

ATAL BIHARI UNIVERSITY BILASPUR, CHHATISGARH

INSTITUTE OF ADVANCED STUDY IN EDUCATION

TWO YEAR

BACHELOR OF EDUCATION (B.Ed-PART -II)

SECOND YEAR

SUBJECT

PAPER -5 CODE 014.4 -U-4

CONTENT OF PEDAGOGY OF MATHEMATICS

UNIT-IV

BY

D.K. JAIN

HEAD OF MATHEMATICS DEPARTMENT

B.Ed. SECOND YEAR -PAPER -5 CODE 014.4 -U-4

B.Ed.-II -PEDAGOGY OF MATHEMATICS-II

Unit 4: Learning and Teaching of Mathematics-

Data Handling and Probabilistic Reasoning:

In this unit we will discuss different ways of representing data, analyzing data and interpreting data. We will try to focus on importance of the choice of data representation based on the context and the information needed. After the representation of data we will also look at various measures of central tendency. The focus of this discussion would be to understand the various measures of central tendency and not on computing them. We will also look at the concept of probability and connections to our daily lives. History of data handling and probability Conceptual understanding of some key topics like data representation in different situations and drawing meaningful conclusions from the organised data. Meaning and significances of representative values of central tendency (mean, mode and median), Probability and chance.

Children's understanding (reasoning patterns and misconceptions) Solving interesting problems based on these to develop a better understanding. These would be around understanding and analysing data and drawing conclusions, understanding and finding probability, analysing situations of probability and distributions and solving simple problems Constructing engaging class-rooms, exercises, problems, worksheets etc. for children. These could be on organising data, representing it and on analyzing it in order for them to get over their difficulties and common misconceptions Teacher's knowledge and challenges.

इस इकाई में हम डेटा का प्रतिनिधित्व करने, डेटा का विश्लेषण करने और डेटा की व्याख्या करने के विभिन्न तरीकों पर चर्चा करेंगे। हम संदर्भ और सूचना के आधार पर डेटा प्रतिनिधित्व की पसंद के महत्व पर ध्यान देने की कोशिश करेंगे। आंकड़ों के प्रतिनिधित्व के बाद हम केंद्रीय प्रवृत्ति के विभिन्न उपायों को भी देखेंगे। इस चर्चा का फोकस केंद्रीय प्रवृत्ति के विभिन्न उपायों को समझना होगा, न कि उनकी गणना करना। हम अपने दैनिक जीवन की संभावना और कनेक्शन की अवधारणा को भी देखेंगे। डेटा हैंडलिंग और प्रायिकता का इतिहास विभिन्न स्थितियों में डेटा प्रतिनिधित्व और ऑर्गनाइज्ड डेटा से सार्थक निष्कर्ष निकालने जैसे कुछ प्रमुख शीर्षों की वैचारिक समझ। केंद्रीय प्रवृत्ति (मतलब, मोड

और मंझला), संभावना और मौका के प्रतिनिधि मूल्यों का अर्थ और महत्व। बच्चों की समझ (तर्क पैटर्न और गलतफहमी) एक बेहतर समझ विकसित करने के लिए इन पर आधारित दिलचस्प समस्याओं का समाधान। ये आंकड़ों को समझने और उनका विश्लेषण करने और निष्कर्ष निकालने, समझने और खोजने के आसपास होंगे

संभाव्यता, संभाव्यता और वितरण की स्थितियों का विश्लेषण करना और सरल समस्याओं को हल करना बच्चों के लिए आकर्षक क्लास-रूम, अभ्यास कार्य, समस्याओं, कार्यपत्रकों आदि का निर्माण करना। ये डेटा को व्यवस्थित करने, उनका प्रतिनिधित्व करने और अपनी कठिनाइयों और सामान्य गलत धारणाओं पर काबू पाने के लिए इसे एनालिसिस पर दे सकते हैं। शिक्षक के ज्ञान और चुनौतियों को पढ़ना

Data Handling and Probabilistic Reasoning in Maths

Data आकड़े

The word 'Data' means collection of information in the form of numerical figures, or a set of given facts. A data is a collection of numbers gathered to give some information. Visual representation of data will help us to understand it better and remember the facts easily.

Example:

- The marks obtained by 10 students of a class in a test are:
76, 83, 95, 100, 56, 32, 80, 67, 75, 46
- The following table gives the data regarding the favourite game of 100 students of a school:

Sports	Cricket	Football	Tennis	Badminton
Number Of Students	40	30	25	5

Recording Data आकड़ों का अभिलेखीकरण या आकड़ों का संचय


The data is collected and recorded by the investigator with a definite objective in mind. The data may be collected individually by sending some questionnaires or some other instrument for data collection.


Raghav	—	Banana	Bhawana	—	Apple
Preeti	—	Apple	Manoj	—	Banana
Amar	—	Guava	Donald	—	Apple
Fatima	—	Orange	Maria	—	Banana
Armita	—	Apple	Uma	—	Orange
Raman	—	Banana	Akhtar	—	Guava
Radha	—	Orange	Ritu	—	Apple
Farida	—	Guava	Salma	—	Banana
Anuradha	—	Banana	Kavita	—	Guava
Rati	—	Banana	Javed	—	Banana







Here, we want to collect data list of the entire student name and their choice of fruit, hence we are recording the choice of fruit of each student. Here each entry is a numerical fact and it is called observation. After tabulating the data and representing it graphically we can reach to some meaningful conclusions.

