' (Statistics) शब्द की उत्पत्ति लैटिन भाषा के शब्द 'स्टेट्स' (Status) या इतालवी शब्द 'स्टेटी' (Stati) से हुई है जिसका अर्थ होता है 'राजनीतिक राज्य'। प्राचीन समय में इसे शासन-कला विज्ञान समझा जाता था तथा राज्य की सरकारें प्रशासनिक उद्देश्य हेतु इसका प्रयोग करते थीं। अत: तब इसका क्षेत्र सीमित था।

वर्तमान में सांख्यिकी का क्षेत्र व्यापक हो गया है। जैसे कि अर्थशास्त्र, उद्योग, सामाजिक संरचना, प्रशासन, शोध आदि क्षेत्र में सांख्यिकी का प्रयोग किया जा रहा है।

टिपेट (Tippett) के अनुसार, "सांख्यिकी विज्ञान तथा कला दोनों है। यह विज्ञान इसलिये है क्योंकि इसकी विधियाँ मूल रूप से व्यवस्थित हैं एवं उनका हर जगह प्रयोग होता है, साथ ही यह कला इसलिये है कि इसकी विधियों का सफल प्रयोग पर्याप्त सीमा तक सांख्यिकीविद् की योग्यता, अनुभव एवं उसके प्रयोग क्षेत्र आदि के ज्ञान पर निर्भर करता है।"

सांख्यिकी के अनुप्रयोग *(Applications of Statistics)*

- राष्ट्रीय विकास के लिये योजना बनाने में
- उपभोक्ता कीमतों के सूचकांक में
- जनसंख्या वृद्धि में
- सांख्यिकीय कोटि नियंत्रण में
- बाज़ार सर्वेक्षण और मतगणना में
- स्वास्थ्य के संबंध में सांख्यिकी
- कृषि उत्पादन में।

सांख्यिकी की सीमाएँ *(Limitations of Statistics)*

- सांख्यिकी द्वारा सामूहिक इकाइयों का अध्ययन किया जा सकता है तथा निष्कर्ष निकाला जा सकता है, किंतु व्यक्तिगत इकाइयों का अध्ययन इसके द्वारा नहीं किया जा सकता।
- सांख्यिकी में आँकड़ों का केवल संख्यात्मक मान प्रयोग किया जाता है। ऐसे गुणात्मक तथ्य जिन्हें संख्यात्मक रूप में प्रस्तुत नहीं किया जा सकता, उनका अध्ययन इसके द्वारा संभव नहीं है।
- सांख्यिकी विधियों द्वारा निकाले गए निष्कर्ष कई बार वास्तविकता से अलग प्रतीत होते हैं तथा भ्रम पैदा करते हैं।
- सांख्यिकी के नियम सार्वभौमिक सत्य नहीं होते, इन्हें केवल औसत रूप में ही सत्य माना जा सकता है।
- सांख्यिकी का प्रयोग आम आदमी नहीं समझ सकता, विशेषज्ञों द्वारा ही इसका प्रयोग समझा जा सकता है।
- सांख्यिकी के परिणाम केवल सामान्य रूप से लागू हो सकते हैं। इसमें विशिष्टता का अभाव होता है।
- आँकड़ों के गलत प्रस्तुतीकरण से सांख्यिकी का दुरुपयोग किया जा सकता है।

तथ्यों के संख्यात्मक रूप को आँकड़ा कहा जाता है। कुछ तथ्य संख्यात्मक रूप में ही होते हैं जबकि कुछ तथ्य गुणों पर आधारित होते हैं, जैसे- सुंदरता, बुद्धिमत्ता, रंग, रूप, खुशी इत्यादि। सांख्यिकी विधियों से विश्लेषण करने के लिये इन्हें संख्यात्मक मान दे दिया जाता है एवं अध्ययन के योग्य बनाया जाता है। इस संख्यात्मक मान को कोटि (Rank) कहा जाता है।

आँकड़ों के प्रकार *(Types of Data)*

1. **भौगोलिक आँकड़े:** वे आँकड़े जो किसी क्षेत्र की भौगोलिक विशेषता के आधार पर प्राप्त होते हैं उन्हें भौगोलिक आँकड़े कहा जाता है। जैसे- राज्य की जनसंख्या, पैदावार, वर्षा की मात्रा, उच्चावच इत्यादि।

अध्याय
2

आँकड़ों का संकलन
(Collection of Data)

सांख्यिकी विश्लेषण के लिये सबसे पहला कार्य आँकड़ों का संग्रह/संकलन करना है। किंतु संग्रहित अथवा संकलित आँकड़े शुद्ध एवं सटीक होने चाहिये क्योंकि अशुद्ध आँकड़े विश्लेषण के बाद गलत परिणाम देते हैं। इसलिये जरूरत के हिसाब से यह निर्धारण अति आवश्यक है कि आँकड़े कब, कहाँ और कैसे एकत्र किये जाए। जिससे कि शुद्धतम परिणाम निकल सके।

संग्रह स्रोत के आधार पर निम्न दो प्रकार से वर्गीकृत किये जाते हैं-

प्राथमिक आँकड़े *(Primary Data)*

प्राथमिक आँकड़े वे आँकड़े होते हैं जिन्हें अनुसंधानकर्त्ता स्वयं प्रेक्षण स्थल पर जाकर संकलित करता है। किसी अनुसंधान में इनका प्रत्यक्ष रूप से प्रयोग पहली बार होता है। इसलिये प्राथमिक आँकड़ों की प्रकृति मौलिक होती है। ये आँकड़े पहले से किसी रिपोर्ट, रिसर्च पेपर या अन्य समाचार पत्र-पत्रिका आदि में प्रकाशित नहीं रहते हैं।

