Statistics

Lecture 3

Feb 19-8:47 AM

I randomly selected 20 students, here are the ages

18 19 20 20 23

24 25 25 25 28

29 30 32 35 35

38 40 42 45 48

1) Sample Size $n = 20$

2) Min = 18, Max = 48

3) Range = Max − Min = 48 − 18 = 30

4) Midrange = $\frac{\text{Max} + \text{Min}}{2} = \frac{48+18}{2} = \frac{66}{2} = 33$

5) Mode (value that repeated the most) = 25

I wish to organize this Sample in a table with 3 classes.

→ frequency table.

we must find class width.

Class width = $\frac{\text{Range}}{\#\text{ of classes}}$ if decimal → Round up; if whole # → Add 1

CW = $\frac{\text{Range}}{3} = \frac{30}{3} = 10$ CW = 11

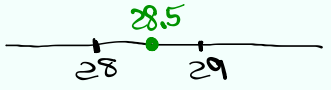
Feb 11-8:08 PM

Freq. table

CW = 11

class limits	class BNDRS	class MP	class F	cum. F	Rel. F	%F
18 – 28	17.5 - 28.5	23	10	10	.50	50%
29 – 39	28.5 - 39.5	34	6	16	.30	30%
40 – 50	39.5 - 50.5	45	4	20	.20	20%

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



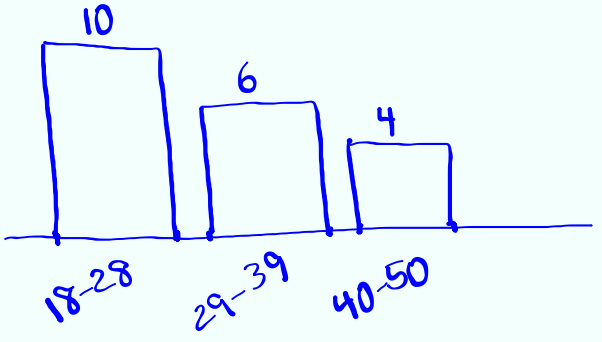
Rel. F = $\frac{f}{n} = \frac{f}{20}$ (n = 20)

Sample Size

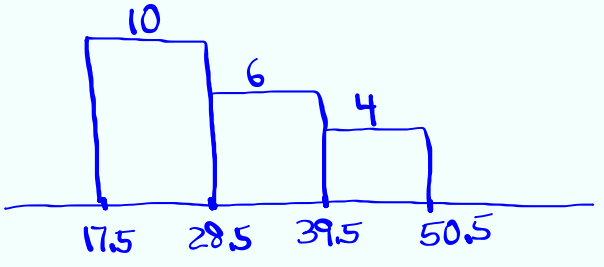
Draw
- Bar chart
- Histogram
- Ogive
- Freq. Polygon
- Pie chart