Organisation of Data ऑकड़ों का संगठन

The data that is collected for the first time is called raw data or ungrouped data. When the data is arranged in any data it becomes grouped data. When the data is very large, it is difficult to arrange them in a grouped form. In that case small bars are used to represent data; such method is called Tally Method.

To get the required information, all observations should be recorded. We depict each observation with the help of tally marks. Tally marks are used to organise the observations. Record every observation by a vertical mark, but every fifth observation should be recorded by a mark across the four earlier marks. The fifth observation tally mark looks like this: 

For example, we have a group of persons and their sizes of shoes. The tabular form representing the tally marks is as shown here. Here  shows the count to be five plus three (i.e. eight).

Shoe size	Tally marks	Number of students
4		5
5	 III	8
6	 	10
7	 II	7
8		2

डेटा हैंडलिंग से तात्पर्य सूचनाओं को इकट्ठा करने, रिकॉर्ड करने और पेश करने की प्रक्रिया से है जो दूसरों के लिए सहायक है – उदाहरण के लिए, ग्राफ या चार्ट में डेटा हैंडलिंग को कभी-कभी आंकड़ों

के रूप में भी जाना जाता है और आप अक्सर मैथ्स और साइंस दोनों के अध्ययन में इसकी जरूरत अनकों बार पड़ेगी विशेष कर आप जब आकड़ों के आधार पर किसी अनुमान या निष्कर्ष तक पहुँचना होता है तक निष्कर्ष निकालने के लिए डेटा का विश्लेषण करना।

डाटा हैंडलिंग के प्रकार—:

Types of Data Handling

- Data and its Frequency Distribution.
- Pictographs.
- Bar Graphs.
- Histogram and Pie-Charts.
- Chance and **Probability**.
- Arithmetic Mean.
- Median and **Mode**.

डाटा को अनेक तरीके से प्रस्तुत किया जाता है , सामान्यतः अंको के रूप में यदि यह हैं तो सारणी या टेबल के रूप में इसे व्यवस्थित किया जाता है। यदि संकेतों या चित्रों या डायग्राम के रूप में डाटा का स्वरूप हो तो सभी को एकत्र कर प्रस्तुत करने के तरीके में भिन्नता हो सकती है।।टेबल्स, चार्ट और ग्राफ डेटा का प्रतिनिधित्व करने के सभी तरीके हैं, और उनका उपयोग दो व्यापक उद्देश्यों के लिए किया जा सकता है। पहला एक वैज्ञानिक अध्ययन की प्रक्रिया के हिस्से के रूप में डेटा के संग्रह, संगठन और विश्लेषण का करना होता है। हम इन्हीं उद्देश्यों के आधार पर उसे निम्न स्वरूपों में बदल लेते हैं

Presenting numbers in tables

- When to use tables. Tables are an effective **way of presenting data**: ...
- Table design. ...
- Examples of poor and better practice in the presentation of **data** in tables. ...
- Bar charts. ...
- Histograms. ...
- Pie charts. ...
- Line graphs. ...
- Scatter plots.

डेटा का विश्लेषण और व्याख्या Analysis and Interpretation of the Data

Data analysis and interpretation is the process of assigning meaning to the collected information and determining the conclusions, significance and implications of the findings. It is an important and exciting step in the process of research. In all research studies, **analysis** follows **data** collection.

डेटा का विश्लेषण और व्याख्या

डेटा विश्लेषण और व्याख्या एकत्रित जानकारी को अर्थ प्रदान करने और प्राप्त डाटा के आधार पर निष्कर्षों तक पहुँचा जाता है जिसके कारण के निष्कर्ष, महत्व और निहितार्थ निर्धारित करने की प्रक्रिया है। यह शोध की प्रक्रिया में एक महत्वपूर्ण और अनिवार्य कार्य है। सभी शोध अध्ययनों में, विश्लेषण डेटा संग्रह का ही अनुसरण करता है। आकड़ों के अर्थापन को प्राप्त करने के लिये आकड़ों को व्यवस्थित रूप में सारणी या तालिका में रूप में प्रस्तुत करते हुए उसका सांख्यिकीय विश्लेषण किया जाता है और निर्धारित किये गये प्रश्न का उत्तर प्राप्त करने या परिकल्पनाओं की पुष्टि करने के लिए की जाती है। इस प्रकार डाटा के विश्लेषण के पश्चात ही व्याख्या संभव है। इनके लिये हम गणित की एक शाखा जिसे सांख्यिकी कहते हैं उसका अध्ययन करेंगे, जिसमें सबसे महत्वपूर्ण पद है, केन्द्रीय प्रवृत्ति की माप अर्थात् Measures of central tendency के अन्तर्गत डाटा के विश्लेषण के बारे में अध्ययन करेंगे।

