Statistics

Lecture 2

Feb 19-8:47 AM

(S6.3)

One way to organize raw data is by making a frequency table.

1) We need to know the range of data
   → Max − Min.

2) We need to know how many classes we are having.
   groups, bins
   # of classes will be given.

3) class width = $\frac{\text{Range}}{\text{\# of classes}}$

If decimal → Round-up to a whole #
If whole → Add 1

Suppose a data set has a min. of 18 and max. of 60.

1) Range = 60 − 18 = 42

2) class width for 3 classes.
   $CW = \frac{\text{Range}}{3} = \frac{42}{3} = 14$ → CW = 15
   ↑ whole #

3) class width for 4 classes.
   $CW = \frac{\text{Range}}{4} = \frac{42}{4} = 10.5$ → CW = 11
   ↑ Decimal #

Mar 7-8:05 AM

A Freq. table consists of

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

Each Row belongs to 1 class.

Combination of certain Columns will be used to do graphs. 1) Bar chart 2) Histogram 3) Ogive 4) Freq. Polygon 5) Pie chart

Mar 7-8:14 AM

I Surveyed 20 students, and here are their ages

18	20	21	24	25
25	28	30	32	32
32	35	38	40	44
48	52	55	58	60

1) $n = 20$ ← Sample Size

2) Range = Max. − Min.
$= 60 - 18 = \boxed{42}$

3) Midrange $= \frac{\text{Max.} + \text{Min.}}{2} = \frac{60+18}{2} = \frac{78}{2} = \boxed{39}$

4) Mode = 32

5) Make STEM plot
Data must be Sorted!

1	8
2	0 1 4 5 5 8
3	0 2 2 2 5 8
4	0 4 8
5	2 5 8
6	0

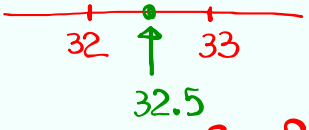
6) Find class width if we wish to have a freq. table with 3 classes.

$CW = \frac{\text{Range}}{3} = \frac{42}{3} = 14 \rightarrow \boxed{CW = 15}$

↑ whole #

Mar 7-8:18 AM

Class limits	Class BNDRS	Class MP	Class F	Cum. F	Rel. F	% F
18 – 32	17.5 – 32.5	25	11	11	.55	55%
33 – 47	32.5 – 47.5	40	4	15	.20	20%
48 – 62	47.5 – 62.5	55	5	20	.25	25%

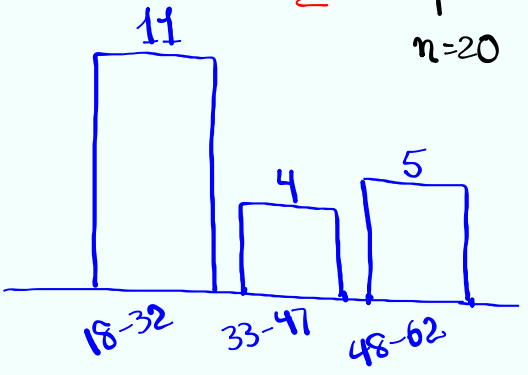


Class MP = $\frac{\text{+ class limits}}{2} = \frac{\text{+ class BNDRS}}{2}$

$n = 20$

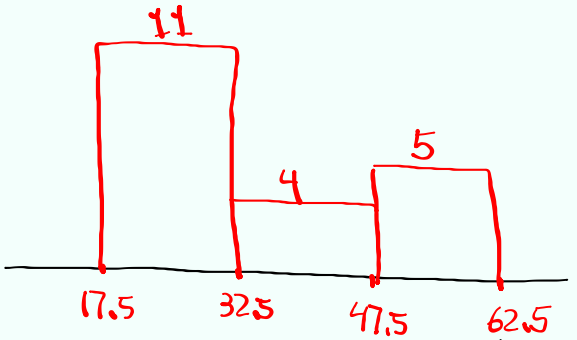
$\text{Rel. F} = \frac{f}{n} = \frac{f}{20}$

