Data Handling

Exemplar 6

Cumulative Frequency Polygon

Objectives   (1)   To construct a cumulative frequency polygon from a set of data

(2)   To interpret a cumulative frequency polygon

Learning Unit   Construction and Interpretation of Simple Diagrams and Graphs

Key Stage   3

Materials Required   *Excel* and the file dh06_e.xls

Prerequisite Knowledge   (1)   Frequency distribution table

(2)   Frequency polygon

Description of the Activity

1.   The teacher distributes the worksheet and the file dh06_e.xls in a diskette to students.

2.   The teacher briefly explains the problem to students and asks them to open the file and complete the Frequency Distribution Table (Table 1) on the spreadsheet. See the figure below.

3. The teacher asks students to do the following investigations.
    (a) How many students have marks less than 30.5?
    (b) How many students have marks less than 40.5?
    (c) How many students have marks less than 50.5?
    (d) Can you observe any relationship between the frequencies in the Frequency Distribution Table (Table 1) and the answers obtained above?

    The teacher introduces the meaning of cumulative frequency.

4. The built-in function of addition in *Excel* is used to complete the column of cumulative frequency in Table 2. The term "cumulative frequency table" is then introduced.

5. Students are guided to plot the cumulative frequency polygon by using the software. See the figure below.
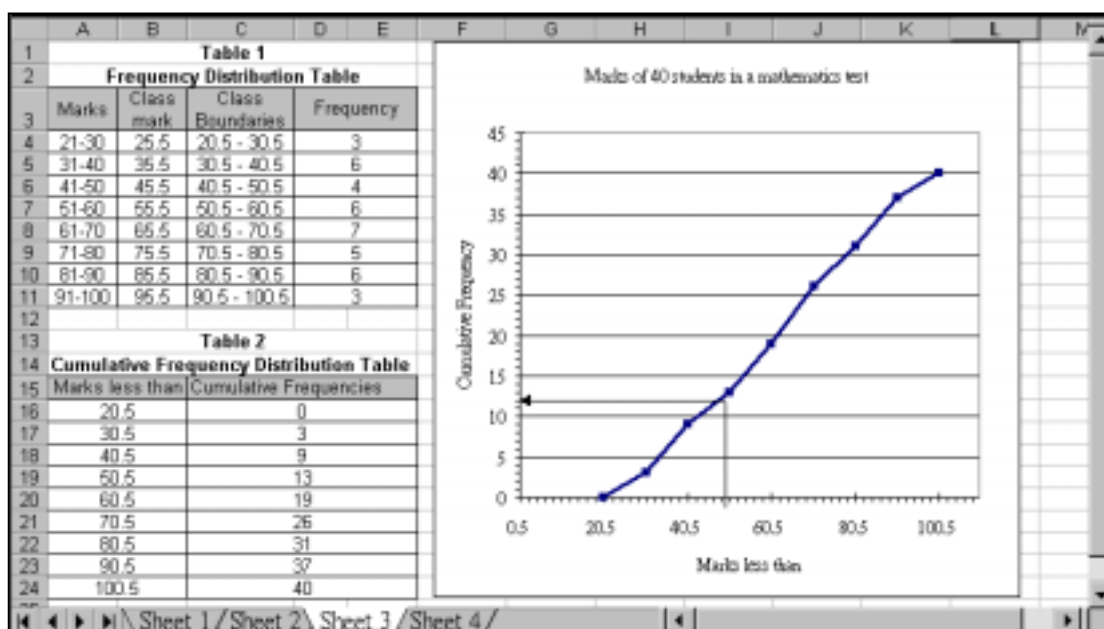
6. Students are divided into groups and are asked to discuss the remaining problems in the worksheet.

7. Each group is invited to present their answers and methods used. The teacher can make comments when appropriate.

Worksheet: Cumulative Frequency Polygon

Problem :

The following marks are obtained by 40 students of S1A in a Mathematics test.

77	81	74	56	63	52	87	90	34	29
57	68	29	34	98	58	43	51	74	64
68	39	45	83	62	94	36	61	88	89
38	54	46	73	67	31	27	45	99	79

1. In the *Excel* file dh06_e.xls, a Frequency Distribution Table (Table 1) with eight class intervals of equal lengths is provided. Complete the table by using the marks provided with 21 – 30 as the lowest class interval.

2. Answer the following questions by using Table 1.
    (a) How many students have marks less than 30.5?　　　＿＿＿＿＿＿＿＿＿
    (b) How many students have marks less than 40.5?　　　＿＿＿＿＿＿＿＿＿
    (c) How many students have marks less than 50.5?　　　＿＿＿＿＿＿＿＿＿
    (d) Explain briefly how you obtain these answers.

    ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

3. By using the built in function "Addition" in *Excel*, complete Table 2 (Cumulative Frequency Distribution Table) on the spreadsheet.

4. Use Table 2 to plot a line graph by *Excel*. The graph generated is called a cumulative frequency polygon.

5. From the graph, answer the following questions:

    (a) How many students have marks less than 50?

    ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

    How do you find it from the graph? (Answer orally)

    (b) If the pass mark is 60, how many students pass?

    ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

    How do you find it from the graph? (Answer orally)

(c) How many students obtain marks between 65 and 85? Show your work on the graph.

_____

(d) If 40% of the students pass this test, what is the pass mark?

_____

How do you obtain your answer? Write down your work below.

_____

_____

_____

(e) Find the percentage of students with marks less than 70.

_____

_____

_____

Discussion

6. Suppose the teacher finds that she has made a mistake and the original marks '72' and '83' were incorrectly typed as '27' and '38', suggest how you could obtain the correct graph. Discuss the procedures involved.

Exercise for consolidation

7. Plot the cumulative frequency polygon for question 6 and answer question 5 again.

Notes for Teachers

1. From questions 2 to 4, students are expected to learn the meaning of cumulative frequency and cumulative frequency polygon. The teacher should explain that the values in the left-hand column of the cumulative frequency distribution table are in fact the upper class boundaries of the class intervals.

2. The teacher should remind students to write down all the necessary labels (the horizontal axis, the vertical axis and the title of the graph) on the graph.

3. The teacher should tell students that the software does not provide a function to add a line in the cumulative frequency polygon. If a line is to be added manually, the **Draw Tool** could be used.

4. The teacher may alter the graph by changing the data in the frequency table or provide another set of data and then ask students similar questions to consolidate their concepts on the cumulative frequency polygon. For instance, in question 7 of the worksheet , by changing certain values in the frequency distribution table, the corresponding cumulative frequencies and a new graph will be displayed immediately.

5. By *Excel*, students may appreciate the following advantages:
    (a) A cumulative frequency polygon can be plotted easily.
    (b) The cumulative frequency polygon will change automatically when the data is changed.
    (c) It provides opportunities to make contrast and comparison between different graphs when some of the data are changed.

6. After completing the worksheet, the teacher can also introduce the cumulative frequency curve and ask students to try to plot it.

7. Suggested answers to the worksheet:

    Question 1.

Table 1			
Frequency Distribution Table			
Marks	Class mark	Class Boundaries	Frequency
21-30	25.5	20.5 - 30.5	3
31-40	35.5	30.5 - 40.5	6
41-50	45.5	40.5 - 50.5	4
51-60	55.5	50.5 - 60.5	6
61-70	65.5	60.5 - 70.5	7
71-80	75.5	70.5 - 80.5	5
81-90	85.5	80.5 - 90.5	6
91-100	95.5	90.5 - 100.5	3

    Question 2   (a)   3

    (b)   9

    (c)   13

    (d)   The answers in question 2(a) to 2(c) are obtained by adding each frequency to the sum of the previous frequencies.

    Question 3

Table 2	
Cumulative Frequency Distribution Table	
Marks less than	Cumulative Frequencies
20.5	0
30.5	3
40.5	9
50.5	13
60.5	19
70.5	26
80.5	31
90.5	37
100.5	40

    Question 5   (a)   About 13

    (b)   About 21

    (c) $34 - 23 = 11$

    (d)   About 68.5

    (e) $\dfrac{26}{40} \times 100\% = 65\%$

Question 6   Adjusting the frequencies in the Frequency Distribution Table (Table 1) automatically updates the corresponding cumulative frequencies and hence a new cumulative frequency polygon is drawn.

