Statistics

Lecture 3

Feb 19-8:47 AM

§6.5

Consider the Sample below

2 4 4 4 8

1) $n = 5$

2) Range $= 8 - 2 = 6$

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

4) Mode $= 4$

5) Median $= 4$

it is a value in the middle

(Data is Sorted)

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

$\sum x \rightarrow$ Sum of data elements

$\sum$ Summation

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

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

$\sum x = 2 + 4 + 4 + 4 + 8 = 22$

$\bar{x} = \frac{\sum x}{n} = \frac{22}{5} = 4.4$

Round to whole # 4

Round up to whole 5

Sep 12-11:35 AM

$x \rightarrow$ Data element

$n \rightarrow$ Sample Size

$\sum x \rightarrow$ Sum of data elements

$\sum x^2 \rightarrow$ Square data elements, then add

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

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

we use Variance to find Sample Standard deviation

Variance is the Sum of Squares of the difference between data elements and $\bar{x}$, then divide by $n-1$.

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

Standard deviation is Square root of Variance $s = \sqrt{s^2}$

$s \geq 0$

Sep 12-11:42 AM

Consider the Sample below

1 4 4 6 6 9

$n = 6$ Range $= 9 - 1 = 8$

Midrange $= \frac{9+1}{2} = 5$ Mode: 4 & 6

Median

value in the middle $\frac{4+6}{2} = \boxed{5}$

Data must be Sorted

$\sum x = 1 + 4 + 4 + 6 + 6 + 9 = \boxed{30}$

$\sum x^2 = 1^2 + 4^2 + 4^2 + 6^2 + 6^2 + 9^2 = \boxed{186}$

$\bar{x} = \frac{\sum x}{n} = \frac{30}{6} = \boxed{5}$ $\qquad s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)}$

$s = \sqrt{s^2}$

$= \sqrt{7.2} \approx \boxed{2.683}$ $\qquad = \frac{6 \cdot 186 - 30^2}{6(6-1)} = \frac{216}{30}$

$= \boxed{7.2}$ $\qquad s^2 \geq 0$

Sep 12-11:50 AM

what is standard deviation?

It is a value that indicates how spread the data elements are from $\bar{x}$.

$S \geq 0$

If S is Small → Data elements are close to $\bar{x}$.

If S is big → data elements are more spread out from the mean.

If S is Zero → all data elements are equal to $\bar{x}$.

(No deviation)

Sep 12-11:59 AM

Consider the Sample below

1 2 2 2 3

$n = 5$ Mode = 2 Median = 2

$\sum x = 1 + 2 + 2 + 2 + 3 = 10$

$\sum x^2 = 1^2 + 2^2 + 2^2 + 2^2 + 3^2 = 22$

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

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

$= \frac{5 \cdot 22 - 10^2}{5(5-1)} = \frac{10}{20} = \boxed{.5}$

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

$\approx \boxed{.707}$

Since S is Small, data elements are close to $\bar{x}$.

Less deviation from $\bar{x}$.

Sep 12-12:03 PM

Consider the Sample below

1 2 2 2 93

$n=5$ Mode = 2 Median = 2

$\sum x = 1+2+2+2+93 = \boxed{100}$

$\sum x^2 = 1^2+2^2+2^2+2^2+93^2 = \boxed{8662}$

$\bar{x} = \frac{\sum x}{n} = \frac{100}{5} = \boxed{20}$

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

$= \frac{5\cdot 8662 - 100^2}{5(5-1)}$

$= \frac{33310}{20} = 1665.5$

$S = \sqrt{S^2}$

$= \sqrt{1665.5} \approx \boxed{40.811}$

Since S is a big number,
data elements are more Spread out from $\bar{x}$.

Sep 12-12:10 PM

Given $n=10$ $\sum x = 80$ $\sum x^2 = 640$

1) $\bar{x} = \frac{\sum x}{n} = \frac{80}{10} = \boxed{8}$

2) $S^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{10\cdot 640 - 80^2}{10(10-1)} = \frac{0}{90} = \boxed{0}$

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

4) what do you conclude about data elements? explain.

