Statistics

Lecture 2

Feb 19-8:47 AM

Making Freq. table and graph

SG 3 & SG 4

To make a freq. table from raw data

1) Max & Min

2) Range = Max − Min

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

If decimal → Round-up to whole #

If whole → Add 1

class limits	class BNDRS	class MPt	classF	Cum. F	Rel. F	%F

Bar chart
Histogram
Ogive

Freq. Polygon
Pie chart

Sep 5-8:04 AM

I randomly Selected 20 Students, here are their ages:

18 19 20 20 24
25 25 25 28 30
32 32 34 36 37
39 40 45 50 54

1) Sample Size $n = 20$

2) Max = 54, Min = 18

3) Range = Max − Min
$= 54 - 18$
$= \boxed{36}$

4) Midrange $= \frac{\text{Max} + \text{Min}}{2} = \frac{54+18}{2} = \frac{72}{2} = \boxed{36}$

5) Find class width if we wish to have 3 classes

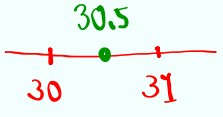
$CW = \frac{\text{Range}}{\#\text{ classes}} = \frac{36}{3} = 12$

whole #

$\boxed{CW = 13}$

Sep 5-8:11 AM

class limits	class BNDRS	class MP	class F	Cum. F	Rel. F	% F
18 – 30	17.5 – 30.5	24	10	10	.50	50%
31 – 43	30.5 – 43.5	37	7	17	.35	35%
44 – 56	43.5 – 56.5	50	3	20	.15	15%

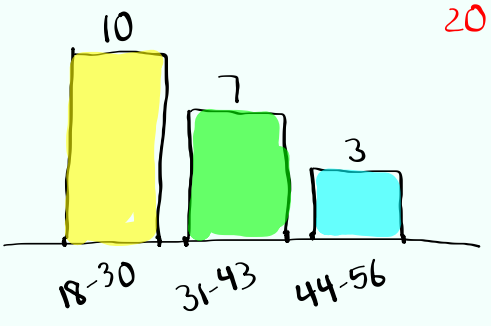


class MP $= \frac{\text{+ class limits}}{2}$

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

20

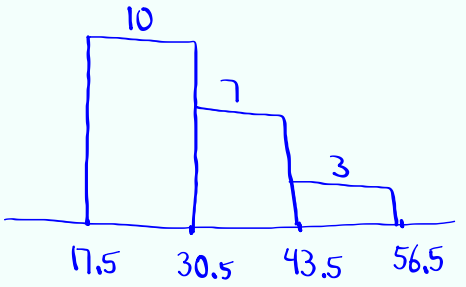
Bar Chart
- class limits
- class F



Sep 5-8:18 AM

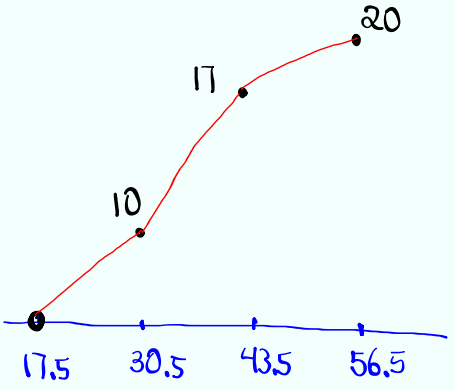
Histogram

- class BNDRS
- class F



Ogive

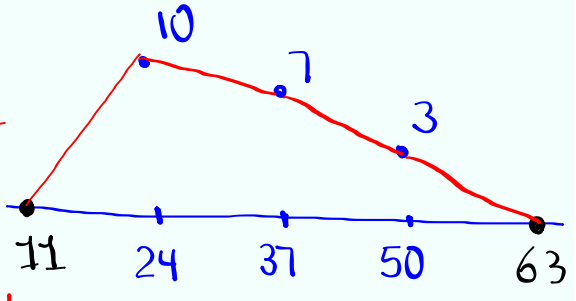
- class BNDRS
- Cum. F
- start at 0 level
- Non-Decreasing



