Statistics

Lecture 3

Feb 19-8:47 AM

Class Quiz 1

Consider the Sample below

2 5 5 8 14 18

1) Range = $18-2=16$ ✓

2) Midrange = $\frac{18+2}{2}=10$ ✓

3) Mode = 5 ✓

4) $n=6$ ✓

5) $\Sigma x = 2+5+5+8+14+18 = 52$ ✓

6) $\Sigma x^2 = 2^2+5^2+5^2+8^2+14^2+18^2 = 638$ ✓

7) $\frac{\Sigma x}{n} = \frac{52}{6}$

$= 8.\overline{6}$

whole 9

1-dec. 8.7

2-dec. 8.67

3-dec. 8.667

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

$= \frac{6\cdot 638-52^2}{6(6-1)}$

$= \frac{1124}{30} = 37.4\overline{6}$

whole 37

1-dec. 37.5

2-dec. 37.47

3-dec. 37.467

9) $\sqrt{\text{Last Ans}} = \sqrt{37.467}$

≈ 6.871

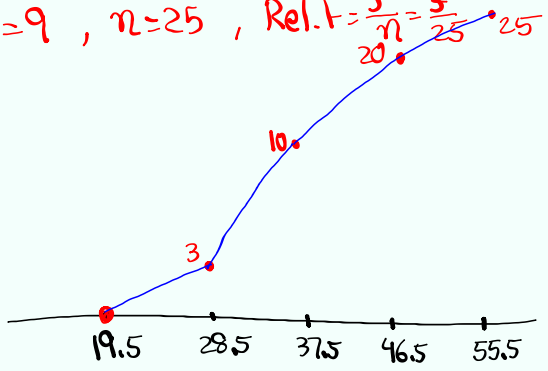
Sep 6-2:26 PM

Complete the chart below

Class limits	Class BNDRS	Class MP	Class F	Cum. F	Rel. F	% F
20 – 28	19.5 – 28.5	24	3	3	.12	12%
29 – 37	28.5 – 37.5	33	7	10	.28	28%
38 – 46	37.5 – 46.5	42	10	20	.40	40%
47 – 55	46.5 – 55.5	51	5	25	.20	20%

4 classes, CW = 9, $n = 25$, Rel. F = $\frac{F}{n} = \frac{F}{25}$ 25

ogive
- class BNDRS
- Cum. F
- start at 0 level



Sep 13-11:44 AM

SG 5-8

x → Data element

$\sum x$ → Sum of data elements

n → Sample Size

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

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

Consider the sample below

1 3 3 3 9

$n = 5$ Mode = 3 Range = 9 − 1 = 8

Midrange = $\frac{9+1}{2} = 5$ $\sum x = 1+3+3+3+9 = 19$

$\bar{x} = \frac{\sum x}{n} = \frac{19}{5} = 3.8$ Round to whole → 4

Sep 13-11:55 AM

Consider the Sample below

2 3 3 3

5 5 5 10

1) $n = 8$ 2) Range = 8 3) Midrange = 6

4) Mode = 3 & 5 5) $\Sigma x = 2+3+3+3+5+5+5+10 = 36$

6) $\bar{x} = \frac{\Sigma x}{n} = \frac{36}{8} = \boxed{4.5}$

Sep 13-11:59 AM

$x \rightarrow$ Data element

$x^2 \rightarrow$ Square data element

$\Sigma x \rightarrow$ Sum of data elements

$\Sigma x^2 \rightarrow$ Sum of square of data elements

$n \rightarrow$ Sample Size

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

$s^2 \rightarrow$ Sample Variance

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

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

Sep 13-12:03 PM

Consider the Sample below

1 3 3 3 5

$n=5$ Mode $=3$ Range $=4$

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

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

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

$= \frac{15}{5} = \boxed{3}$

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

$= \frac{5 \cdot 53 - 15^2}{5(5-1)} = \frac{40}{20} = \boxed{2}$

Sep 13-12:07 PM

Given

$\sum x = 55$ $\sum x^2 = 375$ $n = 10$

Min. $=2$ Max $=12$

1) Range $= 10$

2) Midrange $= 7$

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

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

$= \frac{10 \cdot 375 - 55^2}{10(10-1)}$

$= \frac{725}{90}$

$= 8.0\overline{5}$

Round

Whole 8

1-dec. 8.1

2-dec. 8.06

3-dec. 8.056

Sep 13-12:12 PM

$\bar{x}$ → Sample Mean

s^2 → Sample Variance

s → Sample Standard deviation

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

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

Consider the Sample below

2 4 6 8 10

1) $n = 5$

2) Mode None

3) $\sum x = 30$

4) $\sum x^2 = 220$

5) $\bar{x} = \frac{\sum x}{n} = \frac{30}{5} = \boxed{6}$

6) $s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{5 \cdot 220 - 30^2}{5(5-1)} = \frac{200}{20} = \boxed{10}$