Measures of central tendency

केंद्रीय प्रवृत्ति की एक माप एक सारांश सांख्यिकीय है जो केंद्र बिंदु या किसी डेटासेट के विशिष्ट मूल्य का प्रतिनिधित्व करता है। ये उपाय इंगित करते हैं कि वितरण में वितरण के अधिकांश मूल्य कहाँ होंगे और इन्हें वितरण के केंद्रीय स्थान के रूप में भी जाना जाता है। आप इसे एक मध्य मूल्य के आसपास क्लस्टर करने के लिए डेटा की प्रवृत्ति के रूप में सोच सकते हैं। आंकड़ों में, केंद्रीय प्रवृत्ति के लिए तीन मापन का उपयोग करते

हैं, जिन्हें माध्य, मध्यिका और बहुलक या मोड कहते हैं। इनमें से प्रत्येक मापन को अलग अलग परिस्थिति में किया जाता है, जिसका आकलन करने की विधि भी अलग अलग है। मुख्य उद्देश्य केंद्रीय बिंदु के स्थान की गणना करता है। जिसका हम क्रमशः संक्षेप में अध्ययन करेंगे –:

Mean मध्यमान या माध्य सामान्यतः इसे समान्तर माध्य के रूप में भी जानते हैं, यह केन्द्रिय प्रवृत्ति की माप में सर्वाधिक उपयोग किया जाने वाला व विश्वनीय मान के रूप में जाना जाता है जिसे ज्ञात करने के लिये प्राप्त सभी आँकड़ों के मान का योगफल निकालकर उने आँकड़ों की संख्या से भाग देने पर प्राप्त मान या मूल्य जो कि औसत के रूप में जाना जाता है यहाँ उसे माध्य के रूप में जाना जाता है।

Mean = Sum of All Data Points / Number of Data Points

यदि आकड़ों की संख्या ज्यादा हो और ये व्यवस्थिति रूप से तालिका में हो तो हम इसका उपयोग करते हैं

Mean = Assumed Mean + (Sum of All Deviations / Number of Data Points)

मध्यमान के गुण Merits of Mean

1. मध्यमान का मूल्य हमेशा निश्चित तथा स्थिर रहता है, इसलिए शोधकर्ता एक समान मान को ही प्राप्त करता है।
2. इसका मान समस्त में से प्रत्येक मापनों या आकड़ों पर निर्भर करती है, यदि एक का भी मान बदल जावे तो मूल्य में अन्तर आ जायेगा।
3. मध्यमान को आसानी से समझा व गणना कर प्राप्त किया जा सकता है।
4. मध्यमान के सभी प्राप्तांकों के विचलन का योग शून्य होता है। इसप्रकार मध्यमान एक सन्तुलित बिन्दु होता है।
5. मध्यमान में एक विशेषता यह पायी जाती है कि वह अग्रिम बीजगणितीय विवेचन के योग्य होता है। दूसरे शब्दों में इसमें बीजगणित की प्रक्रियाओं का प्रयोग किया जा सकता है।
6. मध्यमान के मूल्य पर प्रतिदर्श के परिवर्तनों का प्रभाव नहीं पड़ता है।
7. मध्यमान से लिये गये विचलनों के वर्ग का योग किसी अन्य मूल्य से लिये गये विचलनों के वर्गों के योग से कम होता है।
8. मध्यमान की एक अन्य विशेषता यह भी है कि यदि किसी अंकिक श्रृंखला के अंको को उस श्रृंखला के मध्यमान से स्थानापन्न कर दिया जावे तो नवीन श्रृंखलाओं का योग मौलिक श्रृंखला के बराबर होता है।

मध्यमान के दोष Demerits of Mean

मध्यमान की जहाँ अनेक विशेषतायें हैं वहीं इसकी गणना विधि के कारण कुछ दोष भी हैं।

1. इसकी गणना सभी प्राप्तांको के आधार पर की जाती है, अतः कभी-कभी असामान्य तथा सीमान्त पद इसके मान को प्रभावित करते हैं।
2. मध्यमान की गणना सभी प्राप्तांको के आधार पर की जाती है यदि श्रृंखला का एक मान प्राप्त नहीं होगा तो इसकी गणना नहीं की जा सकती।
3. मध्यमान कभी कभी इसप्रकार के मूल्य भी देता है जो कि अवास्तविक प्रतीत होता है।
4. मध्यमान का अनुमान बहुलक की तरह निरीक्षण मात्र से संभव नहीं है।
5. अंको के पूर्ण विवरण की अनुपस्थिति में मध्यमान भ्रमपूर्ण निष्कर्ष भी प्रदान कर सकते हैं।

मध्यमान के उपयोग Use of Mean

1. जब वितरण सामान्य हों, तो शोध में इसका आकलन बिलकुल सही मान देता है, क्योंकि इसके केन्द्र में यह होता है और सभी प्राप्तकों का इसमें योगदान होता है।
2. सर्वाधिक स्थिर और विश्वनीय मान की गणना करना हो तो मध्यमान से ज्यादा उपयुक्त कोई केन्द्रीय प्रवृत्ति नहीं है।
3. जब अन्य सांख्यिकीय मानों की जरूरत होती है, जैसे प्रमाप विचलन, सह-संबंध आदि तो उसके लिए मध्यमान का प्रयोग करना ही पड़ता है।