Sep 5-8:31 AM

Freq. Polygon

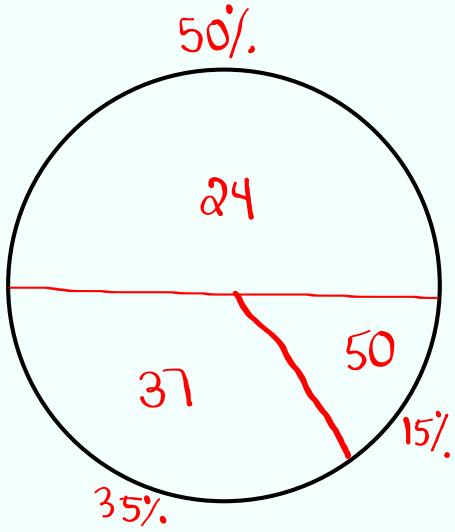
- class MP
- class F
- one extra midpoint on each side
- start & finish at 0 level



Sep 5-8:37 AM

Pie-chart

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



Sep 5-8:40 AM

I randomly selected 25 exams and here are the Scores.

53	58	60	65	65
68	70	72	72	72
75	78	79	82	82
82	85	89	90	93
95	95	98	100	100

1) $n = 25$

2) Range = 100 − 53 = 47

3) Midrange = $\frac{100+53}{2}$ = 76.5

4) Mode = 72 & 82 Bimodal

5) find class width if we wish to have 4 classes

Class width = $\frac{\text{Range}}{4} = \frac{47}{4} = 11.75$

decimal → Round-up

CW = 12

Sep 5-8:45 AM

53 58 60 65 65

68 70 72 72 72

75 78 79 82 82

82 85 89 90 93

95 95 98 100 100

Make Stem Plot

"Data must be Sorted"

5	38
6	0558
7	0222589
8	22259
9	03558
10	00

Sep 5-8:53 AM

Class limits	Class BNDRS	Class MP	Class F	Cum. F	Rel. F (3-dec.)	% F
53 - 64	52.5 - 64.5	58.5	3	3	.12	12%
65 - 76	64.5 - 76.5	70.5	8	11	.32	32%
77 - 88	76.5 - 88.5	82.5	6	17	.24	24%
89 - 100	88.5 - 100.5	94.5	8	25	.32	32%

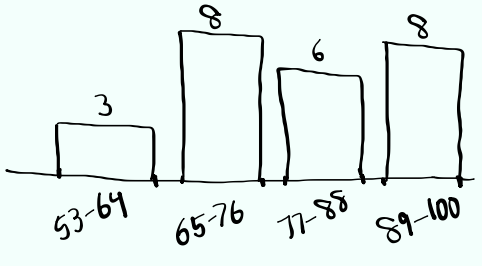
n = 25

64.5 (64 — 65)