→ $s = \sqrt{s^2} = \sqrt{10} \approx \boxed{3.162}$

Sep 13-12:18 PM

Consider the Sample below

4 4 4 4 4

1) $n = 5$

2) $\sum x = 20$

3) $\sum x^2 = 80$

4) $\bar{x} = \frac{\sum x}{n} = \frac{20}{5} = \boxed{4}$

5) $s^2 = \frac{n \cdot \sum x^2 - (\sum x)^2}{n(n-1)} = \frac{5 \cdot 80 - 20^2}{5(5-1)} = \frac{0}{20} = \boxed{0}$

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

Sep 13-12:25 PM

How to estimate S:

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

Range Rule-of-Thumb

Given $n = 10$, $\sum x = 67$, $\sum x^2 = 519$

Min = 2 Max = 12

$\bar{x} = \frac{\sum x}{n} = \frac{67}{10} = 6.7$ $s^2 = \frac{n\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{10 \cdot 519 - 67^2}{10(10-1)} = \frac{701}{90} \approx 7.789$

$s = \sqrt{s^2} = \sqrt{7.789} \approx 2.791$ Estimate $S \approx \frac{\text{Range}}{4} = \frac{12-2}{4} = \frac{10}{4} = 2.5$

Sep 13-12:30 PM

what is Sample Standard deviation?

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

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

If S is big ⇒ data elements are more spread out from $\bar{x}$.

If $S = 0$ ⇒ All data elements are equal to $\bar{x}$. No deviation from $\bar{x}$.

Sep 13-12:51 PM

When data distribution is symmetric and bell-shape (Mean = Mode = Median)

Empirical Rule

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

95% " " " Usual Range $\bar{x} \pm 2s$

99.7% of data " " $\bar{x} \pm 3s$

Ages of voters have a symmetric dist. with $\bar{x} = 42$ & $s = 6$.

68% Range $\Rightarrow \bar{x} \pm s = 42 \pm 6 \Rightarrow$ 36 to 48

95% Range $\Rightarrow \bar{x} \pm 2s = 42 \pm 2(6) \Rightarrow$ 30 to 54

Usual Range

Sep 13-12:56 PM

I randomly selected 40 exams. Scores had a bell-shape dist. with $\bar{x} = 84$ & $s = 7$.

68% Range $= \bar{x} \pm s = 84 \pm 7 \Rightarrow$ 77 to 91

95% Range $= \bar{x} \pm 2s = 84 \pm 2(7) \Rightarrow$ 70 to 98

Usual Range

What % of scores were above 70? 97.5%

How many scores were unusually low?

2.5% (40) = .025(40) = 1

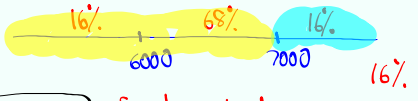
Sep 13-1:04 PM

I randomly selected 80 nurses.
Their Salaries had a Symmetric dist. with $\bar{x} = \$6500$ & $s = \$500$.

1) 68%. Range $= \bar{x} \pm s = 6500 \pm 500 \Rightarrow$ 6000 to 7000

2) 99.7%. Range $= \bar{x} \pm 3s = 6500 \pm 3(500)$
$\Rightarrow$ 5000 To 8000

3) what % of them had a Salary above $7000?

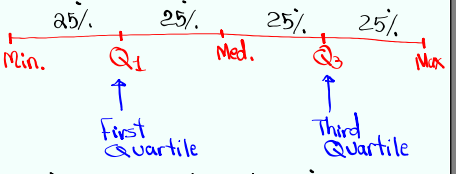


4) How many of them had a Salary below $7000? $84\%(80) = .84(80)$
$= 67.2$
$\approx$ 68

SG 5 ✓

Sep 13-1:12 PM

5 - Number Summary
Min. Q_1 Med. Q_3 Max



50% below Median & 50% above Median
25% below Q_1 & 75% above Q_1
75% below Q_3 & 25% above Q_3