Median ;या माध्यिका

माध्यिका या मेडियन केन्द्रीय प्रवृत्ति का वह माप है जो डेटा में मध्य पर स्थित –सबसे अधिक अवलोकन का मूल्य देता है। अर्थात् मध्यांक मान स्थिति का माध्य होता है। स्थिति का तात्पर्य किसी स्थान के अंकिक मूल्य से होता है। मध्यांक की स्थिति किसी कमबद्ध श्रृंखला से इस प्रकार की होती है कि इसके उपर तथा नीचे समान संख्या में आकड़े होते हैं। अनियंत्रित डेटा के मामले में, हम पहले आरोही क्रम में टिप्पणियों के डेटा मूल्यों को व्यवस्थित करते हैं, जबकि व्यवस्थित आकड़ों के मामलों जबकि आकड़े एक तालिका में हो अलग तरीके से गणना की जाती है।

इसे ज्ञात करने के लिये ऑकड़ों को अवरोही या आरोही क्रम में रखकर यदि आकड़ों की संख्या सम है तो $n/2$ वाँ ऑकड़ा या पद मध्याक होगी परन्तु यदि ऑकड़ों की संख्या विषम है तो उसमें जोड़कर $n+1/2$ वाँ पद को मध्याक माना जाता है

मध्यांक मान के गुण Merits of Median

1. मध्यांक मान का मूल्य सदैव निश्चित रहता है। अर्थात् श्रृंखला के किनारे के मान में परिवर्तन का कोई प्रभाव मध्यांक पर नहीं होता।
 2. मध्यांक मान की गणना करना व समझना बहुत आसान है।
 3. इसके मूल्य पर चरम या सीमान्त मूल्यों का प्रभाव नहीं पड़ता है, क्योंकि यह स्थिति का मापन होता है।
 4. मध्यांक मान की गणना उस समय भी की जा सकती है जबकि सीमान्त मूल्यों की जानकारी न हो। अनियमित वर्गाकार तथा खुली सीमा वाले अंक वितरण हो।
 5. कई बार इसकी गणना निरीक्षण मात्र से ही हो जाती है।
- जैसे— किसी कक्षा के छात्रों को बढ़ती हुई उचाई के क्रम में खड़े करवाकर हम बीच के एक लड़के की उचाई को उस कक्षा के लिए उचाई मान सकते हैं।
6. विशेष प्रकार के तथ्यों के मापन में केवल मध्यांक मान की ही गणना की जा सकती है, उन तथ्यों के अध्ययन में यह अच्छे परिणाम देता है, जिनका हम प्रत्यक्ष रूप से संख्यात्मक मापन नहीं कर सकते जैसे ईमानदारी, सुन्दरता, रंग आदि।

मध्यांक मान के दोष Demerits of Median

1. कई उदाहरणों में मध्यांक श्रृंखला का प्रतिनिधित्व ही नहीं करता।
2. इसका बीजगणितय विवेचना संभव ही नहीं है।

3. यदि एक श्रृंखला के बड़े बड़े और छोटे-छोटे मूल्यों को महत्व दिया जाना है तो मध्यांक मापन अनुचित होगा क्योंकि यह सीमान्त मानों को अपने में शामिल ही नहीं करता ।
4. संख्यात्मक मानों को कभी कभी कमबध्द करना कठिन कार्य होता है ।
5. अंक श्रृंखला में थोड़ा सा भी परिवर्तन मध्यांक में बड़ा परिवर्तन पैदा कर देता है ।
6. व्यवस्थित प्राप्ताकों में से मध्यांक मान की गणना करने पर यह कल्पना की जाती है कि वर्गान्तर की सभी आवृतियों इसके मूल्यों में समान रूप से वितरित होती है ।

मध्यांक मान के उपयोग Use of Median

1. जब वितरण का वास्तविक बिन्दु ज्ञात करना हो ।
2. वितरण क विषम अथवा असामान्य होने की स्थिति में केन्द्रीय प्रवृत्ति की गणना करनी हो तो मध्यांक का मापन ही उपयोगी होता है ।
3. जब इस बात की संभावना हो कि एक निश्चित प्राप्तांक केन्द्रीय प्रवृत्ति को प्रभावित कर सकता है परन्तु उसके विषय में केवल यह ज्ञात हा कि वह मध्यांक मान से उपर या नीचे है ।
4. जिन विशेषताओं की गणना सम्भव नहीं है वरन केवल कम दिया जा सकता है । वहाँ पर भी केन्द्रीय मूल्य के रूप में मध्यांक मान को गणना की जानी चाहिए ।

Median Formula

The formula to calculate the median of the data set is given as follows:

If the total number of observation given is odd, then the formula to calculate the median is:

$$\text{Median} = \{(n+1)/2\}^{\text{th}} \text{term}$$

If the total number of observation is even, then the median formula is:

$$\text{Median} = [(n/2)^{\text{th}} \text{ term} + \{(n/2)+1\}^{\text{th}}]/2$$

Mode बहुलक

किसी वितरण में सर्वाधिक बार दिखाई देने वाला आकड़ा ही बहुलक कहलाता है, अर्थात् किसी वितरण का सर्वाधिक आवृत्ति वाला पद बहुलक या मोड कहलाता है ।

इस प्रकार डेटा सेट का मोड वह संख्या है जो सेट में सबसे अधिक बार होती है। मोड को आसानी से खोजने के लिए, संख्याओं को सबसे कम से कम क्रम में रखें और गिनें कि प्रत्येक संख्या कितनी बार होती है। सबसे अधिक होने वाली संख्या मोड है! परन्तु आकड़े ग्रुप डाटा के रूप में हो तो मोड की गणना कैसे करेंगे आइये इसको समझें –:

Mode is that value among the observations which occurs most often, that is, the value of the observation having the maximum frequency.

