Statistics

Lecture 3

Feb 19-8:47 AM

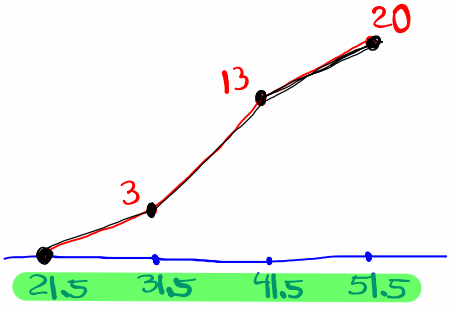
Class QZ 2

Complete the chart below, then draw ogive.

class BNDRS	class F	Cum.F
21.5 – 31.5	3	3
31.5 – 41.5	10	13
41.5 – 51.5	7	20

CW = 31.5 − 21.5 = 10

n = 20



Feb 13-9:12 PM

$x \rightarrow$ Data element

$\sum x \rightarrow$ Sum of data elements

$\sum x^2 \rightarrow$ Sum of data elements^2

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

$s^2 \rightarrow$ Sample Variance

$s \rightarrow$ Sample Standard deviation

$n \rightarrow$ Sample Size

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

$s = \sqrt{s^2}$

Consider the Sample 2, 3, 5, 5, 8

$n = 5$ $\quad \sum x = 2+3+5+5+8 = 23$

$\sum x^2 = 2^2+3^2+5^2+5^2+8^2 = 127$

$\bar{x} = \frac{\sum x}{n} = \frac{23}{5} = \boxed{4.6}$

$s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{5 \cdot 127 - 23^2}{5(5-1)} = \frac{106}{20} = \boxed{5.3}$

$s = \sqrt{s^2} = \sqrt{5.3} \approx \boxed{2.3}$

Feb 20-6:55 PM

Suppose $n = 8$, $\sum x = 28$, $\sum x^2 = 142$, Min $= 1$, Max $= 9$.

1) Range $=$ Max $-$ Min $= 9 - 1 = \boxed{8}$

2) Midrange $= \frac{\text{Max} + \text{Min}}{2} = \frac{9+1}{2} = \boxed{5}$

3) $\bar{x} = \frac{\sum x}{n} = \frac{28}{8} = \frac{14}{4} = \frac{7}{2} = \boxed{3.5}$

4) $s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{8 \cdot 142 - 28^2}{8(8-1)} = \frac{352}{56} \approx \boxed{6.286}$

5) $s = \sqrt{s^2} = \sqrt{6.286} \approx \boxed{2.507}$

The Range Rule-of-Thumb

To estimate s

$s \approx \frac{\text{Range}}{4}$ $\quad s \approx \frac{8}{4} = \boxed{2}$

Feb 20-7:04 PM

Standard deviation S is a non-negative number. S indicates how data elements are spread with respect to Sample mean $\bar{x}$.

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

If S is large $\Rightarrow$ Data elements are more spread out from $\bar{x}$.

If S is Zero $\Rightarrow$ there is no deviation from $\bar{x}$. All data elements are equal to $\bar{x}$.

Feb 20-7:12 PM

Consider the Sample below

2, 3, 3, 3, 4

$n = 5$ $\quad \sum x = 15$ $\quad \sum x^2 = 47$

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

$$S^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{5 \cdot 47 - 15^2}{5(5-1)}$$

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

$$S = \sqrt{S^2} = \sqrt{.5} \approx .707$$

S is fairly small, data elements are close to $\bar{x}$.

Feb 20-7:17 PM

Consider the Sample below

2, 3, 3, 3, 29

$n = 5$ $\quad \Sigma x = 40$ $\quad \Sigma x^2 = 872$

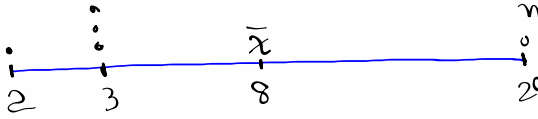
$\bar{x} = \frac{\Sigma x}{n} = \frac{40}{5} = \boxed{8}$ $\quad s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{5 \cdot 872 - 40^2}{5(5-1)}$