Feb 11-8:19 PM

Bar Chart
- class limits
- class F

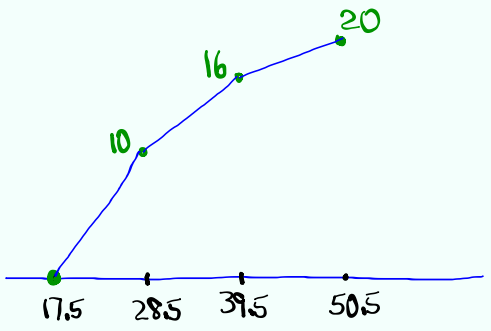


Histogram
- class BNDRS
- class F

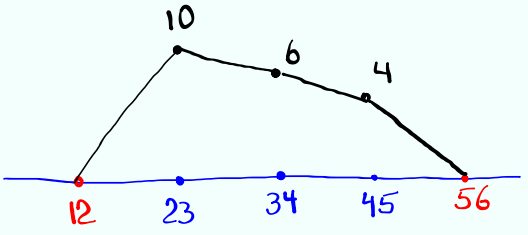


Feb 18-7:02 PM

Ogive
- Class BNDRS
- Cum. F
- Start at 0 level

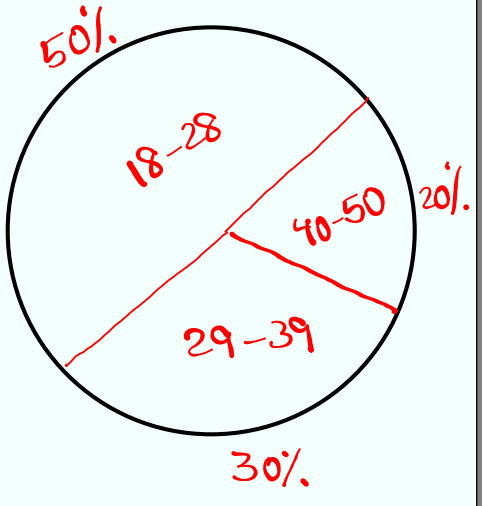


Freq. Polygon
- class MP
- class F
- One additional MP on each side
- Start & Stop at 0 level.



Feb 18-7:06 PM

Pie Chart
- Circle
- class limits to name each slice
- %F for Size of each Slice.



Feb 18-7:12 PM

I randomly selected 25 exams. Here are the Scores:

50 52 58 62 65
68 68 70 73 75
75 75 78 80 83
85 88 88 88 90
92 95 96 100 100

1) $n = 25$

2) Min = 50 Max = 100

3) Range = Max − Min
= 50

4) Midrange
$= \frac{\text{Max} + \text{Min}}{2}$
$= \frac{100 + 50}{2}$
$= \frac{150}{2} = 75$

5) Mode
88 & 75
Bimodal

6) Make a freq. table with 4 classes.

$CW = \frac{\text{Range}}{4} = \frac{50}{4} = 12.5$ ↑ Decimal

$CW = 13$

Feb 11-8:31 PM

CW = 13

Class limits	class BNDRS	Class MP	class F	Cum. F	Rel. F	% F
50 – 62	49.5 – 62.5	56	4	4	.16	16%
63 – 75	62.5 – 75.5	69	8	12	.32	32%
76 – 88	75.5 – 88.5	82	7	19	.28	28%
89 – 101	88.5 – 101.5	95	6	25	.24	24%

$n = 25$

$\text{Rel. F} = \frac{f}{n}$
$= \frac{f}{25}$

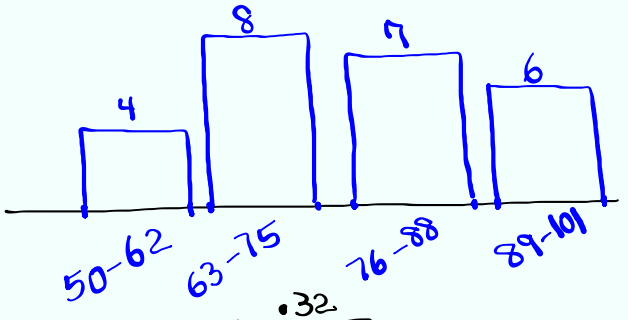
$\text{Class MP} = \frac{\text{+ class limits}}{2}$