In a grouped frequency distribution, it is not possible to determine the mode by looking at the frequencies. Here, we can only locate a class with the maximum frequency, called the **modal class**.

E.g. Looking at data below, we can say that maximum occurrence occur at class 60-80, frequency 61. But we can't tell the most frequent data (mode).

Lifetime (hrs)	0-20	20-40	40-60	60-80	80-100	100-120
Frequency	10	35	52	61	38	29

We can use this formula to find the mode for Grouped data. In the example above modal class is 60-80.

$$\text{Mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right) \times h$$

- where l = lower limit of the modal class,
- h = size of the class interval (assuming all class sizes to be equal),
- f_1 = frequency of the modal class,

- f_0 = frequency of the class preceding the modal class,
- f_2 = frequency of the class succeeding the modal class.

ExamFear.com

$h = 80 - 60 = 20$

Lifetime (hrs)	0-20	20-40	40-60	60-80	80-100	100-120
Frequency	10	35	52	61	38	29

ExamFear.com

f_0 f_1 f_2

Modal Class $l = 60$

For the example above, modal class is 60-80 with frequency 61

Therefore, $l=60$, $h = 80-60 = 20$, $f_1 = 61$, $f_0 = 52$, $f_2 = 38$

$$\text{Mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right) \times h$$

ExamFear.com

$$\text{Mode} = 60 + (61-52)/(2*61-52-38) * 20$$

$$\text{Or Mode} = 65.625$$

एक सरल उदाहरण से हम उपरोक्त सभी को समझने का प्रयास करते हैं :-

- **Find the mean, median, mode, and range for the following list of values:**

13, 18, 13, 14, 13, 16, 14, 21, 13

The mean is the usual average, so I'll add and then divide:

$$(13 + 18 + 13 + 14 + 13 + 16 + 14 + 21 + 13) \div 9 = 15$$

Note that the mean, in this case, isn't a value from the original list. This is a common result. You should not assume that your mean will be one of your original numbers.

The median is the middle value, so first I'll have to rewrite the list in numerical order:

13, 13, 13, 13, 14, 14, 16, 18, 21

There are nine numbers in the list, so the middle one will be the $(9 + 1) \div 2 = 10 \div 2 = 5$ th number:

13, 13, 13, 13, 14, 14, 16, 18, 21

So the median is 14.

The mode is the number that is repeated more often than any other, so 13 is the mode.

The largest value in the list is 21, and the smallest is 13, so the range is $21 - 13 = 8$.

mean: 15

median: 14

mode: 13

range: 8

Note: The formula for the place to find the median is " $([the\ number\ of\ data\ points] + 1) \div 2$ ", but you don't have to use this formula. You can just count in from both ends of the list until you meet in the middle, if you prefer, especially if your list is short. Either way will work.

Second Example

Alex timed 21 people in the sprint race, to the nearest second:

59, 65, 61, 62, 53, 55, 60, 70, 64, 56, 58, 58, 62, 62, 68, 65, 56,
59, 68, 61, 67

To find the [Mean](#) Alex adds up all the numbers, then divides by how many numbers:

$$\begin{aligned} \text{Mean} &= 59 + 65 + 61 + 62 + 53 + 55 + 60 + 70 + 64 + 56 + 58 + \\ &58 + 62 + 62 + 68 + 65 + 56 + 59 + 68 + 61 + 67 \\ &= 61.38095\dots \end{aligned}$$

To find the [Median](#) Alex places the numbers in value order and finds the middle number.



In this case the median is the 11th number:

53, 55, 56, 56, 58, 58, 59, 59, 60, 61, 61, 62, 62, 62, 64, 65, 65,
67, 68, 68, 70

$$\text{Median} = 61$$

To find the [Mode](#), or modal value, Alex places the numbers in value order then counts how many of each number. The Mode is the number which appears most often (there can be more than one mode):

53, 55, 56, 56, 58, 58, 59, 59, 60, 61, 61, 62, 62, 62, 64, 65, 65,
67, 68, 68, 70

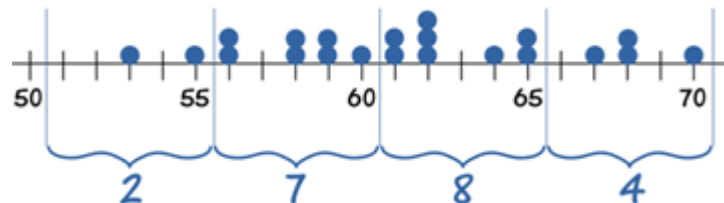
62 appears three times, more often than the other values,
so **Mode = 62**

Grouped Frequency Table

Alex then makes a [Grouped Frequency Table](#):

Seconds Frequency

51 - 55	2
56 - 60	7
61 - 65	8
66 - 70	4



So 2 runners took between 51 and 55 seconds, 7 took between 56 and 60 seconds, etc

Only the Grouped Frequency Table survived ...

