Statistics

Lecture 3

Feb 19-8:47 AM

Complete the chart below:

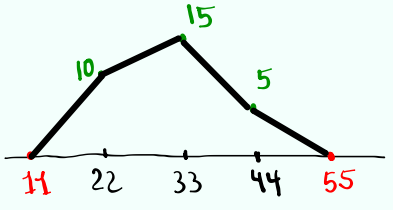
class limits	class BNDRS	class F	Cum. F	Rel. F	class MP
17 - 27	16.5 - 27.5	10	10	.333	22
28 - 38	27.5 - 38.5	15	25	.500	33
39 - 49	38.5 - 49.5	5	30	.167	44

3 - classes , CW = 11 , n = 30

Rel. F = $\frac{f}{30}$

Freq. Polygon

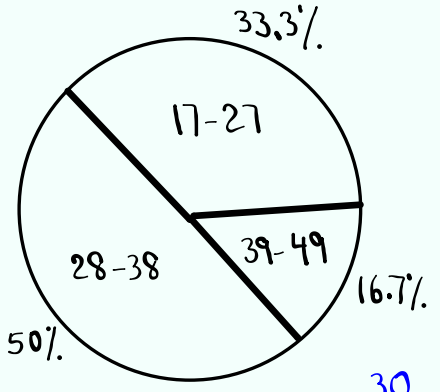
- class MP
- class F
- additional MP, one on each side



Jan 8-4:31 PM

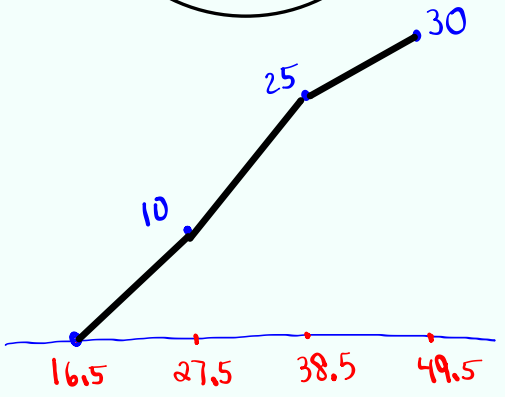
Pie chart

- Circle
- class limits
- % F



Ogive

- class BNDRS
- Cum. F



Jan 8-4:40 PM

Consider the Sample below

4 6 6 8 8 10

1) $n = 6$

2) Range $= 10-4 = 6$

3) Midrange $= \frac{10+4}{2} = 7$

4) Mode $= 6 \& 8$

5) Median $= \frac{6+8}{2} = 7$

6) $\Sigma x = 42$

7) $\Sigma x^2 = 316$

8) $\bar{x} = \frac{\Sigma x}{n} = \frac{42}{6} = 7$ ← Sample Mean

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

$= \frac{6 \cdot 316 - 42^2}{6(6-1)}$

$= \frac{132}{30} = 4.4$

10) $s = \sqrt{s^2} = \sqrt{4.4} \approx 2.098$ ← Sample standard deviation

Jan 8-4:45 PM

More on s^2 & s:

1) $s^2 \geq 0$

When $s^2 = 0$, All data elements are equal to $\bar{x}$.

2) $s \geq 0$

When $s = 0$, all data elements are equal to $\bar{x}$.

When s is small, data elements are close to $\bar{x}$.

When s is big, data elements are more spread out from $\bar{x}$.

Standard deviation indicates how data elements are spread from the mean.

Jan 8-4:53 PM

Consider the Sample below

1 1 2 3 3

$n = 5 \qquad \Sigma x = 10 \qquad \Sigma x^2 = 24$

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

$s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{5 \cdot 24 - 10^2}{5(5-1)} = \frac{20}{20} = 1$

$s = \sqrt{s^2} = \sqrt{1} = \boxed{1}$

Now Consider

1 1 2 3 30

$n = 5 \qquad \Sigma x = 37 \qquad \Sigma x^2 = 915$

$\bar{x} = \frac{37}{5} = 7.4 \qquad s^2 = 160.3 \qquad s = 12.661$

Jan 8-4:58 PM

Empirical Rule:

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 $\bar{x} \pm 3s$

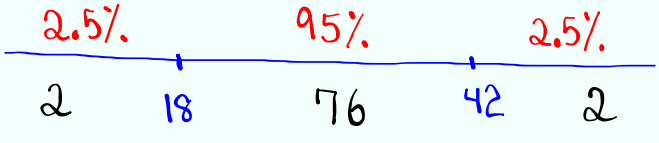
Jan 8-5:04 PM

I randomly selected 80 students. their mean age was 30 with standard deviation 6. $n=80$, $\bar{x}=30$, $s=6$

68% Range → $\bar{x} \pm s = 30 \pm 6$ → 24 to 36

95% Range → $\bar{x} \pm 2s = 30 \pm 2(6)$

Usual Range $= 30 \pm 12$ → 18 to 42



95% of 80 = .95(80)

Jan 8-5:07 PM

Salaries of randomly selected nurses had mean of $6200 with standard deviation of $400. $\bar{x} = 6200$, $S = 400$

Maria makes $7200/mo. Is that usual salary?