62 62.5 63

Feb 11-8:39 PM

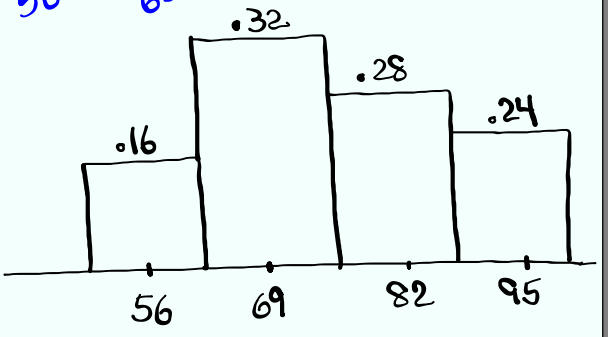
Bar chart

- Class limits
- Class F



Histogram

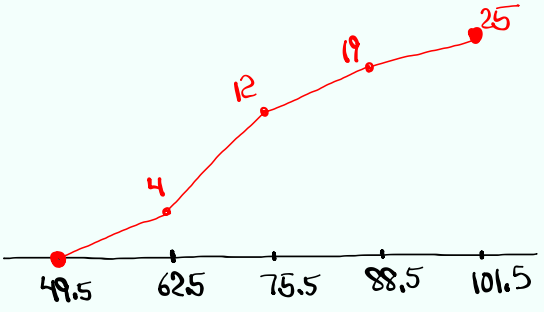
- Class MP
- Rel. F



Feb 18-7:22 PM

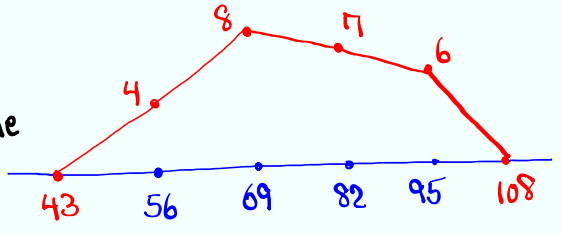
Ogive

- Class BNDRS
- Cum. F.
- Start at 0 level



Freq. Polygon

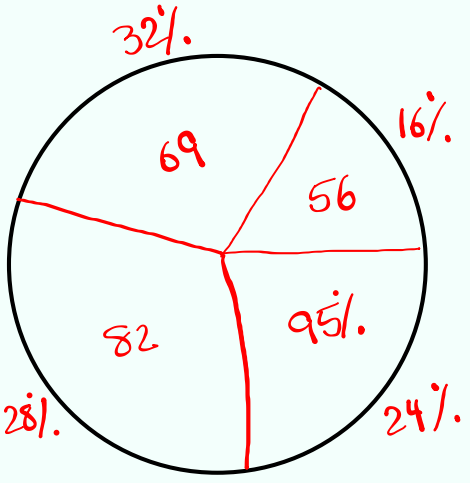
- Class MP
- Class F
- extra MP, one on each side
- Start & Stop at 0 level.



Feb 18-7:27 PM

Pie Chart

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



Feb 18-7:33 PM

50 52 58 62 65
68 68 70 73 75
75 75 78 80 83
85 88 88 88 90
92 95 96 100 100

Stem Plot
Data must be Sorted

5	028
6	2588
7	035558
8	035888
9	0256
10	00

How many data elements are below 70? 7

What % of data elements are below 70?

7 is what % of 25? ← Sample Size

$7 = \frac{P}{100} \cdot 25$ (4)

$7 = \frac{P}{4}$

Multiply by 4

$P = 7 \cdot 4 = 28$

28%

Sc 3 & 4

Feb 18-7:37 PM

Special Calculations/ Computations in Statistics:

SG 5-8

$x \rightarrow$ Data element

$\sum x \rightarrow$ Sum of data elements

$\sum$ → Summation

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

$n \rightarrow$ Sample Size

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

Consider the Sample below

2, 3, 3, 5, 7

$\sum x = 2+3+3+5+7 = \boxed{20}$

$n = 5$

Range $= 7 - 2 = 5$

Midrange $= \frac{7+2}{2} = 4.5$

Mode $= 3$

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

Feb 18-7:44 PM

Consider the Sample below

2, 3, 4, [4, 6] 6, 7, 8

1) $n = \boxed{8}$

2) Range $= 8 - 2 = \boxed{6}$

3) Midrange $= \frac{8+2}{2} = \boxed{5}$

4) Mode $= 4 \& 6$

5) $\sum x = 2+3+4+4+6+6+7+8 = \boxed{40}$

6) $\bar{x} = \frac{\sum x}{n} = \frac{40}{8} = \boxed{5}$

7) Median

Data must be Sorted

It is the value in the middle.

If no value → Add Middle two # / 2

MD $= \frac{4+6}{2} = \boxed{5}$

Feb 18-7:51 PM

$x \rightarrow$ Data element

$\sum x \rightarrow$ Sum of data elements

$x^2 \rightarrow$ (Data element)^2

$\sum x^2 \rightarrow$ Sum of (Data element)^2

$n \rightarrow$ Sample Size

$s^2 \rightarrow$ Sample Variance

$\bar{x} \rightarrow$ Sample Mean

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

$$s^2 = \frac{\sum (x-\bar{x})^2}{n-1} \quad \text{OR} \quad \boxed{s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)}}$$