... can we help Alex calculate the Mean, Median and Mode from just that table?

The answer is ... no we can't. Not accurately anyway. But, we can make **estimates**.

Estimating the Mean from Grouped Data

So all we have left is:

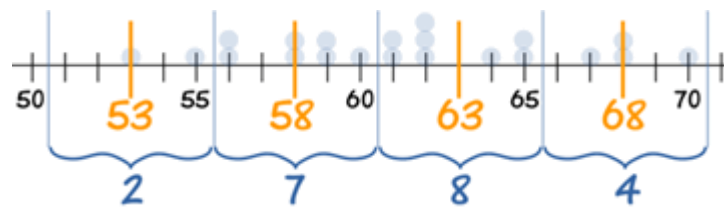
Seconds Frequency

51 - 55	2
---------	---

56 - 60	7
61 - 65	8
66 - 70	4

The groups (51-55, 56-60, etc), also called **class intervals**, are of **width 5**

The **midpoints** are in the middle of each class: 53, 58, 63 and 68



We can estimate the **Mean** by using the **midpoints**.

So, how does this work?

Think about the 7 runners in the group **56 - 60**: all we know is that they ran somewhere between 56 and 60 seconds:

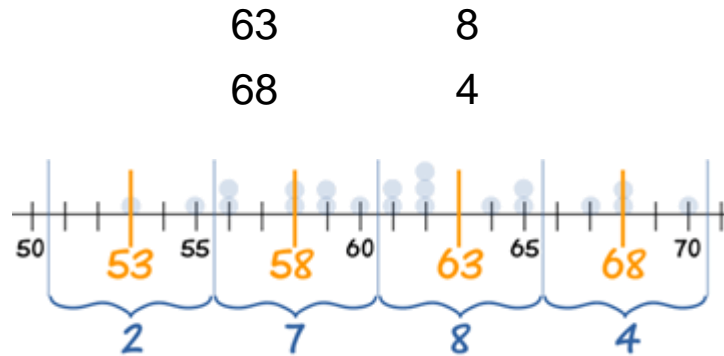
- Maybe all seven of them did 56 seconds,
- Maybe all seven of them did 60 seconds,
- But it is more likely that there is a spread of numbers: some at 56, some at 57, etc

So we take an average and **assume** that all seven of them took 58 seconds.

Let's now make the table using midpoints:

Midpoint Frequency

53	2
58	7



Our thinking is: "2 people took 53 sec, 7 people took 58 sec, 8 people took 63 sec and 4 took 68 sec". In other words we **imagine** the data looks like this:

53, 53, 58, 58, 58, 58, 58, 58, 58, 63, 63, 63, 63, 63, 63, 63, 63, 68, 68, 68, 68

Then we add them all up and divide by 21. The quick way to do it is to multiply each midpoint by each frequency:

Midpoint x	Frequency f	Midpoint x Frequency fx
53	2	106
58	7	406
63	8	504
68	4	272
Totals:	21	1288

And then our **estimate** of the mean time to complete the race is:

$$\text{Estimated Mean} = \frac{1288}{21} = 61.333\dots$$

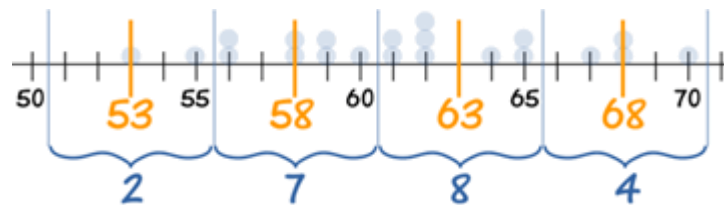
Very close to the exact answer we got earlier.

Estimating the Median from Grouped Data

Let's look at our data again:

Seconds Frequency

51 - 55	2
56 - 60	7
61 - 65	8
66 - 70	4



The median is the middle value, which in our case is the 11th one, which is in the 61 - 65 group:

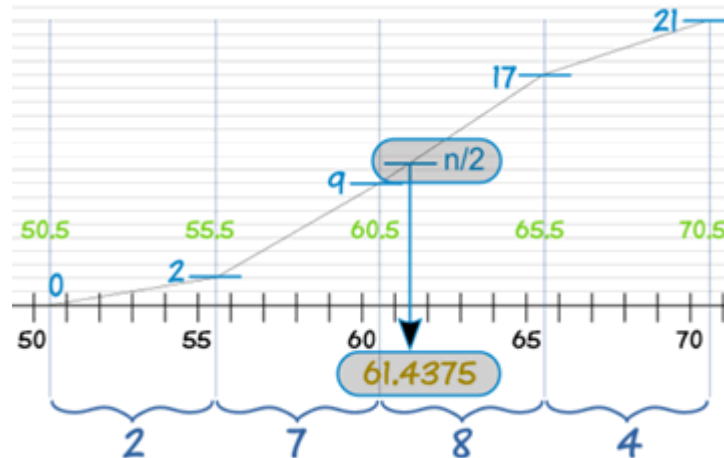
We can say "the **median group** is 61 - 65"

But if we want an estimated **Median value** we need to look more closely at the 61 - 65 group.