Usual Range

95% Range

$\bar{x} \pm 2S =$

$6200 \pm 2(400) =$

$6200 \pm 800 =$

5400 to 7000

Maria's salary is unusually high.

Jan 8-5:13 PM

Z - Score

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

Always round to 3-decimal places.

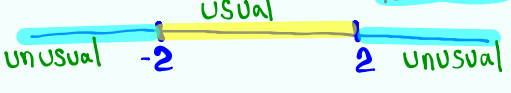
It is a value that indicates how many standard deviation is the data element above or below the mean.

It is a way to standardize data elements.

It allows us to compare data elements from different samples.

when $-2 \leq Z \leq 2$ → data element is usual.

when $Z < -2$ or $Z > 2$ → data element is unusual.



Jan 8-5:17 PM

Exam I: $\bar{x} = 88$, $S = 5$

Isabella got 95.

$$Z = \frac{x - \bar{x}}{S} = \frac{95 - 88}{5} = \frac{7}{5} = 1.4$$

Since $-2 \leq Z \leq 2$ → her exam score is usual.

Exam II: $\bar{x} = 75$, $S = 8$

Isabella got 92.

$$Z = \frac{x - \bar{x}}{S} = \frac{92 - 75}{8} = 2.125$$

Since $Z > 2$, her exam score is unually high.

Jan 8-5:24 PM

Nurses: $\bar{x} = 6200$, $S = 400$

Maria makes \$6600

Sales: $\bar{x} = 8000$, $S = 500$

John makes \$8500

Who is doing better?

Maria: $Z = \frac{x - \bar{x}}{S} = \frac{6600 - 6200}{400} = 1$

John: $Z = \frac{x - \bar{x}}{S} = \frac{8500 - 8000}{500} = 1$

They are doing the Same.

Jan 8-5:29 PM

I randomly Selected 25 exams, here are the Scores:

58	59	63	68	68
70	72	76	76	76
79	80	83	85	85
85	88	89	92	93
93	95	100	100	100

1) $n = 25$

2) Range $= 100 - 58 = 42$

3) Estimate s

$s \approx \frac{\text{Range}}{4} = \frac{42}{4} = \boxed{10.5}$

4) Make STEM Plot

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

Jan 8-5:33 PM

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

How many Scores were below 70? 5

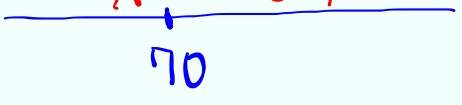
What % of Scores were below 70?

5 is what % of 25?

$5 = \frac{P}{100} \cdot 25$

$P = 20$

20%

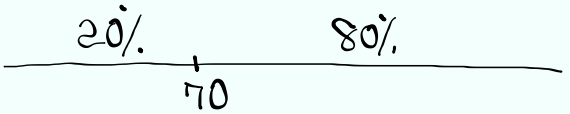


Jan 8-5:40 PM

Percentile :

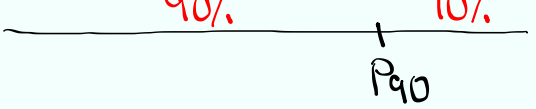
Notation P_K

Requirement Data must be Sorted



$P_{20} = 70$

P_{90}



Jan 8-5:43 PM

How to find P_K

Make Sure data is Sorted

Find location

$$L = \frac{K}{100} \cdot n$$

n: Sample Size

if L is decimal: Round-up

P_K = Lth element

if L is a whole: $P_K = \frac{\text{Lth element} + \text{Next one}}{2}$

Jan 8-5:45 PM

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

Find P_{30}

$L = \frac{30}{100} \cdot 25 = 7.5$

L is decimal → Round-up

$L = 8$

30% | 70%

76

P_{30} = 8th element

= 76

Find P_{80}

$L = \frac{80}{100} \cdot 25 = 20$ ← whole #

80% | 20%

$P_{80} = 93$

$P_{80} = \frac{20\text{th element} + \text{Next one}}{2}$

$= \frac{93+93}{2} = 93$

Jan 8-5:48 PM

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

Find K such that $P_K = 80$

44% | 56%

80

Doing reverse

Find K

Below

$K = \frac{B}{n} \cdot 100$

Sample Size

Round to whole %

$K = \frac{B}{n} \cdot 100 = \frac{11}{25} \cdot 100$

$= 44$

$P_{44} = 80$

Jan 8-5:53 PM

TI Instructions:

1) To clear the screen [Clear]

2) To quit [2nd] [MODE]

3) To clear all lists.