उदाहरण के लिये, एक अनुसंधानकर्त्ता द्वारा किसी सरकारी कार्यालय में जाकर शाकाहारी अथवा माँसाहारी व्यक्तियों के बारे में आँकड़े संकलित करना। इस प्रकार के आँकड़े प्राथमिक आँकड़े कहलाते हैं। प्राथमिक आँकड़े प्रत्यक्ष वैयक्तिक अनुसंधान, अप्रत्यक्ष अनुसंधान, संवाददाता द्वारा, डाक प्रश्नावली द्वारा तथा अनुसूची इत्यादि विधि द्वारा संकलित किये जाते हैं।

द्वितीयक आँकड़े *(Secondary Data)*

द्वितीयक आँकड़े वे आँकड़े हैं जिन्हें अनुसंधानकर्त्ता स्वयं एकत्र नहीं करता बल्कि व उन्हें अन्य स्रोतों से करता है। प्रकाशित अथवा लिखित पत्र-पत्रिका, समाचार-पत्र, डायरी, सरकारी रिपोर्ट आदि किसी से भी इन्हें लिया जा सकता है। जैसे- किसी प्रेक्षण में यदि हम जनगणना, 2011 के आँकड़ों का प्रयोग करें तो वे द्वितीयक आँकड़े होंगे।

आँकड़ों का सर्वेक्षण *(Survey of Data)*

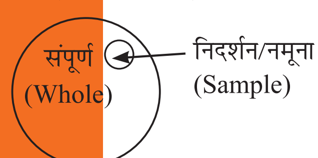
आँकड़ों का सर्वेक्षण दो विधि से किया जाता है।

प्रथम विधि

इस विधि में, जिस समूह का अध्ययन करना है उसके प्रत्येक अवयव को अध्ययन में शामिल किया जाता है तथा सभी अवयवों से कोई निष्कर्ष प्राप्त किया जाता है। भारत की जनगणना इसी पद्धति पर आधारित है। इस कारण इसे जनगणना पद्धति भी कहा जाता है।

द्वितीय विधि

इस विधि से सांख्यिकी अध्ययन में समस्त अवयवों को शामिल न करके कुछ अवयवों/तत्त्वों को नमूने (Sample) के रूप में शामिल किया जाता है तथा इन तत्त्वों का ही अध्ययन कर पूरे समूह के लिये निष्कर्ष निकाल लिया जाता है। इस विधि को निदर्शन पद्धति (Sampling Method) कहते हैं। इस पद्धति में लिया गया सैम्पल पूरे समूह का प्रतिनिधित्व करता है। इसे निदर्शन या न्यायदर्श कहा जाता है।



अभ्यास प्रश्न

1. सांख्यिकी का प्रयोग किया जाता है-
   (a) उद्योगों में
   (b) प्रशासन व शोध के क्षेत्रों में
   (c) बैंकिग व बीमा क्षेत्रों में
   (d) उपर्युक्त सभी में

2. सांख्यिकी है-
   (a) विज्ञान
   (b) कला
   (c) विज्ञान एवं कला दोनों
   (d) इनमें से कोई नहीं