We call it "61 - 65", but it really includes values from 60.5 up to (but not including) 65.5.

Why? Well, the values are in whole seconds, so a real time of 60.5 is measured as 61. Likewise 65.4 is measured as 65.

At 60.5 we already have **9** runners, and by the next boundary at 65.5 we have **17** runners. By drawing a straight line in between we can pick out where the median frequency of **$n/2$** runners is:



And this handy formula does the calculation:

$$\text{Estimated Median} = L + \frac{(n/2) - BG}{G} \times w$$

where:

- **L** is the lower class boundary of the group containing the median
- **n** is the total number of values
- **B** is the cumulative frequency of the groups before the median group
- **G** is the frequency of the median group
- **w** is the group width

For our example:

- **L** = 60.5
- **n** = 21
- **B** = 2 + 7 = 9
- **G** = 8
- **w** = 5

$$\begin{aligned} \text{Estimated Median} &= 60.5 + \frac{(21/2) - 9}{8} \times 5 \\ &= 60.5 + 0.9375 \\ &= 61.4375 \end{aligned}$$

Estimating the Mode from Grouped Data

Again, looking at our data:

Seconds	Frequency
51 - 55	2
56 - 60	7
61 - 65	8
66 - 70	4

We can easily find the modal group (the group with the highest frequency), which is **61 - 65**

We can say "the **modal group** is 61 - 65"

But the actual **Mode** may not even be in that group! Or there may be more than one mode. Without the raw data we don't really know.

But, we can **estimate** the Mode using the following formula:

$$\text{Estimated Mode} = L + \frac{f_m - f_{m-1}}{f_m - f_{m-1} + (f_m - f_{m+1})} \times w$$

where:

- L is the lower class boundary of the modal group
- f_{m-1} is the frequency of the group before the modal group
- f_m is the frequency of the modal group
- f_{m+1} is the frequency of the group after the modal group
- w is the group width

In this example:

- $L = 60.5$
- $f_{m-1} = 7$
- $f_m = 8$

- $f_{m+1} = 4$
- $w = 5$

$$\begin{aligned}\text{Estimated Mode} &= 60.5 + 8 - 7(8 - 7) + (8 - 4) \times 5 \\ &= 60.5 + (1/5) \times 5 \\ &= \mathbf{61.5}\end{aligned}$$

Our final result is:

- Estimated Mean: **61.333...**
- Estimated Median: **61.4375**
- Estimated Mode: **61.5**

(Compare that with the true Mean, Median and Mode of **61.38...**, **61** and **62** that we got at the very start.)

विषयवस्तु की अधिक जानकारी के लिए निम्नलिखित लिंक का अवलोकन करें—:

<https://www.youtube.com/watch?v=9PHKe14tk20>

<https://www.khanacademy.org/math/pre-algebra/pre-algebra-math-reasoning/pre-algebra-representing-data/v/ways-to-represent-data>

<https://www.youtube.com/watch?v=yeh8ndRIA>

http://mpbou.edu.in/slm/B.Ed_SLM/bed02_07_b4_unit-2.pdf

<https://www.youtube.com/watch?v=wrGYheq0whk>

<https://www.youtube.com/watch?v=Urz4NGkXwis>

<https://people.umass.edu/biep540w/pdf/Grouped%20Data%20Calculation.pdf>

Probability प्रायिकता अथवा संभावना

संभाव्यता गणित की वह शाखा है जिसमें संख्यात्मक विवरणों के बारे में बताया जाता है कि कोई घटना कैसे घटित होती है या कितनी संभावना है कि प्रस्ताव सत्य है। एक घटना की संभावना 0 और 1 के बीच की एक संख्या है, जहां, मोटे तौर पर बोलना, 0 घटना की असंभवता को दर्शाता है और 1 घटना को इंगित करता है। इसे 0 प्रतिशत व 100 प्रतिशत के बीच की संभावना के रूप में भी देख सकते हैं। इसे समझने के लिये एक उदाहरण लेते हैं यदि एक सिक्का उछाला जाता है तो सिक्के के दो पहलू होते हैं जिसे हम हेड और टेल के नाम से जानते हैं इसमें संभावना 50–50 प्रतिशत होगा कि वह हेड या टेल हो अर्थात् एक सिक्के के उछालने पर संभावित परिणाम 2 हो सकते हैं परन्तु हमारी अपेक्षा या तो हेड या टेल की होगी इसप्रकार केवल 1 ही होगी इसलिये प्रायिकता का नाम सरल घटना Simple Event कहते हैं, परन्तु यहीं पर एक साथ कई सिक्के उछालें जाएं तो यहाँ स्थिति बदलती है बल्कि कई बार एक घटना दूसरे का प्रभावित भी कर सकती है उस स्थिति में वह पहली घटना पर आश्रित हो जाती है जैसे ताश की गड्डी से कोई एक पत्ता निकालना हो तो उसे बिना गड्डी में रखे पत्ते निकालना प्रभावित करेगा क्योंकि पहले निकाला गया पत्ता वापस रखा नहीं गया है, उसकी निकलने की संभावना समाप्त हो चुकी है। इस प्रकार प्रायिकता के लिए सूत्र भी इन स्थितियों में अलग-अलग होंगे ।

