Frequency Distributions

Distinguish carefully between a variable(x) and its frequency(f)

Table A

|  |  |
| --- | --- |
| Number of Students with this Mark(f) | Mark(x) |
| 1 | 0 |
| 2 | 1 |
| 2 | 2 |
| 0 | 3 |
| 5 | 4 |
| 6 | 5 |
| 8 | 6 |
| 5 | 7 |
| 3 | 8 |
| 1 | 9 |
| 2 | 10 |

This is a frequency Table. The Mark X is the Variable and the number f of students each mark is the frequency of that mark. The sum of the frequencies is the total number of students. We shall use the symbol ∑ to mean ‘sum of’ so ∑f=35

Values of the variable are always marked along the horizontal axis, values of the frequencies along the vertical axis. Note: bars in a histogram are always vertical and are always placed together, with no gaps (unless a frequency is zero and each value of the variable is placed at the center of the corresponding bar. Show figures graphically as a histogram.

Grouped frequency distributions

If the variable can take large number of values group the variables into classes (grouped frequency distribution)

Marks of 50 students in an examination is 11 21 33 26 30 45 32 26 18 3 18 25 28 24 11 38 23 20 13 40 18 12 8 42 21 10 27 24 21 13 9 28 4834 30 28 23 6 1534 26 25 19 16 26 29 35 24 27 32. Group the marks into classes, each class containing an equal spread of marks (class interval). Between 6 and 12 is usually satisfactory. The lowest mark is 3 and the highest is 48. The Range of marks is 45. choose a class interval of marks. The first one being the interval 3-48 marks(inclusive).3 and 48 are the class limits. Make a group frequency table? The first column showing their marks for each class. There are 8 classes. Go through the marks in the list and put a / in the tally column locating each mark in its class. Do this carefully so that no mark is missed or done twice. The 5th tally mark rosses out the previous four to help in counting up at the end. Complete the frequency column.

Grouped Frequency table

|  |  |  |
| --- | --- | --- |
| Marks for each class | Tally | Frequency |
| 3-8 |  | 3 |
| 9-14 |  | 7 |
| 15-20 |  | 7 |
| 21-26 |  | 14 |
| 27-32 |  | 10 |
| 33-38 |  | 5 |
| 39-44 |  | 2 |
| 45-50 |  | 2 |
| Total |  | 50 |

∑f=50

Show the histogram for this distribution with class limits marked? The class intervals are of equal width, the area of each bar in the histogram is propositional to the frequency represented. The total area under a histogram is therefore propositional to the total frequency. Show frequency curve and polygon. Analyze the data?

Unequal Class Intervals

The salary distribution among three grades of employee in a large company are given in Table B. Show these figures on a histogram.

Table B

|  |  |
| --- | --- |
| Salary In 1000 $ | Frequency |
| 10-12 | 8 |
| 12-20 | 16 |
| 20-40 | 10 |

Make of width 2 units and height 8 units. The area of this bar =16 sg. In drawing histograms for distributions with unequal class intervals. Remember that the areas of the bars must be propositional to the frequencies. Note difference between a bar chart and a histogram.

Discrete and continuous data

The weight of 60 schoolchildren were measured to the nearest kg. The frequencies are shown in Table Show these figures on a histogram.

Table C

|  |  |
| --- | --- |
| Weights(kg) | f |
| 25-27 | 2 |
| 28-30 | 8 |
| 31-33 | 16 |
| 34-36 | 20 |
| 37-39 | 10 |
| 40-42 | 4 |

∑f=60

Note: that 24.5 and 27.5 are the class boundaries. show on histogram. Mark the class centers on the variable scale. The class center for the first class is 25+27/2=26kg and so on. Find the mean?

Draw a histogram to illustrate the data in Table Which gives the heights of 100 children to the nearest cm.(a) From your histogram estimate the number of children whose height lies between 142cm and 152cm.(b) If a child is chosen at random from this group, what is the probability that its height was not greater than 156cm?

Table E

|  |  |
| --- | --- |
| Height(cm) | f |
| 120-129 | 6 |
| 130-139 | 15 |
| 140-149 | 31 |
| 150-159 | 37 |
| 160-169 | 9 |
| 170-179 | 2 |

∑f=100

Note: the class boundaries are 119.5,129. 5…….,179. 5cm.Area under the histogram are proportional to the appropriate frequencies. To find the number of children between 142cm and 152cm in height(from 141.5cm to 152.5cm inclusive) we find the shaded area between these lines(shown dotted on the figure).This area = 0.8×31+0.3x37=36.A height not greater than 156cm means a height less than 156.5cm.The number of children with a height less than 156.5cm is given by the area to the left of the dotted line at 156.5.This number = 6+15+31+(0.7x37)=78.Hence the required probability = 78/100 = 0.78