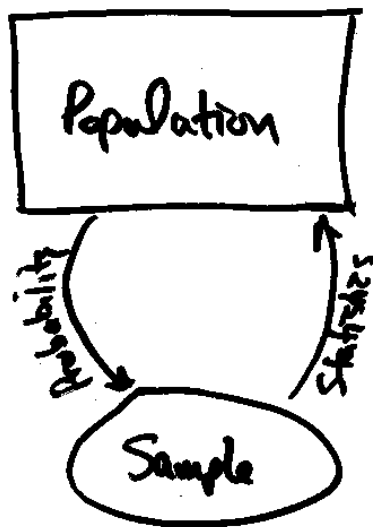


Syllabus & notes:

www.mth.pdx.edu/~fountain

①
Stat 459
1-5

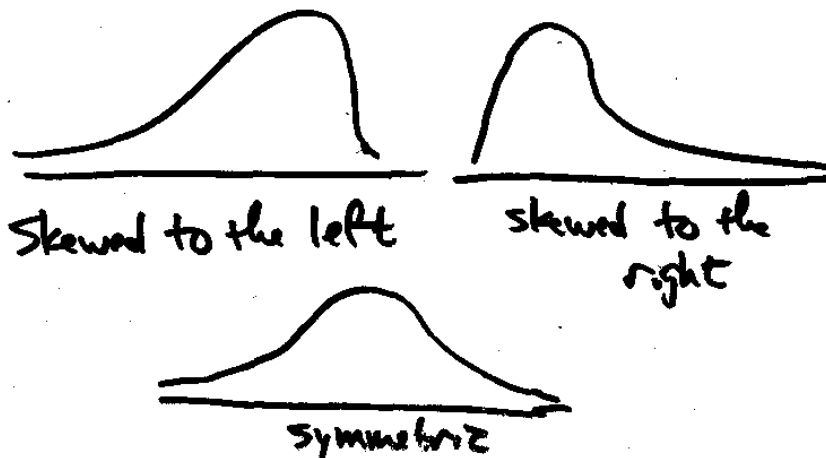


55
75
23
66
65
66
71
79
85

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

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

↑ ↑
Stem & leaf plot



②

Summary table

③

Class	Midpoint	Frequency	Relative Frequency
20-29	24.5	1	.111
30-39	34.5	0	0
40-49	44.5	0	0
50-59	54.5	1	.111
60-69	64.5	3	.333
70-79	74.5	3	.333
80-89	84.5	1	.111
		<u>9</u>	<u>.111</u> ≈ 1

Measures of location

$$\text{Sample mean} = \bar{x} = \frac{\sum x}{n}$$

$$= 65$$

Sample median = Q_2 = "middle value" ④

$$= \begin{cases} \left(\frac{n+1}{2}\right)^{\text{th}} \text{ value if } n \text{ is odd} \\ \text{Average of } \left(\frac{n}{2}\right)^{\text{th}} \text{ and } \left(\frac{n}{2}+1\right)^{\text{th}} \\ \text{values if } n \text{ is even} \end{cases}$$

$$Q_2 = 66$$

Sample mode = most frequent value

midrange = average of highest & lowest values

These last two are rarely used

Trimmed mean :

Cut a certain percentage of values from the ends of a data set and average the remaining values

Example: 10% trimmed mean

10% of 9 ≈ 1

So cut 1 item from each end:

Delete: 23 and 85

Average the other 7 values: 68.14

⑤

Percentiles are markers in the data set, at which a certain percentage of the values lie to the left.

The median is the 50th percentile

The first quartile = Q_1 is the 25th percentile

Q_3 is the 75th percentile

⑥

Measures of variability or dispersion

Range = largest value minus smallest

IQR = interquartile range = $Q_3 - Q_1$

⑦

Sample standard deviation = s

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

Sample variance = s^2

Not used often:

MAD = mean absolute deviation

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

⑧

Other measures:

Skewness (based on $\sum (x - \bar{x})^3$)

Kurtosis (based on $\sum (x - \bar{x})^4$)