इस प्रकार सरल शब्दों में समान रूप से सम्भाव्य अनेक घटनाओं में किसी एक घटना के घटित होने की प्रायिकता वांछित घटना के अनुकूल परिस्थितियों की संख्या के साथ कुल सम्भाव्य घटनाओं की संख्या का अनुपात होता है

$$\text{प्रायिकता} = \frac{\text{वांछित घटनाओं की संख्या}}{\text{कुल घटनाओं की संख्या}}$$

यदि किसी घटना के होने तथा न होने के सभी ढंग सम-प्रायिक (Equally Likely) हों अर्थात् एक दूसरे से संबंधित होते हैं घटना यदि m बार हो रही है, और n बार नहीं हो रही है तो घटना की प्रत्याशा $m+n$ होगी इसे इस प्रकार भी कह सकते हैं कि उस घटना के घटित होने की प्रायिकता होगी $m/m+n$ उस घटना के नही घटित होने की प्रायिकता $n/m+n$ होगी यहाँ यदि हम घटना के घटित होने की प्रायिकता को p और नही घटित होने की प्रायिकता को q से प्रदर्शित करें तो समप्रायिक घटना के लिए

घटित होने की प्रायिकता , $p = m/m+n$

घटित न होने की प्रायिकता, $q = n/m+n$

स्पष्ट है कि p और q का योग हमेशा 1 होगा अतः p और q में से किसी एक के ज्ञात होने से दूसरे की गणना की जा सकती है । प्रायिकता के कुछ नियम भी है, सामान्य प्रायिकता में दो नियमों का पालन होता है

Two Rules of Probability

1-The sum rule shows that the probability of one of two mutually exclusive events, call them A and B, occurring is equal to the the sum of the two individual events' probabilities. This is depicted mathematically as:

$$P(A \cup B) = P(A) + P(B)$$

2-The product rule addresses two independent events (meaning that each does not affect the outcome of the other) that happen together, such as considering the probability that your offspring will have dimples and be male. The probability that the events will occur together can be calculated by multiplying the probabilities of each individual event:

$$P(A \cup B) = P(A) \times P(B)$$

If you were to roll a die twice, the formula to calculate the probability that you roll a 4 the first time and a 1 the second time would look like this:

$$P(A \cup B) = P(\text{rolling a 4}) \times P(\text{rolling a 1}) = (1/6) \times (1/6) = 1/36$$

हम इसे दो उदाहरण से और अधिक स्पष्ट करने का प्रयास करते हैं—:

उदाहरण 1

मान लीजिए कि एक एक डाइस उछाला गया हमें मालूम है यह इसमें 1,2,3,4,5,6, डाट विभिन्न छै फलक में बने होते हैं तो क्या संभावना है कि 5 डाट वाला फलक हीउपर आये ?

हल—: जब पॉसा उपर उछाला जाता है तो कोई एक फलक ही उपर आयेगा इसका मतलब 1,2,3,4,5,6 में से कोई एक ही उपर दिखाई देगा। इसका अर्थ है 6 संभावित परिणाम होंगे उसमें से 5 भी है जो उपर आ सकता है।

$$S = \{ 1, 2, 3, 4, 5, 6 \}.$$

प्रत्येक संभावित परिणाम एक यादृच्छिक चर (X) है, और प्रत्येक परिणाम समान रूप से होने की संभावना है। इस प्रकार, हमारे पास एक समान वितरण है। इसलिए

$$P(X = 5) = 1/6.$$

$$P(X < 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$P(X < 5) = 1/6 + 1/6 + 1/6 + 1/6 = 2/3$$

उदाहरण 2—: यदि उपरोक्त उदाहरण में वर्णित पासा फेंकने के प्रयोग को दोहराते हैं। इस बार, हम पूछते हैं कि क्या संभावना है कि डाइस में डाट की संख्या 5 से छोटी होगी?

हल—: इसका अर्थ है पॉसा उछालने में स्थिति तो वैसी ही होगी मतलब $S = \{ 1, 2, 3, 4, 5, 6 \}$. तो 6 संभावित परिणाम

प्रस्तुत होते हैं, परन्तु पाँच से छोटी संख्या मतलब 1,2,3,4, आने की संभावना अलग अलग बनती है,तो प्रायिकता सभी का योगफल होगी।

$$P(X < 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$P(X < 5) = 1/6 + 1/6 + 1/6 + 1/6 = 2/3$$

प्रायिकता की उपयोगिता

प्रायिकता की उपयोगिता , शिक्षा अनुसंधान के साथ मौसम विज्ञान की भविष्यवाणी में भी होता है। बाजार में भी व्यवसायी इस बात का अनुमान लगाने के लिए कि किस प्रकार की संभावना आगे होगी करते हैं? खेल में तो इसका सबसे ज्यादा उपयोग होता है ,विशेष कर क्रिकेट का खेल तो पूरी तरह से संभावना पर आधारित होता है कहा जाता कि क्रिकेट इज द चांस ।

<https://www.youtube.com/watch?v=tkLBIgFN6mM>

https://www.coconino.edu/resources/files/pdfs/academics/sabbatical-reports/kate-kozak/chapter_5.pdf