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

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

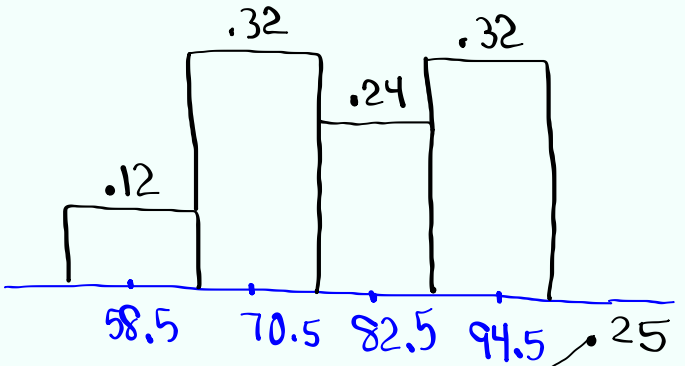
Bar chart
- class limits
- class F



Sep 5-8:57 AM

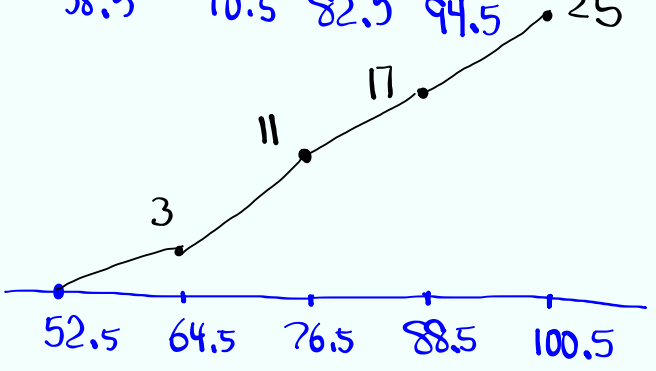
Histogram

- class MP
- Rel. F.



Ogive

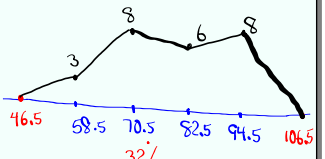
- Class BNDRS
- Cum. F.



Sep 5-9:11 AM

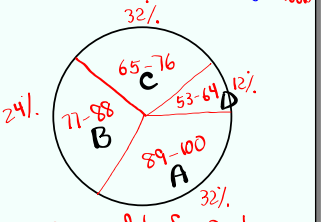
Freq. Polygon

- class MP
- class F
- extra MP



Pie chart

- Circle
- % F
- class limits



Be Careful of Scales

what % of Scores were between 76 & 89?

24%

what % of Scores were below 89?

100% − 32% = 68%

Sg. 3 & Sg 4 ✓

Sep 5-9:15 AM

Certain Computations in Statistics: (SG 5)

n → Sample Size

x → Data element

$\sum x$ → Sum of data elements

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

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

Consider the Sample below

2 3 3 7 10

$n = 5$

$\sum x = 2+3+3+7+10 = 25$

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

Mode = 3

Sep 5-9:40 AM

Consider the Sample below

2 3 4 4 6 6 8 12

1) $n = 8$

2) Range = 10

3) Midrange = 7

4) Mode 4 & 6

5) $\sum x = 2+3+4+4+6+6+8+12 = 45$

6) $\bar{x} = \frac{\sum x}{n} = \frac{45}{8} = 5.625$

Round to

whole → 6

1-Dec → 5.6

2-Dec. → 5.63

Sep 5-9:44 AM

n → Sample Size

x → Data element

$\sum x$ → Sum of data elements

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

x^2 → data element to Second power

$\sum x^2$ → Square each data element, then add

s^2 → Sample Variance

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

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

$s^2 \geq 0$ Always

Sep 5-9:50 AM

Consider the Sample below

1 3 3 3 5

1) $n = 5$

2) Mode = 3

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

4) $\sum x^2 = 1^2+3^2+3^2+3^2+5^2 = 53$

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

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

$= \boxed{2}$

Sep 5-9:55 AM

Given $n = 10$ $\sum x = 55$

$\sum x^2 = 379$ Min $= 1$ Max $= 11$

1) $\bar{x} = \frac{\sum x}{n} = \frac{55}{10} = \boxed{5.5}$

2) Range $= 11 - 1 = \boxed{10}$ 3) Midrange $= \frac{11+1}{2} = \boxed{6}$

4) $s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{10(379) - 55^2}{10(10-1)}$

$= \frac{765}{90} = \boxed{8.5}$ $\frac{17}{2}$

8.5 MATH (1:▶Frac Enter

Sep 5-10:00 AM

$\bar{x}$ → Sample Mean

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

s^2 → Sample Variance

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

s → Sample Standard deviation

$\boxed{s = \sqrt{s^2}}$ $\boxed{s \geq 0}$ Always

Sep 5-10:06 AM

Consider the Sample below

2 3 5 7 9

1) $n = \boxed{5}$

2) Mode None

3) Range $= 9-2 = \boxed{7}$

4) Midrange $= \frac{9+2}{2} = 5.5$

5) $\Sigma x = 26$

6) $\Sigma x^2 = 168$

7) $\bar{x} = \frac{\Sigma x}{n} = \frac{26}{5} = \boxed{5.2}$

8) $s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{5 \cdot 168 - 26^2}{5(5-1)} = \frac{164}{20} = \boxed{8.2}$