Bar Chart
- Class limits
- Class F

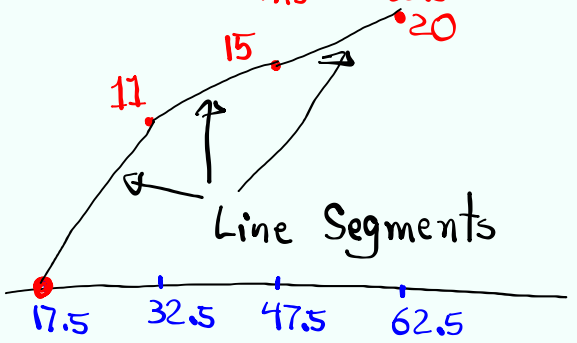


Mar 7-8:26 AM

Histogram
- Class BNDRS
- Class F



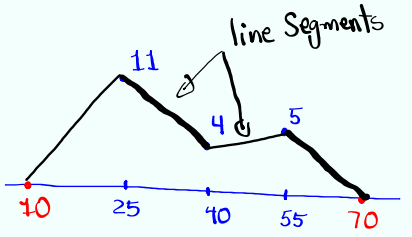
Ogive
- Class BNDRS
- Cum. F.
- Start at 0 level



Mar 7-8:41 AM

Freq. Polygon

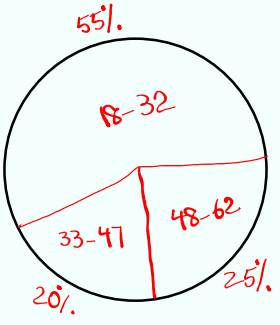
- class MP
- class F
- one extra MP on each side
- Start & Stop at 0 level



Pie chart

- Circle
- class limits or MP to name each Slice
- % F to Size each Slice

Sr 3 & 4 ✓



Mar 7-8:47 AM

Consider the Stem Plot below

5	2 5
6	0 2 5 8 8
7	2 3 5 5 5 7 9
8	0 5 5 5 8 9
9	3 5 8
10	0 0

1) $n = 25$

2) Range = Max − Min = 100 − 52 = 48

3) Midrange = $\frac{\text{Max}+\text{Min}}{2} = \frac{100+52}{2} = 76$

4) Mode = 75 & 85 Bimodal

5) How many are below 70? 7

6) what % of data are below 70?

$\frac{7}{25} \cdot 100 = 28$ **28%**

28% | 72%

70

7) Find class width if we wish to have a freq. table with

a) 4 classes

$CW = \frac{\text{Range}}{4} = \frac{48}{4} = 12$

CW = 13

b) 5 classes

$CW = \frac{\text{Range}}{5} = \frac{48}{5} = 9.6$

CW = 10

Mar 7-8:55 AM

Make a freq. table with 4 classes. CW = 13

class limits	class BNDRS	class MP	class F	Cum. F	Rel. F	% F
52 – 64	51.5 – 64.5	58	4	4	.16	16%
65 – 77	64.5 – 77.5	71	9	13	.36	36%
78 – 90	77.5 – 90.5	84	7	20	.28	28%
91 – 103	90.5 – 103.5	97	5	25	.20	20%

$n = 25$

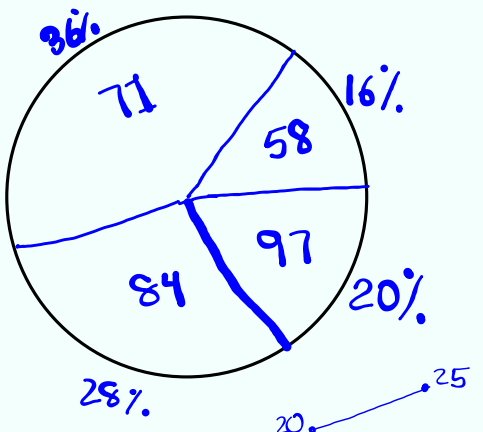
Class MP: $\frac{52+64}{2} = 58$

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

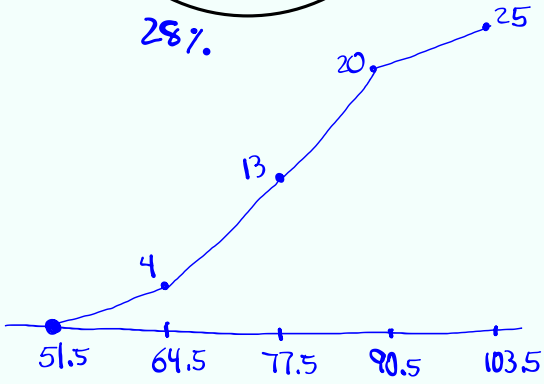
what % of data elements are between 65 and 90, inclusive? 36% + 28% = 64%

