Statistics

Lecture 3

Feb 19-8:47 AM

Class Quiz 1:

Consider the Sample below

3 5 5 8 12 15

1) Range $= 15-3 = \boxed{12}$ ✓ 2) Midrange $= \frac{15+3}{2} = \boxed{9}$ ✓

3) Mode $= \boxed{5}$ ✓ 4) Sample Size $\boxed{n=6}$ ✓

$\sum x = 3+5+5+8+12+15 = \boxed{48}$

$\sum x^2 = 3^2+5^2+5^2+8^2+12^2+15^2 = \boxed{492}$

$\frac{\sum x}{n} = \frac{48}{6} = \boxed{8}$ $\frac{n\cdot\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{6\cdot 492 - 48^2}{6(6-1)}$

$= \frac{648}{30} = \boxed{21.6}$

$\sqrt{\text{Last answer}} = \sqrt{21.6}$

≈ 4.648

Round to

whole #	5
1-dec.	4.6
2-dec.	4.65

Sep 6-10:50 AM

Complete the chart below

class limits	class BNDRS	class MP	class F	Cum.F	Rel.F	%F
21 – 30	20.5 – 30.5	25.5	8	8	.32	32%
31 – 40	30.5 – 40.5	35.5	12	20	.48	48%
41 – 50	40.5 – 50.5	45.5	5	25	.20	20%

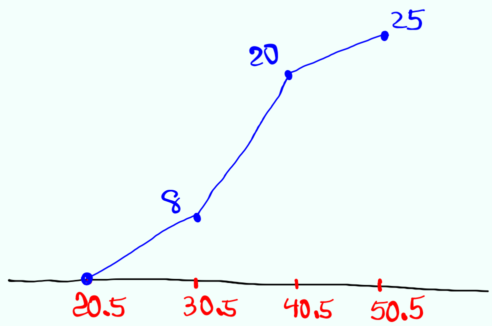
3 classes, CW = 10, $n = 25$, Rel. F. $= \frac{f}{n} = \frac{f}{25}$

Ogive

– Class BNDRS

– Cum. F

– Start at 0 level.



Sep 13-8:11 AM

SG 5–8

$x \to$ Data element

$\sum x \to$ Sum of data elements

$n \to$ Sample Size

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

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

Consider the Sample below

2, 3, 5, 7, 8

$n = 5$ Mode: None $\sum x = 2+3+5+7+8 = \boxed{25}$

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

Sep 13-8:22 AM

Consider the Sample below

2 3 3 3 5 5 5 8

1) $n = 8$

2) Range $= 8 - 2 = 6$

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

4) Mode $= 3 \,\&\, 5$

5) $\sum x = 2+3+3+3+5+5+5+8 = \boxed{34}$

6) $\bar{x} = \frac{\sum x}{n} = \frac{34}{8} = \boxed{4.25}$

Round to

whole $\rightarrow 4$

1-Dec. $\rightarrow 4.3$

Sep 13-8:26 AM

$x \rightarrow$ Data elements

$x^2 \rightarrow$ Square of data elements

$\sum x \rightarrow$ Sum of data elements

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

$n \rightarrow$ Sample Size

$\bar{x} \rightarrow$ Sample Mean

$S^2 \rightarrow$ Sample Variance

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

Sep 13-8:30 AM

Consider the Sample below

2 3 3 3 4

$n=5$ Mode $=3$ Range $=2$ Midrange $=3$

$\sum x = 2+3+3+3+4 = \boxed{15}$

$\sum x^2 = 2^2+3^2+3^2+3^2+4^2 = \boxed{47}$

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

$S^2 = \frac{n\cdot\sum x^2 - (\sum x)^2}{n(n-1)} = \frac{5\cdot 47 - 15^2}{5(5-1)} = \frac{10}{20} = \boxed{.5}$