3. सांख्यिकी द्वारा अध्ययन किया जाता है-
   (a) केवल व्यक्तिगत इकाइयों का
   (b) दो अलग-अलग व्यक्तिगत इकाइयों का
   (c) सामूहिक इकाइयों का
   (d) इनमें से कोई नहीं
4. सांख्यिकी का क्षेत्र है-
   (a) व्यापक (b) सीमित
   (c) (a) व (b) दोनों (d) इनमें से कोई नहीं
5. 'स्टेटी' (Stati) शब्द का अर्थ है:
   (a) राजनीतिक मित्र
   (b) राजनीतिक शिष्टाचार
   (c) राजनीतिक वैश्वीकरण
   (d) राजनीतिक राज्य
6. किसके अनुसार सांख्यिकी, विज्ञान एवं कला दोनों है?
   (a) गॉटफ्रायड ऑकेनवाल
   (b) टिपेट
   (c) डॉ. मैकस ओ.लॉरेंज
   (d) उपर्युक्त कोई नहीं
7. निम्नलिखित में से किसके द्वारा सांख्यिकी का अध्ययन नहीं किया जा सकता है?
   (a) सामूहिक इकाइयों द्वारा
   (b) व्यक्तिगत इकाइयों द्वारा
   (c) उपर्युक्त (a) तथा (b) दोनों द्वारा
   (d) उपर्युक्त में से कोई नहीं
8. "सांख्यिकी गणना का विज्ञान है।" किसने कहा है?
   (a) एफ.एन. डेविड
   (b) प्रो. केल्विन सेड्रूज
   (c) बॉउले
   (d) मार्शल
9. जब कोई अनुसंधानकर्त्ता स्वयं प्रेक्षण स्थल पर जाकर आँकड़े एकत्र करता है, तो आँकड़े कहलाते हैं-
   (a) प्राथमिक आँकड़े (b) द्वितीयक आँकड़े
   (c) विशिष्ट आँकड़े (d) इनमें से कोई नहीं
10. यदि वर्तमान में किसी सर्वे में भारतीय जनगणना 2011 में प्रयुक्त आँकडों का पुन: प्रयोग किया जाए तो उन आँकड़ों को कहा जाएगा-
   (a) प्राथमिक आँकड़े
   (b) द्वितीयक आँकड़े
   (c) विशिष्ट आँकड़े
   (d) इनमें से कोई नहीं
11. किसी समूह या समंकों के अध्ययन में प्रत्येक अवयव को शामिल करना कौन-सी पद्धति कहलाती है?
   (a) उत्पादन पद्धति (b) निदर्शन पद्धति
   (c) आय-व्यय पद्धति (d) जनसंख्या पद्धति
12. आँकड़ों का सर्वेक्षण कितने प्रकार से किया जा सकता है?
   (a) एक (b) दो
   (c) तीन (d) चार
13. समस्त अवयवों या समंकों को शामिल न करके कुछ अवयवों/तत्वों को नमूने के रूप में शामिल करना कौन-सी पद्धति कहलाती है?
   (a) निदर्शन पद्धति (b) आय-व्यय पद्धति
   (c) उत्पादन पद्धति (d) जनगणना पद्धति
14. सांख्यिकीय विश्लेषण के लिये सबसे पहला तथा प्रमुख कार्य है-
   (a) वस्तुओं का मूल्यांकन करना
   (b) आँकड़ों का संग्रह करना
   (c) जनगणना के आँकड़ों का निर्धारण करना
   (d) इनमें से कोई नहीं
15. त्रुटिपूर्ण आँकड़े विश्लेषण के बाद कैसे परिणाम देते है?
   (a) शुद्ध (b) गलत
   (c) सटीक (d) उपर्युक्त कोई नहीं
16. राम किसी स्थान पर जाकर कुछ आँकड़े एकत्र करता है, तो प्राप्त आँकड़े कहलाते हैं-
   (a) प्राथमिक आँकड़े (b) द्वितीयक आँकड़े
   (c) विभिष्ट आँकड़े (d) इनमें से कोई नहीं
17. आँकड़ों को कितने प्रकार से वर्गीकृत किया जा सकता है?
   (a) एक (b) दो
   (c) तीन (d) चार
18. विभिन्न प्रकार के आँकड़ों को एकत्र करना कहलाता है-
   (a) आँकडों का संग्रह (b) द्वितीयक आँकड़े
   (c) दंड बारंबारता वक्र (d) बहुस्तंभी चार्ट

उत्तरमाला

1. (d)	2. (c)	3. (c)	4. (a)	5. (d)
6. (b)	7. (b)	8. (c)	9. (a)	10. (b)
11. (d)	12. (b)	13. (a)	14. (b)	15. (b)
16. (a)	17. (b)	18. (a)		

अभ्यास प्रश्नों के हल

1. सांख्यिकी का उपयोग उद्योग में, प्रशासन व शोध के क्षेत्रों में तथा बैंकिग व बीमा क्षेत्रों में किया जाता है।
2. सांख्यिकी विज्ञान एवं कला दोनों है।
3. सांख्यिकी द्वारा केवल सामूहिक इकाइयों का अध्ययन किया जाता है।
4. सांख्यिकी का क्षेत्र व्यापक है।
5. 'स्टेटी' शब्द का अर्थ 'राजनीतिक राज्य' होता है।
6. टिपेट के अनुसार सांख्यिकी विज्ञान तथा कला दोनों है।
7. व्यक्तिगत इकाइयों के द्वारा सांख्यिकी का अध्ययन नहीं किया जाता है।
8. "सांख्यिकी गणना का विज्ञान है"-बाउले ने कहा है।
9. जब कोई अनुसंधानकर्त्ता स्वयं प्रेक्षण स्थल पर जाकर आँकड़े एकत्र करता है, तो उन आँकड़ों को प्राथमिक आँकड़े कहते हैं।
10. वर्तमान में किसी सर्वे में भारतीय जनगणना 2011 में प्रयुक्त आँकड़ों का पुनः प्रयोग किया जाए तो उन आकड़ों को द्वितीयक आँकड़े कहा जाएगा।
11. समूह या समंकों के अध्ययन में प्रत्येक अवयव को शामिल करना जनसंख्या पद्धति कहलाती है।
12. आँकड़ों का सर्वेक्षण दो प्रकार से किया जाता है।
13. प्रयुक्त प्रश्न में उपयोग की जाने वाली पद्धति निदर्शन पद्धति कहलाती है।
14. सर्वप्रथम सांख्यिकी विश्लेषण के लिये आँकड़ों का संग्रह किया जाता है।
15. त्रुटिपूर्ण आँकड़े हमेशा गलत परिणाम प्रदर्शित करते हैं।
16. राम के द्वारा एकत्र किये गए आँकड़े प्राथमिक आँकड़े कहलाते हैं।
17. आँकड़े दो प्रकार से वर्गीकृत किये जाते हैं-
    (a) प्राथमिक आँकड़े
    (b) द्वितीयक आँकड़े
18. विभिन्न प्रकार के आँकड़ों को एकत्रत करना आँकड़ों का संग्रह कहलाता है।

अध्याय 3

आँकड़ों का संगठन (Organisation of Data)