Feb 18-7:58 PM

Consider the sample below

1, 3, 3, 3, 5

1) $n = 5$

2) Range $= 4$

3) Midrange $= 3$

4) Mode $= 3$

5) $\sum x = 1+3+3+3+5 = \boxed{15}$

6) $\bar{x} = \frac{\sum x}{n} = \frac{15}{5} = \boxed{3}$

7) $\sum x^2 = 1^2+3^2+3^2+3^2+5^2 = \boxed{53}$

8) $s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{5 \cdot 53 - 15^2}{5(5-1)} = \frac{40}{20} = \boxed{2}$

Feb 18-8:02 PM

Consider the Sample below

2, 3, 3, 5, 7, 8, 10

1) $n = 7$

2) Range $= 10-2 = 8$

3) Midrange $= \frac{10+2}{2} = 6$

4) Mode 3

5) Median $= 5$

6) $\sum x = 2+3+3+5+7+8+10 = 38$

7) $\sum x^2 = 2^2+3^2+3^2+5^2+7^2+8^2+10^2 = 260$

8) $\bar{x}$, Round to whole #

$$\bar{x} = \frac{\sum x}{n} = \frac{38}{7} = 5.429 \approx 5$$

9) s^2, Round up to whole

$$s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{7 \cdot 260 - 38^2}{7(7-1)} = \frac{376}{42} = 8.952 \approx 9$$

Feb 18-8:08 PM

$\bar{x}$ → Sample Mean

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

s^2 → Sample Variance

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

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

s → Sample Standard deviation

$$s = \sqrt{s^2}$$

Feb 18-8:17 PM

Consider the Sample below

1, 2, 2, 2, 3

1) $n = 5$

2) Range $= 2$

3) Midrange $= 2$

4) Mode $= 2$

5) Median $= 2$

6) $\Sigma x = 10$

7) $\Sigma x^2 = 22$

8) $\bar{x} = \frac{\Sigma x}{n} = \frac{10}{5} = \boxed{2}$

9) $s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{5 \cdot 22 - 10^2}{5(5-1)}$

$= \frac{10}{20} = \boxed{.5}$

10) s, Round to $\boxed{2\text{-dec}}$

$s = \sqrt{s^2} = \sqrt{.5} \approx .707 \approx \boxed{.71}$

Feb 18-8:20 PM

Given $n = 10$, $\Sigma x = 60$, $\Sigma x^2 = 360$

1) $\bar{x} = \frac{\Sigma x}{n} = \frac{60}{10} = \boxed{6}$

2) $s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{10 \cdot 360 - 60^2}{10(10-1)} = \frac{0}{90} = \boxed{0}$

3) $s = \sqrt{s^2} = \sqrt{0} = \boxed{0}$

How to estimate s:

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

Range Rule-of-thumb

Feb 18-8:27 PM

What is Sample Standard deviation?

- use s
- $s \geq 0$
- It measures the variation of data elements from $\bar{x}$.

If s is small $\Rightarrow$ Data elements are close to $\bar{x}$.

If s is big $\Rightarrow$ data elements are spread out from $\bar{x}$.

If s is zero $\Rightarrow$ All data elements are equal to the mean.

Feb 18-8:33 PM

Empirical Rule:

It is best when mean = mode = median.

Data dist. will be Symmetric about $\bar{x}$.

About 68% of data fall within $\bar{x} \pm s$.

Usual Range

About 95% of data fall within $\bar{x} \pm 2s$

About 99.7% of data fall within $\bar{x} \pm 3s$

Feb 18-8:37 PM

I randomly Selected 80 exams. Scores were distributed Symmetrically with $\bar{x} = 82$ and $S = 6$.

68% Range → $\bar{x} \pm S = 82 \pm 6 \Rightarrow$ 76 to 88

95% Range → $\bar{x} \pm 2S = 82 \pm 2(6) = 82 \pm 12$

$\Rightarrow$ 70 to 94

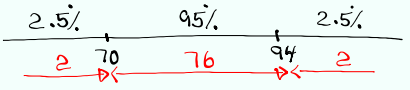
Usual Range

What % of Scores were above 70? 97.5%

How many Scores were Usual Scores?

95% of 80 = .95(80) = 76

↑ Sample Size



SG 5

You can do Problems from SG 6 - 8.

Feb 18-8:42 PM