[2nd] [+] [4: Clear All Lists] [Enter]

4) To reset all lists

[STAT] Edit

[5: Set Up Editor] [Enter]

Jan 8-6:10 PM

How to store data in a list.

I want to store the Sample below in L1.

12 18 5 8 10

15 4 19 20 19

[STAT] Edit

[1: Edit]

Let's quit

[2nd] [Mode]

Clear the Screen

[Clear]

Jan 8-6:15 PM

Let's view L1:

2nd 1 Enter

{12 18 5 8 10 19}

How to Sort L1:

STAT Edit

2:SortA(2nd 1 Enter

Let's view L1:

2nd 1 Enter {4 5 8 10 12

Jan 8-6:20 PM

How to find $\bar{x}$ & S:

STAT → CALC

1:1-Var Stats

with Menu

List: L1

2nd 1

No Menu

L1 Enter

FreqList: clear

Calculate

$\bar{x}=13$

$\sum x=130$

$\sum x^2=2020$

$S=S_x=6.055$

$n=10$

Min = 4

$Q_1=8$

Median = 13.5

$Q_3=19$

Max = 20

5-number Summary

Jan 8-6:26 PM

How to find S^2:

VARS VARS xy χ^2

5: Statistics 3: S_x Enter

$S^2 = 36.66666\ldots$

Convert to reduced fraction

Math 1: ▶Frac Enter $S^2 = \frac{110}{3}$

Clear the Screen clear

clear all lists 2nd + 4: clear all lists

Enter

Jan 8-6:31 PM

I randomly selected 12 students. Here are their ages.

24	32	28	18
20	30	35	19
21	40	34	29

1) Store in L1

STAT Edit

1: Edit

quit & clear Screen

2nd Mode clear

L1
24
32
28
⋮
29

Jan 8-6:38 PM

Sort L1, then view it.

STAT Edit

2:SortA(2nd 1 Enter

to view it

2nd 1 Enter

{18 19 20 → → →

STEM Plot

1	8 9
2	0 1 4 8 9
3	0 2 4 5
4	0

Jan 8-6:41 PM

Find $\bar{x}$ & s

STAT → CALC

1:1-Var Stats

List: L1

FreqList: clear

Calculate

2nd 1

No Menu

L1 Enter

$\bar{x} = 27.5$

$\sum x = 330$

$\sum x^2 = 9632$

$s = s_x = 7.116$

$n = 12$

5-Number Summary

Min = 18

$Q_1 = 20.5$

Med. = 28.5

$Q_3 = 33$

Max = 40

Jan 8-6:45 PM

Find S^2 in reduced fraction.

VARS → 5: Statistics → 3: S_x

χ^2 → Enter

$S^2 = 50.\overline{63}$

Math → 1: ▶Frac → Enter

$S^2 = \frac{557}{11}$

Jan 8-6:50 PM

How to find $\bar{x}$ & S from a grouped data

freq. table

class limits	class MP	class F
12 - 20	16	4
21 - 29	25	10
30 - 38	34	6

1) clear all lists.
2nd → + → 4: clear All Lists → Enter

2) class MP → L1
class F → L2

3) STAT → CALC
1: 1-Var Stats
List: L1
Freq List: L2
Calculate

No Menu
L1 , L2
,
Enter

$\bar{x} = 25.9$
$S = S_x = 6.464$
$n = 20$

Jan 8-6:56 PM

Find S^2 in reduced fraction

VARS | 5: Statistics | 3: S_x

x^2 | Math | 1: ▶Frac | Enter

$$S^2 = \frac{3969}{95}$$

Jan 8-7:06 PM

Class QZ 1

Consider the Sample below

2 4 6 **6** 8 9 10

1) $n = 7$

2) Range $= 10 - 2 = 8$

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

4) Mode $= 6$

5) Median $= 6$

Jan 8-7:09 PM