9) $s = \sqrt{s^2} = \sqrt{8.2} \approx \boxed{2.864}$

whole → 3

1-Dec. → 2.9

2-Dec. → 2.86

Sep 5-10:09 AM

Given $n = 8$ $\Sigma x = 32$ $\Sigma x^2 = 128$

1) $\bar{x} = \frac{\Sigma x}{n} = \frac{32}{8} = \boxed{4}$

2) $s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{8 \cdot 128 - 32^2}{8(8-1)} = \frac{0}{56} = \boxed{0}$

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

Google what it means when standard deviation is 0.

Sep 5-10:18 AM

How to estimate S:

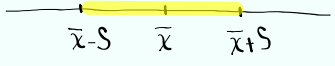
$$S \approx \frac{\text{Range}}{4}$$

the range rule-of-thumb

Sep 5-10:34 AM

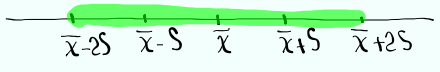
Empirical Rule:

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



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

Usual Range



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

Empirical Rule is best when data distribution is symmetric and bell-shape.

Sep 5-10:36 AM

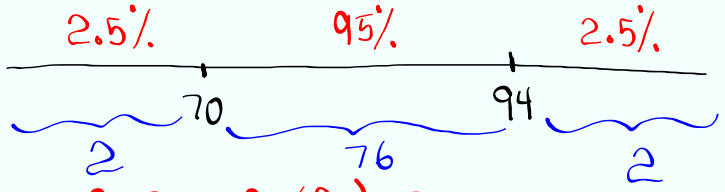
Suppose I randomly selected 80 exams

Scores were Symmetric with $\bar{x} = 82$ and $S = 6$.

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

95% Range $\Rightarrow \bar{x} \pm 2S = 82 \pm 2(6) \rightarrow$ 70 to 94

Usual Range



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

Sep 5-10:42 AM

Z - Score

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

Always round to 3-dec. Places

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

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

Sep 5-10:48 AM

I graded exam 1, $\bar{x} = 82$, $S = 8$

Lisa got 95 → Is that usual or unusual? explain

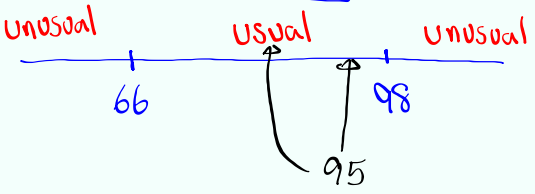
$$Z = \frac{x - \bar{x}}{S} = \frac{95 - 82}{8} = \frac{13}{8} = 1.625$$

Since $-2 \le Z \le 2$ → usual

Usual Range

$\bar{x} \pm 2S = 82 \pm 2(8) = 82 \pm 16$

⇒ 66 - 98



Sep 5-10:50 AM

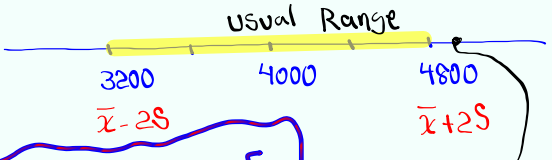
Lisa makes $5000/mo. as a nurse.

Salaries of nurses $\bar{x} = 4000$, $S = 400$

$$Z = \frac{x - \bar{x}}{S} = \frac{5000 - 4000}{400}$$

$$= \frac{1000}{400} = 2.5$$

Since $Z > 2$ ⇒ her Salary is unusual.



You can work on SG 5

Sep 5-10:56 AM