Statistics

Spring 2023

Lecture 2

Feb 19-8:47 AM

Making a freq. dist. table: SG 3.4

It is a method to organize the collected data.

class limits	class BNDRS	class MP	class F	Cum. F	Rel. F	% F

with combination of columns, we can do drawing such as

1) Bar chart
2) Histogram
3) Ogive
4) Freq. Polygon
5) Pie chart

To make a freq. table, we need to know class width.

$$\text{class width} = \frac{\text{Range}}{\#\text{ of classes}} = \frac{\text{max} - \text{min}}{\text{Given}}$$

If whole # $\Rightarrow$ Add 1, If decimal $\Rightarrow$ Round-up

Feb 14-6:50 PM

I randomly selected 20 students, and here are their ages:

18 19 20 20 23
25 25 25 27 29
30 30 31 33 33
35 37 40 40 42

1) Sample Size $n = 20$

2) Min = 18

3) Max. = 42

4) Range = Max − Min
$= 42 - 18 = 24$

5) Midrange $= \frac{\text{Max} + \text{Min}}{2} = \frac{42+18}{2} = 30$

Let's make a freq. table with 3 classes.

Class width $= \frac{\text{Range}}{3} = \frac{24}{3} = 8$ ← Since 8 is a whole # → Add 1

CW = 9

Feb 14-6:57 PM

class limits	class BNDRS	class MP	class F	Cum. F	Rel. F	% F
18 – 26	17.5 – 26.5	22	8 →	8	.40	40%
27 – 35	26.5 – 35.5	31	8 →	16	.40	40%
36 – 44	35.5 – 44.5	40	4 →	20	.20	20%

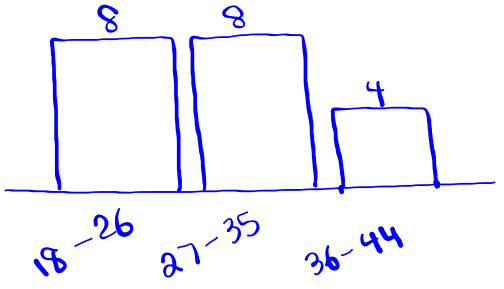
$n = 20$

26 27

Class MP $= \frac{\text{class limits}}{2} = \frac{\text{class BNDRS}}{2}$

Rel. F $= \frac{f}{n} = \frac{f}{20}$

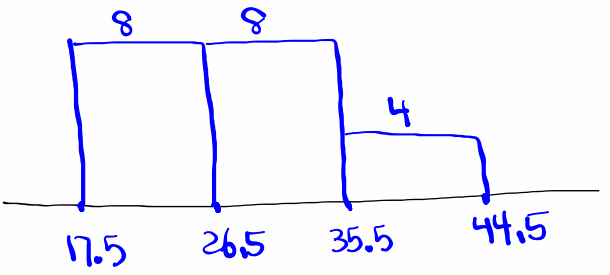
Bar chart
- class limits
- class F



Feb 14-7:03 PM

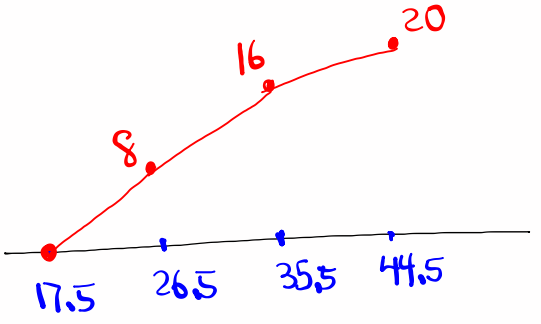
Histogram

- class BNDRS
- class F



Ogive

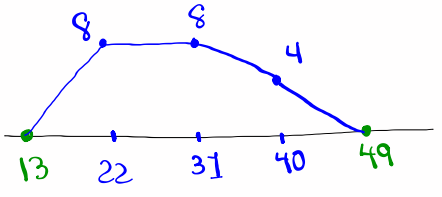
- class BNDRS
- Cum. F.
- Start at 0 level
- Increasing.



Feb 14-7:13 PM

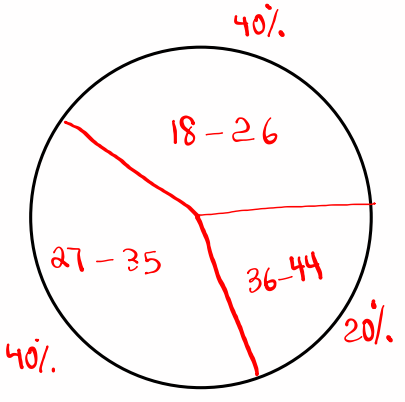
Freq. Polygon

- class MP
- class F
- extra MP, one on each side
- start & stop at 0 level.