$= \frac{2760}{20} = \boxed{138}$

$s = \sqrt{s^2} = \sqrt{138} \approx 11.747$

$\approx \boxed{12}$

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



Feb 20-7:23 PM

Consider the Sample below

4 4 4 4 4

$n = 5$ $\quad \Sigma x = 20$ $\quad \Sigma x^2 = 80$

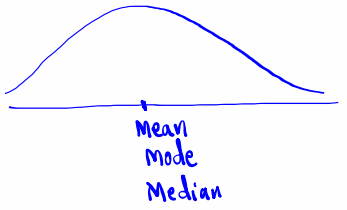
$\bar{x} = \frac{\Sigma x}{n} = \frac{20}{5} = \boxed{4}$ $\quad s^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{5 \cdot 80 - 20^2}{5(5-1)}$

$s = \sqrt{s^2} = \sqrt{0} = \boxed{0}$ $\quad = \frac{0}{20} = \boxed{0}$

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

Feb 20-7:29 PM

whenever mean = mode = Median, data dist. is symmetric and bell-shaped.



Empirical Rule:

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

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

Usual Range

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

Feb 20-7:36 PM

I randomly selected 300 students, Dist. of their ages had a bell-shape graph with $\bar{x} = 34$ and $s = 6$.

68% Range ⇒ $\bar{x} \pm s = 34 \pm 6$ ⇒ 28 to 40

68% of 300 ⇒ 204 students age are between 28 and 40.

Usual Range

95% Range ⇒ $\bar{x} \pm 2s = 34 \pm 2(6)$ ⇒ 22 to 46

95% of 300 ⇒ 285 of students are between 22 and 46

32% are above 40 or below 28.

How many of them were below 40 yrs old?

48 + 204 = 252

Feb 20-7:43 PM

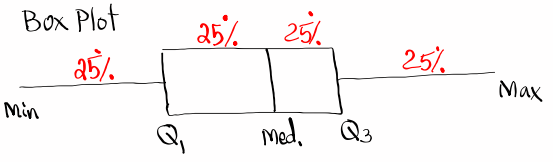
5 – Number Summary

Min. Q_1 Median Q_3 Max.

Q_1: First Quartile

Q_3: Third Quartile

Draw Box Plot



IQR (Inter-Quartile-Range) : $Q_3 - Q_1$

Upper Fence = $Q_3 + 1.5(IQR)$

Lower Fence = $Q_1 - 1.5(IQR)$

Anything below LF or above UF is called outlier.

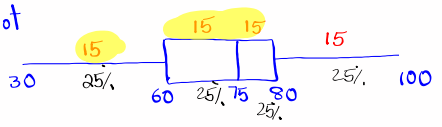
Feb 20-8:06 PM

I randomly selected 60 exams. Here are the 5-Number Summary of the Scores

60 ÷ 4 = 15

30 (Min.) 60 (Q_1) 75 (Median) 80 (Q_3) 100 (Max)

Box Plot



$IQR = Q_3 - Q_1 = 80 - 60 = \boxed{20}$

Upper Fence = $Q_3 + 1.5(IQR) = 80 + 1.5(20) = \boxed{110}$

Lower Fence = $Q_1 - 1.5(IQR) = 60 - 1.5(20) = \boxed{30}$

Any Score below 30 or above 110 is outlier.

None 30 (Min, LF) 100 (Max) 110 (UF) None

No outliers

How many Scores were below 80? 45

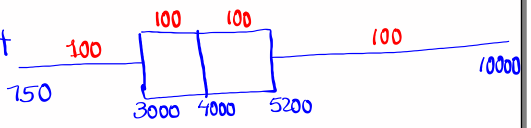
Feb 20-8:12 PM

I randomly selected 400 nurses. Below is 5-Number Summary of their monthly Salary.

400 ÷ 4 = 100

750	3000	4000	5200	10000
Min.	Q_1	Median	Q_3	Max

Box Plot

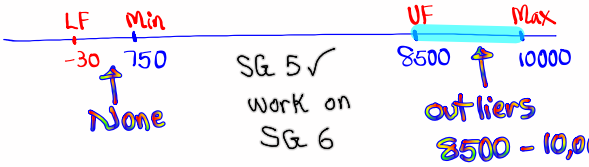


$IQR = Q_3 - Q_1 = 5200 - 3000 = 2200$

$UF = Q_3 + 1.5(IQR) = 5200 + 1.5(2200) = 8500$

$LF = Q_1 - 1.5(IQR) = 3000 - 1.5(2200) = -300$

below -300 or above 8500 $\Rightarrow$ outliers



None

SG 5 ✓

work on SG 6

outliers

8500 - 10,000

Feb 20-8:21 PM

TI Instructions

1) Clear the Screen. [Clear]