Draw Box Plot

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

Upper Fence $= Q_3 + 1.5(IQR)$

Lower Fence $= Q_1 - 1.5(IQR)$

below LF or above UF
Outliers

Sep 13-1:23 PM

I randomly selected 60 students. The 5-Number Summary of their ages were

$\frac{60}{4} = 15$

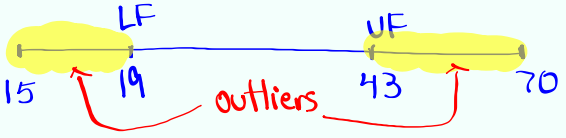
15	28	30	34	70
Min	Q_1	Med.	Q_3	Max

Box Plot

$IQR = Q_3 - Q_1 = 34 - 28 = \boxed{6}$

Upper Fence $= Q_3 + 1.5(IQR) = 34 + 1.5(6) = \boxed{43}$

Lower Fence $= Q_1 - 1.5(IQR) = 28 - 1.5(6) = \boxed{19}$



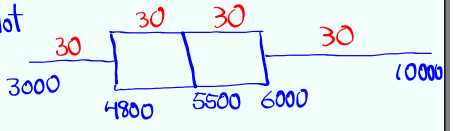
Sep 13-1:30 PM

Salaries of 120 nurses had the following 5-Number Summary:

$\frac{120}{4} = 30$

3000	4800	5500	6000	10,000

1) Draw Box Plot



2) $IQR = Q_3 - Q_1 = 6000 - 4800 = 1200$

3) Upper Fence $= Q_3 + 1.5(IQR) = 6000 + 1.5(1200)$

$= \boxed{7800}$

Lower Fence $= Q_1 - 1.5(IQR) = 4800 - 1.5(1200)$

$= \boxed{3000}$

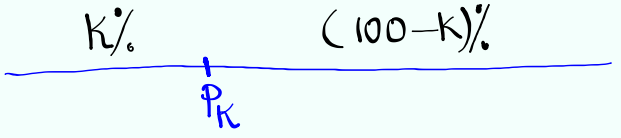
4) Discuss outliers

Sep 13-1:38 PM

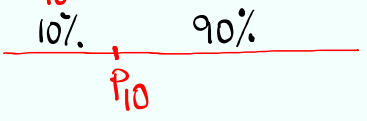
Percentile

Notation P_K

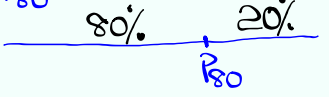
Data must be Sorted



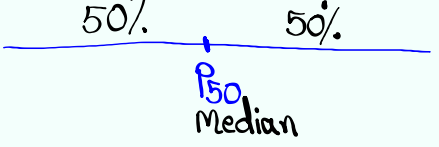
P_{10}



P_{80}



P_{50}



Sep 13-1:45 PM

How to Find P_K:

1) Find location

Sample Size

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

If L is decimal → Round-up

P_K = Lth element

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

Reverse Situation

below

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

Round to whole %

Sep 13-1:59 PM

Consider the Stem Plot below

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

1) $n = 20$

2) P_{40}

40% | 60%

$P_{40} = 50$

$L = \frac{40}{100} \cdot 20 = 8$ ← whole

$P_{40} = \frac{8\text{th} + 9\text{th}}{2} = \frac{48+52}{2} = 50$

3) P_{65}

$L = \frac{65}{100} \cdot 20 = 13$

$P_{65} = \frac{13\text{th} + 14\text{th}}{2} = \frac{56+57}{2} = 56.5$

65% | 35%

$P_{65} = 56.5$

4) P_{88}

$L = \frac{88}{100} \cdot 20 = 17.6 \rightarrow L = 18$

P_{88} = 18th element = 65

88% | 12%

$P_{88} = 65$

Find K such that $P_K = 55$

$K = \frac{B}{n} \cdot 100 = \frac{10}{20} \cdot 100 = 50$ (B = 10 below)

$P_{50} = 55$

Sep 13-2:03 PM

STEM Plot below are for 25 randomly selected exam Scores:

5	2 8
6	0 3 5 5 7
7	2 4 8 8 8 8 9
8	0 2 5 6 8
9	0 2 5 7 8
10	0

1) $n = 25$

2) P_{30}

30% | 70%

$P_{30} = 72$

$L = \frac{30}{100} \cdot 25 = 7.5 \rightarrow L = 8$

P_{30} = 8th element = 72

80% | 20%

$P_{80} = 91$

3) P_{80}

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

$P_{80} = \frac{20\text{th element} + \text{Next one}}{2} = \frac{90+92}{2} = 91$

4) Find K such that $P_K = 80$

$K = \frac{B}{n} \cdot 100 = \frac{14}{25} \cdot 100 = 56$

$P_{56} = 80$

56% | 44%

$P_{56} = 80$

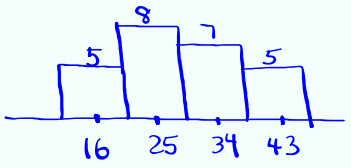
Sep 13-2:12 PM

Class Quiz 2

Complete the table below

class limits	class MP	class F	Cum. F
12 – 20	16	5	5
21 – 29	25	8	13
30 – 38	34	7	20
39 – 47	43	5	25

Draw histogram using class MP & class F.



Sep 13-2:27 PM