Mar 7-9:05 AM

Pie chart
- class MP
- % F



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



Mar 7-9:16 AM

Computations in Statistics

SG 5

$x \rightarrow$ Data element

$\sum x \rightarrow$ Summation of $x \rightarrow$ Sum of data elements

$n \rightarrow$ Sample Size

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

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

ex: Consider the Sample below

1 3 5 7 9

$n = 5$ Range $= 9 - 1 = 8$

Midrange $= \frac{9+1}{2} = 5$ Mode = None

$\sum x = 1 + 3 + 5 + 7 + 9 = 25$

$\bar{x} = \frac{\sum x}{n} = \frac{25}{5} = \boxed{5}$

Mar 7-9:41 AM

Consider the Sample below

1 3 5 5 5 7 8 9

1) $n = \boxed{8}$ 2) Range $= 9 - 1 = \boxed{8}$

3) Midrange $= \frac{9+1}{2} = \boxed{5}$ 4) Mode $= \boxed{5}$

5) $\sum x = 1 + 3 + 5 + 5 + 5 + 7 + 8 + 9 = \boxed{43}$

6) $\bar{x} = \frac{\sum x}{n} = \frac{43}{8} = 5.375$

Round to

whole $\rightarrow 5$

1-Dec. $\rightarrow 5.4$

2-Dec. $\rightarrow 5.38$

Mar 7-9:47 AM

$x \rightarrow$ Data elements

$n \rightarrow$ Sample Size

$\sum x \rightarrow$ Sum of data elements

$\sum x^2 \rightarrow$ square each data element, then add.

$\bar{x} \rightarrow$ x-bar $\rightarrow$ Sample Mean

$s^2 \rightarrow$ Sample Variance $\rightarrow s^2 \geq 0$

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

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

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

Mar 7-9:52 AM

Consider the Sample below

1 3 3 3 5

1) $n = 5$

2) Range $= 5 - 1 = 4$

3) Midrange $= \frac{5+1}{2} = 3$

4) Mode $= 3$

5) $\sum x = 1+3+3+3+5 = 15$

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

7) $s^2 = \frac{\sum (x-\bar{x})^2}{n-1} = \frac{(1-3)^2+(3-3)^2+(3-3)^2+(3-3)^2+(5-3)^2}{5-1}$

$= \frac{(-2)^2+0^2+0^2+0^2+2^2}{4} = \frac{4+0+0+0+4}{4}$

$= \frac{8}{4} = \boxed{2}$

8) $\sum x^2 = 1^2+3^2+3^2+3^2+5^2$

$= 1+9+9+9+25 = \boxed{53}$

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

Mar 7-9:56 AM

Consider the Sample below

1 2 3 3 5 5 8 11

1) $n = \boxed{8}$

2) Range $= 11 - 1 = \boxed{10}$

3) Midrange $= \frac{11+1}{2} = \boxed{6}$

4) Mode $= 3$ & 5 Bimodal

5) $\Sigma x = 1+2+3+\cdots+8+11 = \boxed{38}$

6) $\Sigma x^2 = 1^2+2^2+3^2+\cdots+8^2+11^2 = \boxed{258}$

7) $\bar{x} = \frac{\Sigma x}{n} = \frac{38}{8} = \boxed{4.75}$

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

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

$= \frac{(1-4.75)^2+(2-4.75)^2+\cdots+(11-4.75)^2}{8-1}$

$= \frac{8 \cdot 258 - 38^2}{8(8-1)}$

$= \frac{620}{56} \approx 11.071$

Mar 7-10:04 AM

Suppose $n = 10$, $\Sigma x = 40$, $\Sigma x^2 = 160$

Find

1) $\bar{x} = \frac{\Sigma x}{n} = \frac{40}{10} = \boxed{4}$

we don't have the raw data

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

$= \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{10 \cdot 160 - 40^2}{10(10-1)} = \frac{1600-1600}{10 \cdot 9}$