पिछले अध्याय में हमने जाना कि तथ्यों के संख्यात्मक रूप को आँकड़े कहा जाता है, अत: जब किसी विषय से जुड़े हुए संख्यात्मक तथ्यों का चयन कर उन्हें इस प्रकार संगठित अथवा व्यवस्थित किया जाए जिससे कि उनका अध्ययन एवं विश्लेषण आसानी से किया जा सके तो इस प्रक्रिया को आँकड़ों का संगठन कहा जाता है। आँकड़ों के संगठन को आँकड़ों का वर्गीकरण या आँकड़ों का व्यवस्थीकरण भी कहा जाता है। सर्वप्रथम आँकड़ों के वर्गीकरण से जुड़ी शब्दावली को समझते हैं-

बारंबारता *(Frequency)*

किसी आँकड़े के समूह में किसी चर का मान जितनी बार आता है उसे उस चर की आवृत्ति या बारंबारता (Frequency) कहते हैं। इसे 'f' अक्षर से निरूपित करते हैं।

उदाहरण: माना एक कक्षा में 15 विद्यार्थी हैं जिनमें से 5 की आयु 14 वर्ष, 3 की आयु 16 वर्ष, 4 की आयु 15 वर्ष तथा 3 की आयु 13 वर्ष है। तो आयु के लिये चर के मानों की आवृत्ति होगी-

हल:

आयु (x)	13	14	15	16
आवृत्ति (f)	3	5	4	3

यदि आँकड़ों का व्यवस्थीकरण वर्गीकृत रूप में हो तो वहाँ चर की बारंबारता न लेकर वर्ग-अंतराल की बारंबारता लेते हैं। उदाहरण के लिये यदि पिछले आँकड़ों के चरों को व्यवस्थित किया जाए तो वर्ग-अंतरालों (Class Interval) की आवृत्ति निम्न प्रकार से होगी-

वर्ग-अंतराल	बारंबारता (f)
10-14	8
15-19	7

मिलान चिह्न *(Tally Marks)*

चरों के मान की बारंबारता को दर्शाने के लिये मिलान चिह्नों का प्रयोग किया जाता है। मिलान चिह्न रेखाओं को गिनकर किसी मान या वर्ग समूह की बारंबारता को भी जाना जा सकता है।

टैली मार्क को प्रदर्शित करने हेतु बारंबारता को 5-5 के समूह में दर्शाते हैं। 1 से 4 तक की बारंबारता को क्रमश: I, II, III, IIII चिह्नों द्वारा दर्शाते हैं जबकि बारंबारता 5 के लिये तिरछी रेखा युक्त ~~IIII~~ चिह्न का प्रयोग करते हैं। जैसे कि-

आयु	बारंबारता	मिलान चिह्न
13	3	III
14	5	~~IIII~~
15	4	IIII
16	3	III

परिसर/परास *(Range)*

किसी भी आँकड़ा समूह में चर के अधिकतम मान तथा न्यूनतम मान के अंतर को परास कहा जाता है। इसका अर्थ है किसी भी चर का मान दी गई परास से बाहर नहीं हो सकता। जैसे- यदि किसी कक्षा में सबसे कम उम्र के विद्यार्थी की आयु 15 वर्ष तथा सबसे अधिक उम्र के विद्यार्थी की आयु 18 वर्ष है तो,

कक्षा के विद्यार्थियों की आयु के लिये परास = 18 – 15
= 3 वर्ष

अर्थात् कक्षा के प्रत्येक विद्यार्थी की आयु 15 से 18 के बीच ही होगी।

परास (R) = वर्गों की संख्या (n) × वर्ग की लंबाई (h)

आँकड़ों के वर्गीकरण/व्यवस्थीकरण की विधियाँ *(Method/Classification or Arrangement of Data)*

आँकड़ों को व्यवस्थित करने के लिये उन्हें विभिन्न समूहों या वर्गों में बाँटते हैं। आँकड़ों के व्यवस्थीकरण की दो विधियाँ हैं-

1. *अवर्गीकृत विधि (Unclassified Method)*

इसमें आँकड़ों को समूह में न रखते हुए प्रत्येक अवयव को उसके व्यक्तिगत मान के साथ आरोही या अवरोही क्रम में व्यवस्थित किया जाता है। इस विधि में आँकड़े की हानि नहीं होती क्योंकि प्रत्येक अवयव का प्रयोग अध्ययन में किया जाता है।

अध्याय
4

आँकड़ों का प्रस्तुतीकरण (Presentation of Data)

विभिन्न प्रेक्षणों से प्राप्त आँकड़ों को संग्रहित कर एक व्यवस्थित रूप में प्रदर्शित अथवा प्रस्तुत करना ही आँकड़ों का प्रस्तुतीकरण है। आँकड़ों के प्रस्तुतीकरण से संग्रहित किये गए वृहद् आँकड़ों का आसानी से समझा एवं प्रयोग किया जा सकता है।

आँकड़ों का प्रस्तुतीकरण निम्न प्रकार से किया जा सकता है–

वर्णात्मक विधि द्वारा आँकड़ों का प्रस्तुतीकरण

इस विधि में किसी प्रेक्षण के आँकड़ों को भाषा के रूप में दिया जाता है। इस विधि से प्रस्तुत आँकड़ों से कोई निष्कर्ष निकालने में काफी समय लगता है। सामान्यत: अंकगणित आदि के प्रश्नों में इसका प्रयोग देखा जा सकता है।

सारणी के रूप में आँकड़ों का प्रस्तुतीकरण

इस विधि में वर्गीकृत आँकड़ों को पंक्तियों तथा स्तंभों की सहायता से दर्शाया जाता है। सारणी आँकड़ों को संक्षिप्त रूप प्रदान करती है, साथ ही सारणी के प्रयोग से आँकड़ों का तुलनात्मक अध्ययन सरल बनता है, जिससे विश्लेषण में आसानी होती है।

