SEMESTER – III

STATISTICAL METHODS IN GEOGRAPHY

UNIT : I

TOPIC: THEOROTICAL DISTRIBUTION

PRESENTED BY

DR. SIPRA BISWAS

What or Why Frequency Distribution

Frequency Distribution is a statistical table that shows the values of the variables and the corresponding number of observations or frequencies.

The collected data are required to be organised and summarized for carrying out a meaningful statistical analysis before any study .

Continued…

Frequency distribution is generally used to:

i. **Condense or summarise a huge body of data,**
ii. **Visualize the range, average, spread and shape of the distribution and**
iii. **Approximate the nature of probability distribution in the population.**

Types

Frequency distribution may be of two types:

1. **Simple or Descrete or ungrouped frequency distribution**
2. **Grouped or Continuous frequency distribution**

Simple frequency distribution: In simple frequency distribution, the values of the variables are arranged individually in order of magnitude.

This arrangement of raw data is called ‘array’.

Simple frequency distribution

Such as

33,40,30,36,40,40,30,33,50,36,50,30,40,38,36,33,40,50,38,38,40,33

Data	Tally Mark	Frequency
30	I I I	3
33	I I I I	4
36	I I I	3
38	I I I	3
40	~~I I I I~~ I	6
50	I I I	3
Total		22

Grouped frequency distribution

In grouped frequency distribution, the values of variables are arranged in groups in order of magnitude.

To do this values are chosen at equal intervals as limits of successive classes within the observed range.

The number of individual values falling within each class is counted and termed the frequency of that class.

The process in which the frequencies are distributed over the classes is called frequency distribution.

Grouped frequency distribution

Such as:

5,11,16,21,26,7,12,17,22,27,7,13,17,23,28,8,14,24,29,9,15,19,25,30,5,11,21,16,26,12,7,17,22,8,12,13

Class	Tally Mark	Frequency
5-10	𝍸 I I I	8
10-15	𝍸 I I I I	9
15-20	I 𝍸 I	6
20-25	𝍸 I I I	7
25-30	𝍸 I I	6
		N=36

Construction of grouped distribution

Total number of classes:

$n=1+3.3*\log N$

where, n=Number of classes to be selected

and N=The total frequency

The size of classes or the class width (w):

$w=R/n$

Where R=Range of data i.e., the difference between the highest and lowest values of the series and

n=Number of classes

Continued….

* **It should be noted that the value of the range is approximated to its nearest round figure.**
* **Another important point is that in some special cases there may occur a no. of empty classes. This is because the spread of data is much larger and there is marked concentration of values at certain levels. The distribution is very skewed and the numerical differences between two values far exceeds the class width chosen. For such a series of data, usually unequal classes are taken.**

Cumulative frequency distribution

- Sometimes, it maybe useful to know how many or what proportion of districts less than or more than certain specified population density levels exist.
- For this, frequencies below or above a certain class boundary need to be calculated.
- These accumulated frequencies above or below a certain class boundary are known as cumulative frequencies.
- The cumulative frequencies below or above the class boundaries are arranged in ascending or descending order.
- Sometimes instead of absolute frequencies, percentage of frequencies are used in cumulative frequency distribution, that is called Cumulative Percentage Distribution or Relative Cumulative Frequency Distribution.

Cumulative Frequency Distribution Table

Class Limit	Class Boundary	Class width	Class Mark (x)	Frequency	Cumulative Frequency	
					Less Than	More Than
101-200	100.5-200.5	100	150.5	4	4	92
201-300	200.5-300.5	100	250.5	12	16	80
301-400	300.5-400.5	100	350.5	24	40	56
401-500	400.5-500.5	100	450.5	40	80	16
501-600	500.5-600.5	100	550.5	16	96	0
				N=96		