Sep 13-8:34 AM

Consider the Sample below

1 3 3 5 5 7

1) $n = 6$

2) $\sum x = 24$

3) $\sum x^2 = 118$

4) $\bar{x} = \frac{\sum x}{n} = \frac{24}{6} = \boxed{4}$

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

$= \frac{6\cdot 118 - 24^2}{6\cdot(6-1)} = \frac{132}{30}$

$= \boxed{4.4}$

Sep 13-8:39 AM

$\bar{x} \to$ Sample Mean

$s^2 \to$ Sample Variance

$s \to$ Sample Standard deviation

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

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

Sep 13-8:43 AM

Given $n=10$ $\sum x = 46$ $\sum x^2 = 284$

Min. $=1$ Max $=11$

1) Range $= 11 - 1 = \boxed{10}$

2) Midrange $= \frac{11+1}{2} = 6$

3) $\bar{x} = \frac{\sum x}{n} = \frac{46}{10} = \boxed{4.6}$

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

$= \frac{10 \cdot 284 - 46^2}{10(10-1)}$

$= \frac{724}{90} = 8.0\overline{4}$

≈ 8.044

5) $s = \sqrt{s^2} = \sqrt{8.044} \approx \boxed{2.836}$

whole 3
1-Dec. 2.8
2-Dec. 2.84

Sep 13-8:46 AM

Consider the Sample below

4 4 4 4 4

1) $n = 5$

3) $\sum x = 20$

2) Mode $\begin{cases} \to 4 \\ \to \text{None} \end{cases}$

4) $\sum x^2 = 80$

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

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

$= \frac{5 \cdot 80 - 20^2}{5(5-1)} = \frac{0}{20}$

$= \boxed{0}$

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

Sep 13-8:54 AM

How to estimate Sample Standard deviation:

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

Range rule-of-Thumb

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

Min. = 2 , Max = 12

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

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

$= \frac{725}{90} = 8.0\overline{5} \approx \boxed{8.056}$

3) $S = \sqrt{S^2} \approx \boxed{2.838}$

4) Estimate $S \approx \frac{\text{Range}}{4} = \frac{12-2}{4} = \frac{10}{4} = \boxed{2.5}$

Sep 13-8:59 AM

What is Sample Standard deviation?

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

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

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

If $S = 0$ $\Rightarrow$ All data elements are the Same as $\bar{x}$, No deviation

Sep 13-9:21 AM

When data distribution is Symmetric and bell-Shape (Mean = Mode = Median)

Empirical Rule:

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

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

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

Suppose $\bar{x} = 120$ & $S = 15$

68% Range $\rightarrow$ $\bar{x} \pm S = 120 \pm 15 \Rightarrow$ 105 to 135

95% Range $\rightarrow$ $\bar{x} \pm 2S = 120 \pm 2(15) \Rightarrow$ 90 to 150

Sep 13-9:26 AM

I randomly Selected $\boxed{40 \text{ exams}}$. Scores had a Symmetric dist. with $\bar{x}=82$ and $S=6$.

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

$\underbrace{\text{Usual Range}}_{95\% \text{ Range}} \Rightarrow \bar{x} \pm 2S = 82 \pm 2(6) \Rightarrow \boxed{70 \text{ to } 94}$

What % of Scores were above 70? $\boxed{97.5\%}$

2.5% | 95% | 2.5%

Low | 70 | Usual Range | 94 | High

How many Scores were unusually low?

$2.5\%(40) = .025(40) = \boxed{1}$

SG 5 ✓ Due Monday

Sep 13-9:33 AM

5 – Number Summary

Min. | Q_1 | Med. | Q_3 | Max.

Q_1: First Quartile

Med.: Median

Q_3: Third Quartile

25% | 25% | 25% | 25%

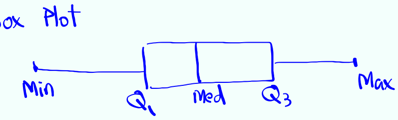
Min | Q_1 | Med. | Q_3 | Max.