सारणीयन *(Tabulation)*

पंक्तियों तथा स्तंभों के रूप में आँकड़ों का प्रस्तुतीकरण ही सारणीयन कहलाता है।

सारणियाँ तीन प्रकार की हो सकती हैं–

1. **सरल सारणी:** किसी एक गुण पर आधारित आँकड़ों को सरल सारणी द्वारा प्रदर्शित किया जाता है। जैसे– कक्षा के विद्यार्थियों के गणित में प्राप्तांक

अनुक्रमांक	–	–	–	–	–
प्राप्तांक	–	–	–	–	–

2. **द्विगुणी सारणी:** इसमें एक ही आँकड़ों के दो गुणों को प्रदर्शित किया जाता है। जैसे–

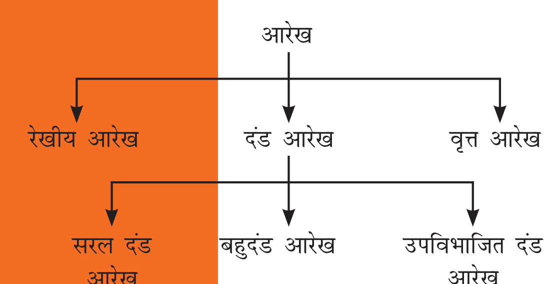
वर्ष	जनसंख्या		
	शिक्षित	अशिक्षित	योग

3. **बहुगुणी सारणी:** इस सारणी में एक ही प्रकार के आँकड़ों के विभिन्न गुणों को प्रदर्शित किया जाता है। जैसे– किसी देश की जनगणना के आँकड़ों से प्राप्त सारणी

आयु वर्ग	पुरुष		महिला		ट्रांसजेंडर	
	साक्षर	निरक्षर	साक्षर	निरक्षर	साक्षर	निरक्षर
योग	–	–	–	–	–	–

आरेख द्वारा आँकड़ों की प्रस्तुति

इस विधि में आँकड़ों को उनके निकटतम रूप में प्रस्तुत किया जाता है। इस विधि में आँकड़ों की तुलना काफी आसान हो जाती है। परंतु निकटतम रूप होने के कारण आँकड़ों का गणितीय उपयोग संभव नहीं हो पाता।



रेखीय आरेख *(Line Diagram)*

रेखीय आरेख आँकड़ों के एक गुण को प्रदर्शित करता है। इसमें समान दूरी पर विभिन्न रेखाएँ स्थित होती हैं। जो आँकड़ों के मान को दर्शाती हैं। यह आरेख एक वित्तीय होता है।

अध्याय 5

केंद्रीय प्रवृत्ति की माप (Measures of Central Tendency)

अभी तक हमने पिछले अध्याय में आँकड़ों को अवर्गीकृत तथा वर्गीकृत बारंबारता बंटनों में व्यवस्थित करना सीखा है, साथ ही आँकड़ों को चित्रीय रूप में विभिन्न आलेखों तथा बहुभुजों के रूप में निरूपित करना भी जाना है। विभिन्न आँकड़ों की तुलना करने तथा कोई निष्कर्ष निकालने के क्रम में हमें सभी आँकड़ों के अध्ययन की आवश्यकता न पड़े इसलिये हम आँकड़ों से कुछ प्रतिनिधि मान प्राप्त करते हैं जो समस्त आँकड़ों का प्रतिनिधित्व करते हैं। इन्हीं प्रतिनिधि मानों को केंद्रीय प्रवृत्ति की माप कहा जाता है।

यह प्रतिनिधि मान आँकड़ों का औसत अथवा केंद्रीय मान कहलाता है। इसके इर्द-गिर्द ही आँकड़ों का संकेंद्रण रहता है। किसी आँकड़ा समूह के निम्न केंद्रीय मान होते हैं–

केंद्रीय प्रवृत्ति की माप

गणितीय औसत
- समांतर माध्य (Arithmetic Mean)
- गुणोत्तर माध्य (Geometric Mean)
- हरात्मक माध्य (Harmonic Mean)

स्थितिक औसत
- माध्यिका (Median)
- बहुलक (Mode)

समांतर माध्य *(Arithmetic Mean)*

माध्य का सामान्य अर्थ औसत होता है। आँकड़ा समूह के प्रत्येक मान के योगफल को आँकड़ों की संख्या से भाग देने पर माध्य प्राप्त होता है। माध्य को ही अंकगणितीय माध्य या समांतर माध्य कहा जाता है। यह दो प्रकार का होता है–

- सरल माध्य (Simple Arithmetic Mean)
- भारित माध्य (Weighted Arithmetic Mean)

सरल माध्य में सभी पदों के मान को समान महत्त्व दिया जाता है, जबकि भारित माध्य में पदों के मान का महत्त्व अलग-अलग होता है।

$$\text{सरल माध्य} = \frac{\text{समस्त पदों का योग}}{\text{पदों की संख्या}}$$

$$\overline{x} = \frac{\sum x}{n}$$

माध्य को $\overline{x}$ (x बार) से प्रदर्शित करते हैं। Σx सभी पदों के योग को दर्शाता है। कई बार पदों की संख्या अधिक होने पर पदों के योगफल को निम्न प्रकार से निरूपित किया जाता है।