Since $S=0$, All data elements are the same as $\bar{x} = \boxed{8}$

How to estimate S:

$S \approx \frac{\text{Range}}{4}$ "The range rule-of-thumb"

Sep 12-12:17 PM

Given $n=20$ $\Sigma x=1583$ $\Sigma x^2=128743$

Min $=52$ Max $=100$

Find

1) $\bar{x} = \frac{\Sigma x}{n} = \frac{1583}{20} = 79.15$

2) $S^2 = \frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)} = \frac{20 \cdot 128743 - 1583^2}{20(20-1)} = \frac{68971}{380}$

3) $S = \sqrt{S^2} = \sqrt{181.503} \approx 13.472 \approx 13 \approx 181.503$

4) Estimate $S \approx \frac{\text{Range}}{4} = \frac{100-52}{4} = \frac{48}{4} = 12$

Sep 12-12:26 PM

whenever Mean, Mode, and median are the Same, data dist. will be symmetric and bell-shape looking.

Empirical Rule (68-95-99.7) Ranges

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

About 95% - - - - - - - - - $\bar{x} \pm 2S$

Usual Range

About 99.7% - - - - - - - - - $\bar{x} \pm 3S$

Sep 12-12:32 PM

I randomly selected 120 teachers, their monthly Salaries had a Symmetric dist. with $\bar{x} = \$8200$ and $S = \$500$.

68% Range $\rightarrow \bar{x} \pm S = 8200 \pm 500$

$\Rightarrow$ 7700 to 8700

68% (120) $\approx$ 82

82 teachers have Salaries

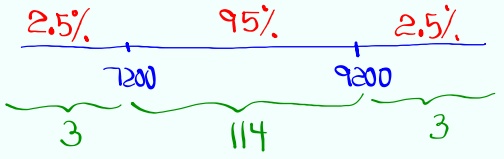
Usual Range "95% Range"

$\bar{x} \pm 2S$

$= 8200 \pm 2(500)$

$\Rightarrow$ 7200 to 9200

95% (120) = 114



Sep 12-12:36 PM

Z-Score Always round to 3 dec. Places

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

Z-Score indicates how many Standard deviation is the data element from the mean.

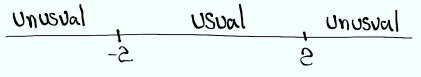
$Z > 0$ above $\bar{x}$

$Z < 0$ below $\bar{x}$

$Z = 0$ You are the mean.

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

$Z < -2$ or $Z > 2$ $\rightarrow$ Unusual data element



Z-Score allows us to Compare data elements from different Samples

It is a way to Standardize data elements with respect to $\bar{x}$.

Sep 12-12:43 PM

Lisa makes $4000/mo. as a nurse.

Mike " $6000/mo. as a manager.

Who is performing better?

Nurse $\bar{x} = 3500$, $S = 400$

manager $\bar{x} = 5000$, $S = 1000$

Usual

Lisa $z = \frac{x - \bar{x}}{S} = \frac{4000 - 3500}{400} = \frac{500}{400} = 1.25$

Usual

Mike $z = \frac{x - \bar{x}}{S} = \frac{6000 - 5000}{1000} = \frac{1000}{1000} = 1$

I work as a manager with z-Score of -2.2. What is my Salary?