2) To quit. [2nd] [Mode]

3) To clear all lists. [2nd] [+] [4: ClearAllLists] [Enter]

4) To reset all lists. [STAT] Edit [5: Set up Editor] [Enter]

Let's quit & clear the Screen [2nd] [Mode] [clear]

Feb 20-8:30 PM

How to store data elements in a list

Suppose we have the following Sample

15 20 18 30 25 24 32 28 30

STAT Edit

1:Edit

L1
15 Enter
20 "
18 "
⋮
30 "

Let's quit

2nd Mode

How to view L1:

2nd 1 Enter

{15 20 18 30 . . . 30}

Feb 20-8:39 PM

How to Sort data using TI:

STAT Edit

2: SortA(

2nd 1 (L1) Enter

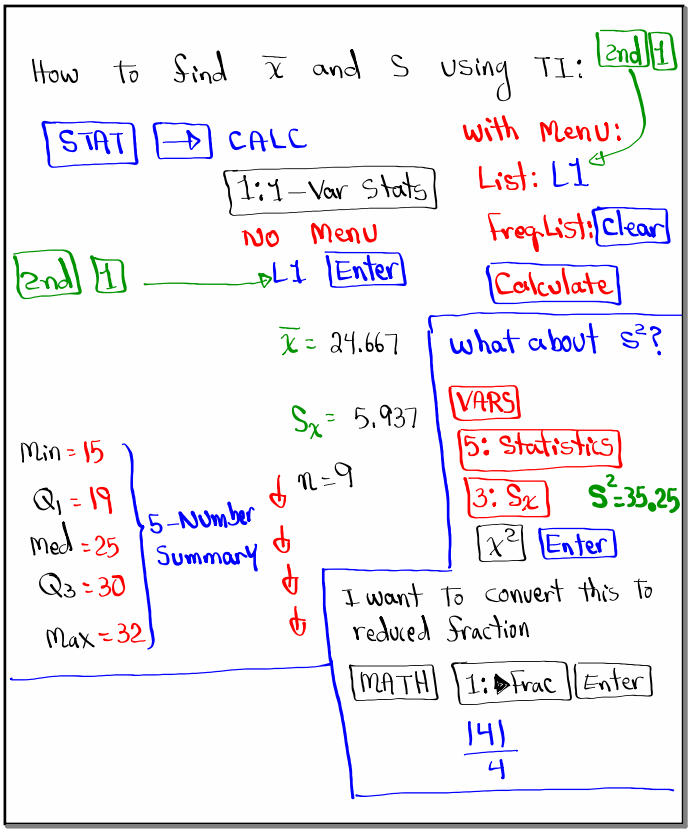
Let's quit & clear the Screen

2nd Mode clear

Let's view L1:

2nd 1 (L1) Enter {15 18 20 24 . . . 32}

Feb 20-8:46 PM



Feb 20-8:56 PM

Clear all lists.

2nd + 4:clear all lists Enter

Store the following in L1.

STAT Edit
1:Edit

72	58	80	75	90
60	65	100	95	88
78	98			

L1
72
58
80
75
⋮
98

Let's quit 2nd MODE

Feb 20-9:09 PM

Find $\bar{x}$ & S:

STAT → CALC

1: 1-Var Stats

NO Menu

L1 Enter

2nd 1 with Menu

List: L1

freqList: clear

Calculate

Min = 58

Q_1 = 68.5

Med = 79

Q_3 = 92.5

Max = 100

$\bar{x}$ = 79.917

S = 14.507

n = 12

Find S^2

VARS

5: Statistics

3: S_x

x^2 Enter

210.4469697...

Convert to reduced fraction.

MATH 1: ▶Frac Enter

$S^2 = \frac{27779}{132}$

SG 5 ✓

Work on

SG 6, 7, and 8

watch videos on the RHS of Study guides

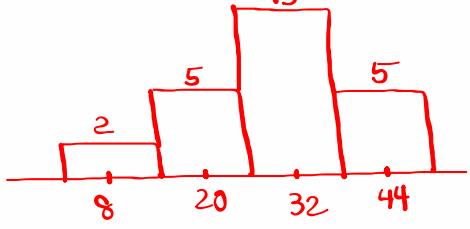
Feb 20-9:13 PM

Class QZ 3

Consider the table below

class MP	class F
8	2
20	5
32	13
44	5

1) Draw histogram



2) Sample Size

$n = 2 + 5 + 13 + 5 = 25$

Feb 20-9:26 PM