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + x_4 + \ldots + x_n$$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\text{भारित माध्य } \frac{\sum wx}{\sum w}$$

$$\overline{x} = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3 + \ldots + x_n w_n}{w_1 + w_2 + w_3 + \ldots + w_n}$$

भारित माध्य निकालने के लिये सबसे पहले प्रत्येक आँकड़े के मान को उसके भार से गुणा करते हैं। फिर सभी भारित मानों को जोड़कर कुल भारों के योगफल से भाग दिया जाता है।

उदाहरण 1: राशियों 1, 2, 3, 4, 5, 6, ... 15 का माध्य ज्ञात कीजिये।

हल: माध्य $\overline{x} = \frac{1+2+3+4+5+\ldots+15}{15}$

$$= \frac{1}{2}(15 \times 16) \times \frac{1}{15}$$

$$= \frac{1}{2} \times 16 = 8$$

उदाहरण 2: एक व्यक्ति ने 50 कलम औसत मूल्य ₹ 12 प्रति कलम की दर से खरीदे तथा 30 कलम औसत मूल्य ₹ 10 प्रति कलम की दर से खरीदे। सभी कलम का माध्य मूल्य ज्ञात करें।

हल: सामूहिक माध्य, $\overline{x} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2}{n_1 + n_2}$

$$= \frac{50 \times 12 + 30 \times 10}{50 + 30} = \frac{900}{80}$$

= ₹ 11.25

अध्याय 6

विक्षेपण की माप (Measures of Dispersion)

विक्षेपण *(Dispersion)*

पिछले अध्याय में हमने अध्ययन किया कि कैसे आँकड़ों को प्रतिनिधि मूल्य के रूप में एकत्र किया जा सकता है। परंतु वह मूल्य आँकड़ों में विदित परिवर्तनशीलता को नहीं बता पाती है, लेकिन विक्षेपण की मापों में आप उन सभी मापों का अध्ययन करेंगे जो समंकों में परिवर्तनशीलता को मापने में सक्षम है।

विक्षेपण सांख्यिकी में एक विशेष सतत चर के लिये अपेक्षित परिणामों की सीमा के आकार का वर्णन करता है। यह केंद्रीय प्रवृत्ति की भिन्न-भिन्न शाखाओं, जैसे- माध्य, माध्यिका और बहुलक की दी गई बारंबारता के चरों (Variables) के वास्तविक आँकड़ों पर निर्भर करता है। इसलिये इसे प्रथम कोटि का औसत (Average of first order) भी कहा जाता है।

समांतर माध्य, गुणोत्तर माध्य तथा हरात्मक माध्य के द्वारा हम विभिन्न समंकों का औसत वितरण ज्ञात करते हैं।

एल.आर.कोर्नर के अनुसार "विक्षेपण, उसी सीमा तक की माप है, जिसमें भिन्न-भिन्न पद अलग-अलग होते हैं।"

केंद्रीय प्रवृत्ति की माप या औसत द्वारा हमें वितरण के मध्य भाग के बारे में प्रेक्षण की एकाग्रता का एक सरल विचार मालूम होता है।

विक्षेपण के द्वारा किसी आँकड़े के अलग-अलग पदों की परिवर्तनशीलता को मालूम किया जा सकता है। किसी भी सतत या असतत आँकड़ों के लिये औसत को अर्थपूर्ण होने के लिये उसका विश्लेषण पदों के बिखराव के समान होना चाहिये।

यदि पदों में बिखराव कम-से कम रहता है तो उन पदों का औसत अत्यधिक विशिष्ट तथा आवश्यक होता है।

इस प्रकार हम समझ सकते हैं कि औसत किसी वितरण का मुखिया होता है तथा विक्षेपण से हम यह मालूम कर सकते हैं कि पदों का औसतन बिखराव कितना है।

उपर्युक्त जानकारी से यह मालूम होता है कि माध्य, माध्यिका तथा बहुलक किसी भी आवृत्ति वितरण की पूरी जानकारी देने में नाकाम है। इस प्रकार हम भली-भाँति कह सकते हैं कि आँकड़ों के सम्यक विश्लेषण के लिये पदों की विचरणशीलता की जानकारी अति आवश्यक है।

माध्य, माध्यिका तथा बहुलक के द्वारा विभिन्न पदों के लिये विचलन ज्ञात कर विक्षेपण की माप के लिये औसत निकाला जा सकता है ताकि दिये गए आँकड़ों के संबंध में अधिक से अधिक जानकारी प्राप्त हो सके। इसलिये इसे द्वितीय कोटि का औसत (Average of Second order) भी कहा जाता है।

विक्षेपण की मापों के महत्त्वपूर्ण गुण

(i) यह गणना के लिये आसान तथा सरल होना चाहिये।

(ii) यह सख्ती या कठोरतापूर्वक (Rigidly) परिभाषित होना चाहिये।

(iii) इसमें नमूनाकरण स्थिरता होनी चाहिये।

(iv) यह चरम (Unduly) पदों से अनावश्यक रूप से प्रभावित नहीं होना चाहिये।

(v) यह सभी वस्तुओं पर आधारित होना चाहिये ताकि अधिक प्रतिनिधि हों।

(vi) यह आगे बीजगणितीय संक्रियाओं के योग्य होना चाहिये।

नोटः यदि केवल हमें औसत की जानकारी आँकड़े में मिलती है तो हम आवृत्ति वितरण के बारे में पूरी तरह से मूल्यांकन नहीं कर सकते हैं।