Question 7   (a)   About 11

(b)   About 23

(c) $34 - 21 = 13$

(d)   About 70

(e) $\dfrac{24}{40} \times 100\% = 60\%$
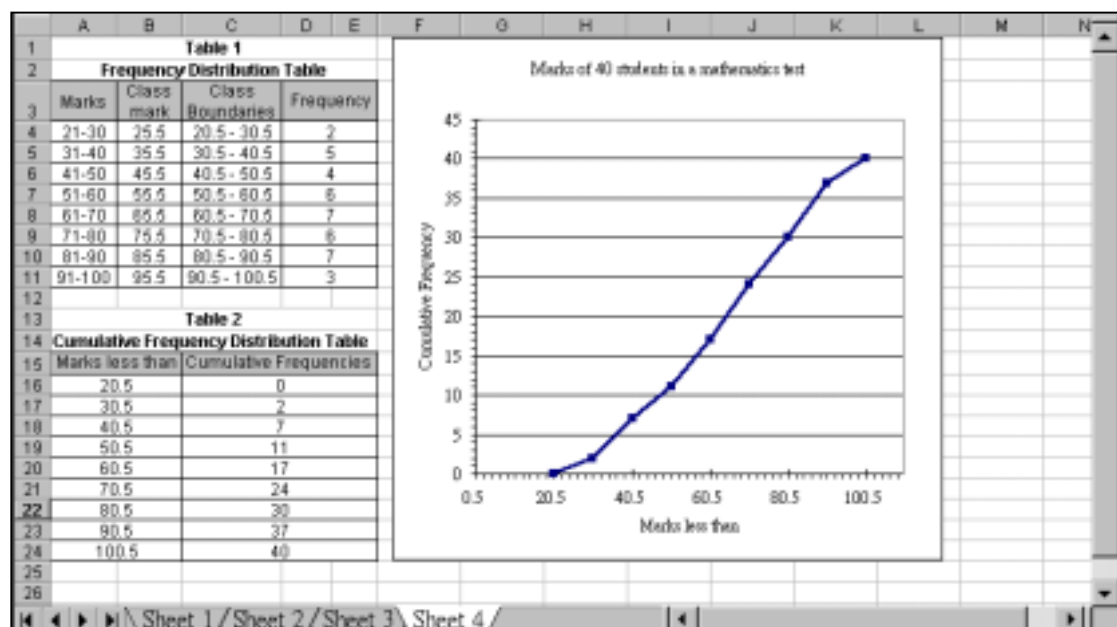
(See sheet 4 of the file dh06_e.xls)

Operation Procedure

(I)   Construct the frequency table

Input the data of the frequency table from cell A4 to D11.

(II)   Construct the cumulative frequency table

1. Key in the values from cell A16 to A24 in the "Marks less than" column. Key in "0" in cell C16 directly.
2. Key in the formula =C16+D4 in cell C17.
3. Select cell C17. Move the mouse pointer to the bottom right hand corner of this cell. Hold down the left button of the mouse and drag the pointer down to cell C24 to copy the formula in cell C17 to cell C18 to C24.

(III)   Draw the cumulative frequency polygon

1. Highlight cell A16 to C24.
2. Click the **ChartWizard** button. A dialog box will appear. Select the chart type **XY Scatter** on the left, select the subtype polygon on the right and click the **Next** button.
3. In the dialog box of Step 2, all the necessary data are filled in automatically. Click the **Next** button.
4. In the dialog box of Step 3, select the page **Titles**.
    Enter as follows:
Chart title	: Marks of 40 students in a mathematics test
Value (X) axis	: Marks less than
Value (Y) axis	: Cumulative Frequency
5. Select the page **Legend**, uncheck the box of **Show legend** and click the **Finish** button.

(IV)   Edit the scale of the x-axis

Double-click the line of the x-axis. Then a dialog box **Format Axis** will appear on the screen. Select the page of **Scale**, enter as follows:
Minimum	: 0.5
Maximum	: 110
Major unit	: 20
Minor unit	: 2

Click the **OK** button.

(V)   Edit the patterns of the x-axis and the y-axis

Double-click the line of the x-axis. Select the page **Patterns**, and then tick the following:

Major tick mark type   : Outside
Minor tick mark type   : Outside
Tick mark labels   : Next to axis

(VI)   Modify the appearance of the graph

1. To change the background colour of the chart, double-click the background of the chart.
2. To change the colour of the polygon, double-click the polygon.
3. To change the scale and font size of the x-axis or the y-axis, double-click the required axis.
4. To delete the labels, double-click the label and then press the delete key on the keyboard.
5. To move a label to another position, double-click and drag the label.

(VII) Modify the layout of the graph

Right-click a clear region on the graph. A small dialog box will appear:
1. To change the chart type:
    Click **Chart type**. The dialog box **Chart type** will appear. Select the page of **Standard type** and then select any chart type and chart sub-type. The graph will be changed accordingly.
2. To change the source data:
    Click **Source data.** The dialog box **Source data** will appear. Select the page of **Data range** and enter the new range of data.
3. To add the gridlines:
    Click **Chart options**. The dialog box **Chart options** will appear. Select the page of **Gridlines**. In the **Value (X) axis**, check the boxes of **Major gridlines** and/or **Minor gridlines**. In the **Value (Y) axis**, check the boxes of **Major gridlines** and/or **Minor gridlines** as required. Click the **Finish** button.

(VIII) Draw horizontal and vertical lines to show a point on the cumulative frequency graph

1. Click the **Line** button on the Draw Toolbar to draw a line at the desired position on the graph to show the point on the cumulative frequency graph.
2. Highlight the line. Click the **Dash Style** on the Draw Toolbar to change it to a dotted line.