$= \frac{0}{90} = \boxed{0}$

Mar 7-10:13 AM

$x \rightarrow$ Data elements $\quad n \rightarrow$ Sample Size

$\bar{x} \rightarrow x$- bar $\rightarrow$ Sample Mean $\quad \bar{x} = \frac{\sum x}{n}$

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

$s^2 \geq 0$

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

$s \rightarrow$ Sample Standard Deviation

$s = \sqrt{s^2}$

$s \geq 0$

Mar 7-10:18 AM

Consider the sample below

1 3 3 3 7

$n = 5$

$\sum x = 17$

$\bar{x} = \frac{\sum x}{n} = \frac{17}{5} = \boxed{3.4}$

$\sum x^2 = 77$

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

$= \frac{5 \cdot 77 - 17^2}{5(5-1)} = \frac{96}{20}$

$= \boxed{4.8}$

$s = \sqrt{s^2}$

$= \sqrt{4.8}$

$\approx \boxed{2.191}$

whole $\rightarrow 2$

1-Dec. $\rightarrow 2.2$

2-Dec. $\rightarrow 2.19$

Mar 7-10:22 AM

How to estimate S:

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

the range rule-of-thumb.

Suppose Min. = 15, Max = 50

Range = 50 − 15 = 35

Estimate S $S \approx \frac{\text{Range}}{4} = \frac{35}{4} = 8.75$

Mar 7-10:41 AM

Empirical Rule

use this when data dist. is symmetric

Mean ≈ Mode ≈ Median

- About 68% of data elements are within $\bar{x} \pm s$
- About 95% of data elements are within $\bar{x} \pm 2s$ — Usual Range
- About 99.7% of data elements are within $\bar{x} \pm 3s$

Mar 7-10:45 AM

I randomly selected 80 exams, results had a symmetric dist with $\bar{x}=82$ & $S=6$.

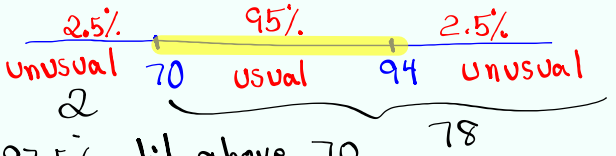
68% Range $\bar{x} \pm S = 82 \pm 6$

$\Rightarrow$ 76 to 88

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

$\Rightarrow$ 70 to 94

Usual Range



2

78

97.5% did above 70

$97.5\%\ (80) = .975(80) = 78$

did above 70

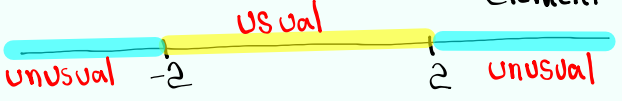
Mar 7-10:50 AM

Z - Score

Always Round to 3-decimal places

If $-2 \leq Z \leq 2$ $\rightarrow$ usual data element

If $Z < -2$ or $Z > 2$ $\rightarrow$ unusual data element



Z Scores allows us to compare data elements from different Samples.

Z Scores are a method to standardize data element.

$$Z = \frac{x - \bar{x}}{S}$$

Mar 7-10:57 AM

Lisa got 85 on a Math exam & 88 " " English exam.

Math $\bar{x} = 78$ & $S = 3$ $\quad Z = \frac{x - \bar{x}}{S} = \frac{85 - 78}{3} = \frac{7}{3} \approx 2.333$

Unusual Score

English $\bar{x} = 84$ & $S = 5$

$Z = \frac{x - \bar{x}}{S} = \frac{88 - 84}{5} = \frac{4}{5} = .8$

Usual Score

Work on SG 3 & 4

We do SG 5, 6, 7, and 8 next week

Come to class with TI-83 or TI-84.

Mar 7-11:01 AM