विचलन के प्रकार *(Types of Deviation)*

1. परास *(Range)*

किसी आवृत्ति वितरण में चर के अधिकतम (Largest) एवं न्यूनतम (Smallest) मानों के अंतर को परास कहते हैं।

परास (R) = अधिकतम मान (L) – न्यूनतम मान (S)

नोटः

(i) यदि आँकड़े वर्ग-अंतराल में दिये हुए हों तो वर्ग-अंतराल की उच्च सीमा तथा निम्न सीमा के अंतर को परास कहते हैं।

अध्याय 7

सह-संबंध (Correlation)

जब दो चरों (x तथा y) में से एक चर (x) में परिवर्तन होने पर दूसरे चर (y) पर प्रभाव पड़ता है, सह-संबंध कहलाता है। अन्य शब्दों में, सांख्यिकी में 'सह-संबंध' का आशय दो चरों के बीच संबंध की माप से है। दो चरों के बीच संबंधों या डिग्री की माप सह-संबंध के द्वारा ज्ञात की जाती है।

"यदि दो या दो से अधिक राशियाँ सहानुभूति में परिवर्तित होती हैं, जिनमें एक में होने वाले परिवर्तनों के फलस्वरूप दूसरी राशि में भी परिवर्तन की प्रवृत्ति पाई जाए, तो वे राशियाँ सह-संबंधित राशियाँ कहलाती हैं।"

– कार्नर

- सह-संबंध के माध्यम से चरों में गहनता तथा दिशा के गुणों की माप की जाती है। दोनों चरों के बीच संबंध की गहनता व दिशा को सह-संबंध गुणांक की सहायता से मात्रात्मक रूप में ज्ञात किया जाता है।
- सह-संबंध के गुणों का आशय उन चरों से है जिनका संख्यात्मक मान ज्ञात नहीं किया जा सकता, जैसे- लोगों की ईमानदारी, चरित्र, बौद्धिक क्षमता इत्यादि।

  इसके माध्यम से केवल सह-प्रसरण के गुणों का मापन किया जा सकता है।
- सह-संबंध गुणांक का मान इसके गुणांक की गहनता व दिशा के अनुक्रमानुपाती होता है।

  सह-संबंध (Correlation) का मान हमेशा -1 व $+1$ के मध्य $(-1 \le r \ge +1)$ होता है। सह-संबंध विश्लेषण हमे दो चरों के बीच संबंध की दिशा और डिग्री की माप को दर्शाता है। हम इसे उदाहरण के माध्यम से समझ सकते हैं।

  1. अधिक गर्मी होने पर आइसक्रीम व ठंडे पेय पदार्थों की मांग बढ़ जाना।
  2. किसी खाद्य पदार्थों की कीमतों में कमी होने पर उसकी मांगों में बढ़ोतरी होना।
  3. चर x में परिवर्तन चर y में परिवर्तन के अनुक्रमानुपाती होता है।

उपर्युक्त उदाहरण (1) में यदि गर्मी पड़ती है तो आइसक्रीम तथा ठंडे पेय पदार्थों की मांग बढ़ जाती है। अत: गर्मी पड़ने का आइसक्रीम व ठंडे पेय पदार्थों की मांग बढ़ने वाली घटना से प्रत्यक्ष संबंध है। इसी प्रकार उदाहरण (2) में यदि खाद्य पदार्थों में कमी होती है, तो इसकी मांग स्वत: ही बढ़ जाती है, खाद्य पदार्थों की कीमतों में कमी होने वाली घटना का संबंध मांग बढ़ने वाली घटना से है अर्थात् एक घटना का प्रभाव दूसरी घटना पर पड़ रहा है।

स्पष्ट है कि एक घटना का प्रभाव दूसरी घटना पर प्रत्यक्ष रूप से पड़ रहा है, यह क्रम सह-संबंध को प्रदर्शित करता है।

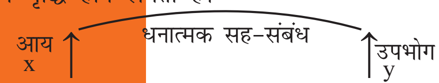
सह-संबंध चरों के बीच घटनाओं के परस्पर संबंध को दर्शाता है। यह तभी मान्य होगा जब दोनों घटनाओं के बीच कारण-प्रभाव का विशेष संबंध है, यदि इस प्रकार के कारण-प्रभाव या अन्य गुणों का अभाव हो तो सह-संबंध की माप को ज्ञात नहीं किया जा सकता।

सह-संबंध के प्रकार-

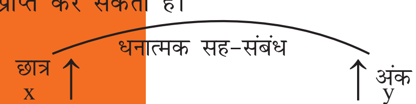
संह-सबंध मुख्यत: दो प्रकार के होते है।

1. धनात्मक सह-संबंध *(Positive Correlation)*

(a) यदि दो चर (x तथा y) का मान समान दिशा में बदलता है, धनात्मक सह-सबंध कहलाता है। जैसे- सामान्यत: किसी परिवार की आय बढ़ती है, तो उसके उपभोग में वृद्धि होने लगती है।



(b) यदि छात्र अधिक परिश्रम करे तो उसके सफल होने की संभावना अत्यधिक बढ़ जाती है। वह अच्छे अंक प्राप्त कर सकता है।



2. ऋणात्मक सह-संबंध *(Negative Correlation):*

जब दो चरों (x तथा y) के परिवर्तन की दिशा एक दूसरे के विपरीत हो, तो ऋणात्मक सह-संबंध कहलाता है। अन्य शब्दों में यदि चर x में कमी होगी तो y में वृद्धि होगी तथा चर x में वृद्धि होगी तो चर y में कमी होगी।