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

$$-2.2 = \frac{x - 5000}{1000}$$

Cross-Multiply

$x - 5000 = -2.2(1000)$

isolate x

$x = 5000 - 2.2(1000)$

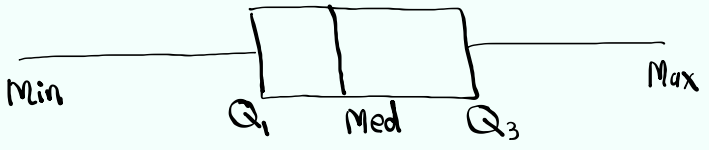
$x = 2800$

Sep 12-12:49 PM

5 - Number Summary

Min Q_1 Med. Q_3 Max

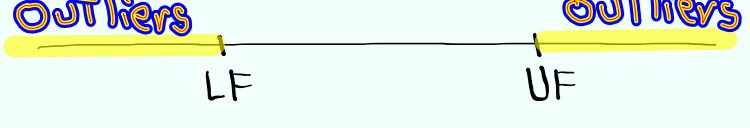
Box Plot



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

upper fence $= Q_3 + 1.5(IQR)$

Lower fence $= Q_1 - 1.5(IQR)$



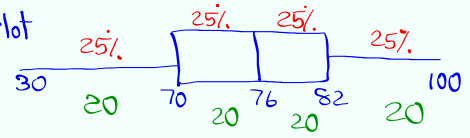
Sep 12-12:56 PM

I randomly selected 80 exams. 80÷4=20

The 5-Number Summary of Scores were

30	70	76	82	100
Min	Q_1	Med	Q_3	Max

Box Plot



$IQR = Q_3 - Q_1 = 82 - 70 = 12$

Upper Fence $= Q_3 + 1.5(IQR)$

$= 82 + 1.5(12) =$ **100**

Lower Fence $= Q_1 - 1.5(IQR)$

$= 70 - 1.5(12) =$ **52**

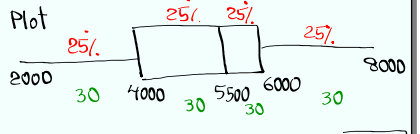
outliers

Sep 12-1:01 PM

5-Number Summary of monthly Salares of 120 nurses were 120÷4=30

2000	4000	5500	6000	8000

Box Plot



$IQR = Q_3 - Q_1 = 6000 - 4000 =$ **2000**

Upper Fence $= Q_3 + 1.5(IQR)$

$= 6000 + 1.5(2000) =$ **9000**

Lower Fence $= Q_1 - 1.5(IQR)$

$= 4000 - 1.5(2000) =$ **1000**

SG 5 ✓

NO outliers

You can work on SG 6.

Look for my email

Look for couple of online QZ.

Sep 12-1:19 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 [Enter]

5: Setup Editor

Sep 12-1:27 PM

How to store data elements in a list:

Consider the Sample below

8	12	5	20	15
10	5	14	15	10

[STAT] Edit

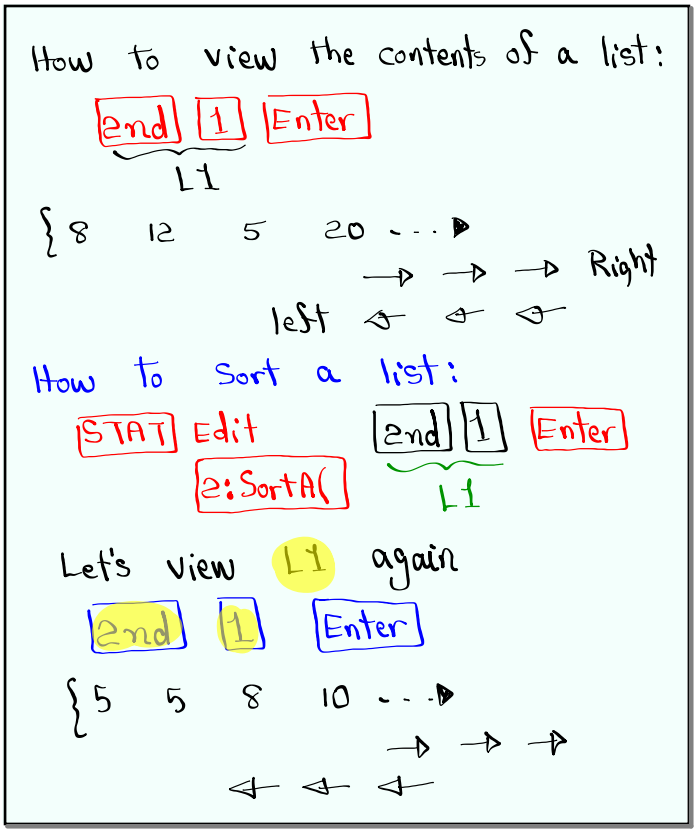
1: Edit

L1
8 [enter]
12 "
5 "
⋮
10 "