25% below Q_1, 75% above Q_1

50% below Median, 50% above Median

75% below Q_3, 25% above Q_3

1) Draw Box Plot



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

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

Lower Fence $= Q_1 - 1.5(IQR)$

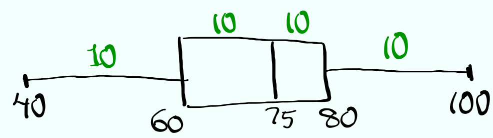
4) below LF OR above UF $\Rightarrow$ Outlier

Sep 13-9:44 AM

I randomly Selected 40 exams. The 5-Number Summary of Scores were $\frac{40}{4} = 10$

40	60	75	80	100
Min	Q_1	Med.	Q_3	Max

Draw Box Plot



$IQR = Q_3 - Q_1 = 80 - 60 = 20$

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

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

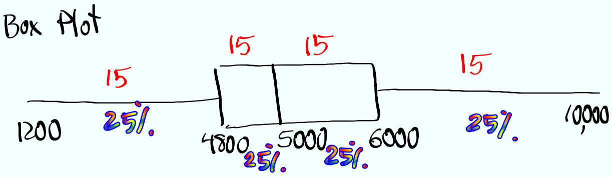
No Score below LF
" " above UF $\Rightarrow$ No outlier

Sep 13-9:51 AM

I randomly Selected 60 nurses. The 5-Number Summary of their Salaries were $\frac{60}{4} = 15$

1200	4800	5000	6000	10,000
Min	Q_1	Med.	Q_3	Max

Box Plot



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

Upper Fence $= Q_3 + 1.5(IQR) = 6000 + 1.5(1200)$
$= \boxed{7800}$

Lower Fence $= Q_1 - 1.5(IQR) = 4800 - 1.5(1200)$
$= \boxed{3000}$

Sep 13-9:58 AM

Percentile

Notation P_K

Data must be Sorted

$K\%$ fall below P_K

$(100-K)\%$ fall above P_K

P_{10}

10% | 90%

P_{10}

P_{80}

80% | 20%

P_{80}

Median P_{50}

50% | 50%

P_{50}

$K\%$ | $(100-K)\%$

P_K

Sep 13-10:20 AM

How to find P_K

1) Location $L = \frac{K}{100} \cdot n$ ← Sample Size

If L is decimal: Round-up

$P_K = L$th element

If L is whole #: $P_K = \frac{L\text{th element} + \text{Next one}}{2}$

when looking for K (Doing Reverse)

$K = \frac{B}{n} \cdot 100$ Round to whole %.

B: # below

n: Sample size

Sep 13-10:23 AM

Consider the Stem Plot below

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

1) $n=25$

2) P_{10}

10% | 90%, 20

$L = \frac{10}{100} \cdot 25 = 2.5 \rightarrow L = 3$

P_{10} = 3rd element

$\boxed{P_{10} = 20}$

3) P_{80}

80% | 20%, 50

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

$P_{80} = \frac{20\text{th} + 21\text{st}}{2} = \frac{48+52}{2} = 50$

4) Find K such that $P_K = 40$

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

#below

$= \frac{15}{25} \cdot 100$

$= 60$

60% | 40%, 40

$P_{60} = 40$

Sep 13-10:27 AM

Consider the Stem Plot below

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

1) $n=27$

2) P_{20}

$L = \frac{20}{100} \cdot 27 = 5.4 \rightarrow L = 6$

P_{20} = 6th element $= \boxed{65}$

3) Median $= P_{50}$

$L = \frac{50}{100} \cdot 27 = 13.5 \quad L = 14$

Median = 14th element $= \boxed{75}$

4) Find K such that $P_K = 90$

Below

$K = \frac{B}{n} \cdot 100 = \frac{22}{27} \cdot 100 = 81.48\ldots \approx 81$

$P_{81} = 90$

SG 6 ✓

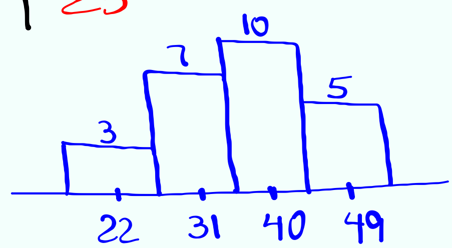
Sep 13-10:35 AM

Class QZ 2

Complete the Chart below

Class limits	Class MP	class F	Cum F.
18 – 26	22	3	3
27 – 35	31	7	10
36 – 44	40	10	20
45 – 53	49	5	25

Draw histogram using class MP & class F.



Sep 13-10:46 AM