अध्याय 8

प्रायिकता (Probability)

प्रयोग *(Experiment)*

ऐसी प्रत्येक क्रिया जिसे करने पर कुछ परिणाम प्राप्त हों, प्रयोग कहलाती है। प्रयोग दो प्रकार के हो सकते हैं-

1. **निर्धारणात्मक प्रयोग:** ऐसे प्रयोग जो समान परिस्थितियों के अंतर्गत दोहराने पर समान परिणाम उत्पन्न करें, निर्धारणात्मक प्रयोग कहलाते हैं। जैसे- 2 और 2 को जोड़ना।
2. **यादृच्छिक प्रयोग:** ऐसे प्रयोग, जिन्हें एक समान परिस्थितियों में दोहराने पर भी समान परिणाम आना निश्चित न हो, उन्हें यादृच्छिक प्रयोग कहते हैं, जैसे एक सिक्के को उछालकर टॉस करना, एक पासे को फेंकना।

प्रतिदर्श समष्टि *(Sample Space)*

किसी प्रयोग को करने पर प्राप्त हो सकने वाले सभी संभव परिणामों के समुच्चय को प्रतिदर्श समष्टि (Sample Space) कहते हैं। इसे 'S' से निरूपित करते हैं।

उदाहरण:

1. किसी सिक्के को उछालने पर प्राप्त हो सकने वाले परिणाम = चित (Head) या पट (Tail)
   अत: प्रतिदर्श समष्टि, $S = \{H, T\}$
   कुल परिणामों की संख्या, $n(S) = 2$
2. एक पासे को फेंकने पर प्राप्त हो सकने वाले परिणाम = 1, 2, 3, 4, 5 या 6
   $S = \{1, 2, 3, 4, 5, 6\}$
   प्रतिदर्श समष्टि में परिणामों की संख्या $= n(S) = 6$
3. दो सिक्कों को एक साथ उछालने पर प्राप्त हो सकने वाले परिणाम $= \{H, T\} \times \{H, T\}$
   $= \{HH, HT, TH, TT\}$
   प्रतिदर्श समष्टि में परिणामों की संख्या $= n(S) = 4$

घटना *(Event)*

किसी भी प्रयोग के लिये, उसके प्रतिदर्श समष्टि के प्रत्येक उपसमुच्चय (सदस्य) को एक घटना कहते हैं। इसे 'E' से निरूपित करते हैं।

उदाहरण:

1. एक पासे को फेंकने पर 4 आना, एक घटना है।
   $E = \{4\}$
   अनुकूल परिणामों की संख्या $= n(E) = 1$
2. किसी पासे को फेंकने पर उस पर सम संख्या आने की घटना $E = \{2, 4, 6\}$
   अनुकूल परिणामों की संख्या $= n(E) = 3$

घटनाओं के प्रकार *(Types of Events)*

1. **सरल घटना (Elementary or Simple Event):** ऐसी घटना, जिसमें प्रयोग का केवल एक परिणाम होता है, अर्थात् $n(E) = 1$, को सरल घटना कहते हैं।
   जैसे- पासे को फेंकने पर 4 आना
   $E = \{4\} \Rightarrow n(E) = 1$
2. **संयुक्त घटना (Complex Event):** वे सभी घटनाएँ जो सरल घटनाएँ नहीं होतीं उन्हें संयुक्त घटना कहते हैं। जैसे किसी पासे को फेंकने पर उस पर विषम संख्याएँ आना, $E = \{1, 3, 5\} \Rightarrow n(E) = 3$
3. **स्वतंत्र घटनाएँ (Mutually Exclusive Events):** यदि दो घटनाएँ इस प्रकार हों कि एक घटना के घटित होने का प्रभाव दूसरी घटना पर नहीं पड़े तो वे स्वतंत्र घटनाएँ कहलाती हैं।
   जैसे कोहली का शतक बनाना और नरेंद्र मोदी का प्रधानमंत्री बनना एक-दूसरे से स्वतंत्र घटनाएँ हैं तथा कोहली का शतक बनाना और भारतीय टीम का मैच जीतना परतंत्र घटनाएँ हैं।
4. **पूरक घटनाएँ (Complementary Events):** किसी घटना E की पूरक घटना को E' या $\overline{E}$ से निरूपित करते हैं। घटना E की पूरक घटना E' का अर्थ है कि जब घटना E घटित नहीं होती है।

डी.एल.पी. बुकलेट्स की विशेषताएँ

- आयोग के नवीनतम पैटर्न पर आधारित अध्ययन सामग्री।
- पैराग्राफ, बुलेट फॉर्म, सारणी, फ्लोचार्ट तथा मानचित्र का उपयुक्त समावेश।
- विषयवस्तु की सरलता, प्रामाणिकता तथा परीक्षा की दृष्टि से उपयोगिता पर विशेष ध्यान।
- क्विक रिवीज़न हेतु प्रत्येक अध्याय में महत्त्वपूर्ण तथ्यों का संकलन।
- प्रत्येक अध्याय के अंत में विगत वर्षों में पूछे गए एवं संभावित प्रश्नों का समावेश।

Website : www.drishtiIAS.com
E-mail : online@groupdrishti.com

DrishtiIAS
YouTube Drishti IAS
drishtiias
drishtithevisionfoundation

641, First Floor, Dr. Mukherjee Nagar, Delhi-110009
Phones : 011-47532596, +91-8130392354, 813039235456