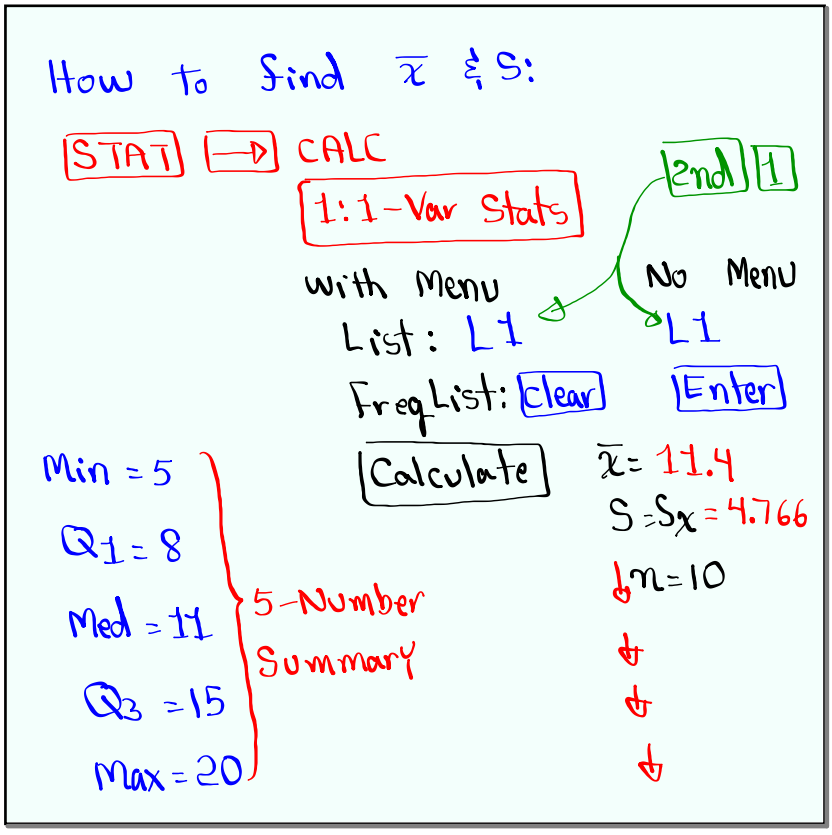
Let's quit & clear Screen

[2nd] [MODE] [Clear]

Sep 12-1:33 PM



Sep 12-1:37 PM



Sep 12-1:43 PM

what about s^2?

VARS | 5: Statistics | 3: S_x | x^2 | Enter

$$s^2 = 22.7\overline{1}$$

Convert to reduced fraction

MATH | 1: ▶Frac | Enter

$$s^2 = \frac{1022}{45}$$

Clear all lists

2nd | + | 4: clear all lists | Enter

Sep 12-1:49 PM

I randomly selected 20 students, here are their ages:

23	27	18	20	30
35	32	25	19	40
48	50	45	38	19
24	25	50	46	44

1) Store in L1

STAT | Edit | 1: Edit

L1

23 Enter

27 Enter

⋮

44

quit & clear Screen

2nd | MODE

Clear

Sep 12-1:55 PM

Sort L1, View L1, then make STEM plot

STAT Edit
2:SortA(

2nd 1 Enter

2nd 1 Enter

{18 19 19 20

1	8 9 9
2	0 3 4 5 5 7
3	0 2 5 8
4	0 4 5 6 8
5	0 0

Sep 12-2:00 PM

Find $\bar{x}$ & s

STAT → CALC
1:1-Var Stats

with Menu
List:L1
FreqList: clear
Calculate

2nd 1
No Menu
L1
Enter

Min = 18
Q_1 = 23.5
Med = 31
Q_3 = 44.5
Max = 50

$\bar{x}$ = 32.9
$S = S_x$ = 11.369
n = 20

Sep 12-2:04 PM

Find S^2 in reduced fraction:

VARS | 5: Statistics | 3: S_x | x^2 | Math

1: ▶Frac | Enter

$$S^2 = \frac{12279}{95}$$

Sep 12-2:09 PM

Complete the table below

class limits	class MP	class F	Cum. F
18 - 30	24	7	7
31 - 43	37	13	20
44 - 56	50	12	32
57 - 69	63	8	40

+13

clear all lists

class MP → L1 class F → L2

L1	L2
24	7
37	13
50	12
63	8

STAT → CALC 1: 1-Var Stats

with Menu

End 1 List: L1

End 2 FreqList: L2

Calculate

L1, L2

Enter

7

Sep 12-2:12 PM

$\bar{x} = 43.825$

$S = S_x = 13.161$

$n = 40$

$s^2 = \frac{6929}{40}$ ✓

VARS

5: Statistics

3: S_x χ^2

Math 1: ▶Frac

Enter

Sep 12-2:20 PM