Pie chart

- Class limits to name each slice
- % F for size of each slice
- Circle



Feb 14-7:17 PM

I randomly selected 25 exams. Here are the scores given in the form of Stem Plot

Data must be Sorted

Stem-and leaf

Stem	Leaf
5	8
6	0 4 8
7	0 0 2 5 5 8 9
8	3 4 5 5 5 6 8 9
9	0 2 8 9
10	0 0

1) $n = 25$

2) Min. = 58

3) Max = 100

4) Range = Max − Min. = 42

5) Midrange: $\frac{\text{Max} + \text{Min}}{2} = 79$

6) Mode = 85

Make a Freq. table with 4 classes.

Class width = $\frac{\text{Range}}{\#\text{classes}} = \frac{42}{4} = 10.5$

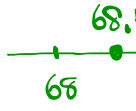
it is decimal ⇒ Round-up

CW = 11

Feb 14-7:22 PM

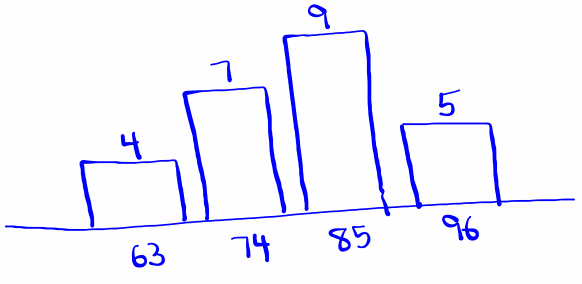
Class limits	class BNDRS	class MP	class F	Cum. F	Rel. F	% F
58 – 68	57.5 – 68.5	63	4	4	.16	16%
69 – 79	68.5 – 79.5	74	7	11	.28	28%
80 – 90	79.5 – 90.5	85	9	20	.36	36%
91 – 101	90.5 – 101.5	96	5	25	.20	20%

$n = 25$



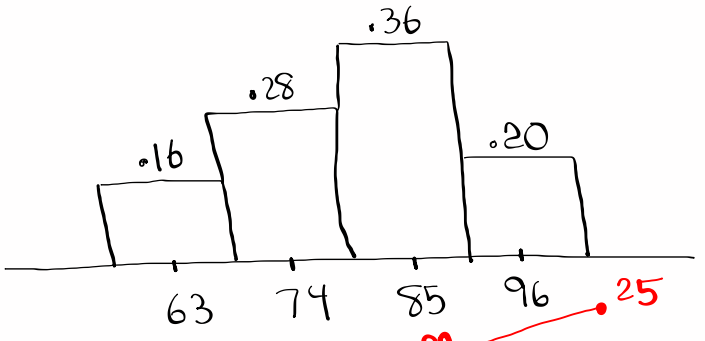
, class MP = $\frac{\text{t limits}}{2}$, Rel. F = $\frac{F}{n} = \frac{f}{25}$

Bar chart
- class MP
- class F

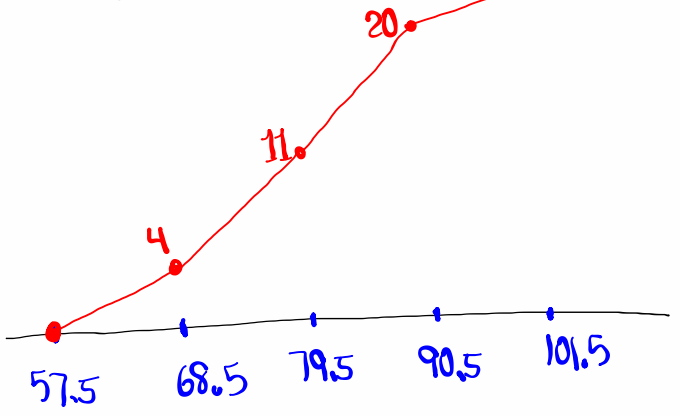


Feb 14-7:29 PM

Histogram
- class MP
- Rel. F.

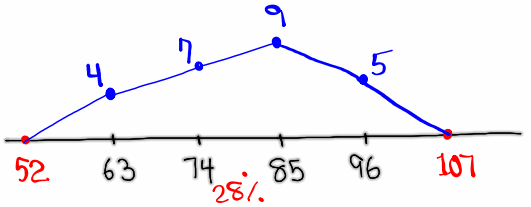


Ogive
- class BNDRS
- Cum. F.

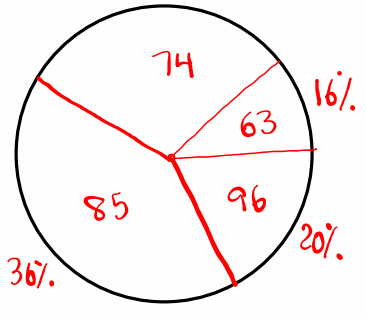


Feb 14-7:42 PM

Freq. Polygon
- class MP
- extra MP
- class F



Pie chart
- class MP
- % F
- Circle



What percentage of Score were between 69 & 90?

28% + 36% ⇒ 64%

Feb 14-7:46 PM

A data set had a range of 36.
Find class width if we wish to have a freq. table with

a) 4 classes

$CW = \frac{\text{Range}}{4} = \frac{36}{4} = 9$

whole # → CW = 10

b) 5 classes.

$CW = \frac{\text{Range}}{5} = \frac{36}{5} = 7.2$

decimal → CW = 8

SG 3 & 4 ✓

Feb 14-7:52 PM

Introduction to Some computing: SG 5-8

x → Data element

$\sum x$ → Summation of x

n → Sample Size

$\bar{x}$ → "x-bar" → Sample Mean (Average)

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

ex: Consider the Sample below

2, 4, 4, 4, 6

1) $n = 5$
2) Min. = 2
3) Max = 6
4) Range = Max − Min = 4
5) Midrange $= \frac{\text{Max} + \text{Min}}{2} = 4$
6) Mode = 4

7) $\sum x = 2 + 4 + 4 + 4 + 6 = 20$

8) $\bar{x} = \frac{\sum x}{n} = \frac{20}{5} = 4$

Feb 14-8:16 PM

Consider the Sample below:

1 3 3 3 4 4 4 11

1) $n = 8$

2) Min. = 1

3) Max = 11

4) Range = Max − Min = 10

5) Midrange = $\frac{\text{Max} + \text{Min}}{2}$ = 6

6) Mode = 3 & 4 Bimodal

7) $\Sigma x = 1+3+3+3+4+4+4+11 = \boxed{33}$

8) $\bar{x} = \frac{\Sigma x}{n} = \frac{33}{8} = \boxed{4.125}$

Round
whole → 4
1-decimal → 4.1
2-decimals → 4.13

Feb 14-8:21 PM

$x \rightarrow$ Data element

$\Sigma x \rightarrow$ Add all data elements

$x^2 \rightarrow$ Data element to Second Power

$\Sigma x^2 \rightarrow$ Square every data element, then find the Sum

$s^2 \rightarrow$ Sample Variance

$$s^2 = \frac{\Sigma (x - \bar{x})^2}{n-1}$$

$\bar{x}$: Sample Mean

n: Sample Size

$$s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)}$$

Feb 14-8:27 PM

Consider the Sample below

3 5 7 9 11 1) $n = 5$

2) Range $= 8$ 3) Midrange $= 7$ 4) Mode None

5) $\Sigma x = 3 + 5 + 7 + 9 + 11 = 35$

6) $\Sigma x^2 = 3^2 + 5^2 + 7^2 + 9^2 + 11^2 = 285$

7) $\bar{x} = \frac{\Sigma x}{n} = \frac{35}{5} = 7$

8) $s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{5 \cdot 285 - 35^2}{5(5-1)}$

$= \frac{200}{20} = \boxed{10}$

Feb 14-8:30 PM

Given $n = 10$, $\Sigma x = 286$ $\Sigma x^2 = 8944$

Min $= 18$, Max $= 45$

1) Range $= 45 - 18$

$= \boxed{27}$

2) Midrange $= \frac{45+18}{2}$

$= \boxed{31.5}$

3) $\bar{x} = \frac{\Sigma x}{n} = \frac{286}{10} = \boxed{28.6}$

4) $s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{10 \cdot 8944 - 286^2}{10(10-1)}$

$= \frac{7644}{90}$

$= 84.9\overline{3}$

$\approx \boxed{84.933} \approx \boxed{85}$

Feb 14-8:37 PM

$\bar{x}$ → Sample Mean $\quad \bar{x} = \frac{\Sigma x}{n}$

s^2 → Sample Variance $\quad s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)}$

s → Sample Standard deviation $\quad s = \sqrt{s^2}$

Given $n = 8$, $\Sigma x = 56$, $\Sigma x^2 = 422$

1) $\bar{x} = \frac{\Sigma x}{n} = \frac{56}{8} = \boxed{7}$

2) $s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{8 \cdot 422 - 56^2}{8(8-1)} = \frac{240}{56}$

$s^2 \approx 4.286$

3) $s = \sqrt{s^2} = \sqrt{4.286} \approx \boxed{2.070}$

Feb 14-8:43 PM

How to estimate Sample Standard deviation:

$s \approx \frac{\text{Range}}{4}$

"The range rule-of-thumb"

A data set has a min. value of 12 and max. value of 80.

1) Range = Max − Min
$= 80 - 12 = \boxed{68}$

2) Midrange $= \frac{\text{Max} + \text{Min}}{2}$
$= \frac{80 + 12}{2} = \boxed{46}$

3) Estimate s

$s \approx \frac{\text{Range}}{4} = \frac{68}{4} = \boxed{17}$

Feb 14-8:50 PM