

HS167 GUEST LECTURE 2/12/08 TOPIC: CH 4

ASSIGN: READ PP 63-84

EX 4.5, 4.6, 4.16, 4.18, 4.19 SHOW ALL WORK; TURN IN ODDS & EVENS! DO YOUR OWN WORK, AND BE AWARE OF DETAILS

PLS. USE HS167-S08@yahoogroups.com TO ASK QUESTIONS ABOUT HALF THE CLASS HAS FAILED TO SUBSCRIBE - BIG MISTAKE

SUMMARY MEASURES OF:

CENTRAL LOCATION ( $\bar{x}$ , M, mode)

SPREAD ( $S^2/S$ , QUANTILES)

SHAPE (RARELY USED IN PRACTICE; COMPARE  $\bar{x}$  & M INSTEAD)

\* MUCH OF THIS SHOULD BE REVIEW FROM PRE-REQ. COURSE

NOTATION:  $n, X, x_i, \Sigma$

EXAMPLE  $\{X: 21, 42, 5, 11\}$

$n = 4$

$x_1 = 21, x_2 = 42, x_3 = 5, x_4 = 11$

$\Sigma x_i = x_1 + x_2 + \dots + x_n$

MEAN

$\bar{x} = \frac{1}{n} \Sigma x_i = \frac{1}{4} (21 + 42 + 5 + 11) = \frac{1}{4} \cdot 79 = 19.75$

$\bar{x}$  TELLS YOU:

a) "GUESSTIMATE" OF  $\mu$

b) "GUESSTIMATE" OF AN INDIVIDUAL DRAWN @ RANDOM FROM SAMPLE "POP'N"

c) "

"BALANCING POINT CONCEPT (SLIDE 5)

$\bar{x}$  &  $\mu$  ARE DIFFERENT

MEDIAN

ALSO AN AVERAGE

ORDER DATA FIRST: 5 ... 11 ... 21 ... 42

DEPTH (D) OF MEDIAN

$D(M) = \frac{n+1}{2}$

e.g.  $D(M) = \frac{4+1}{2} = 2.5$

MEDIAN  $\rightarrow$  WHEN  $n$  IS EVEN, AVERAGE ADJACENT VALUES, e.g.,  $(11+21)/2 = 16$

ROBUSTNESS OF MEDIAN DEMONSTRATED ON SLIDE 9

MODE - MOST FREQUENT VALUE; USEFUL ONLY IN LARGE DATA SETS

COMPARISON OF  $\bar{x}$ , M, MODE: SLIDE 11

### SPREAD

{ SOLE RELIANCE ON AVERAGES (CENTERS) IS FOOLISH → SLIDE 12  
{ ALSO NOTE SIDE-BY-SIDE STEM PLOTS IN SLIDE 12

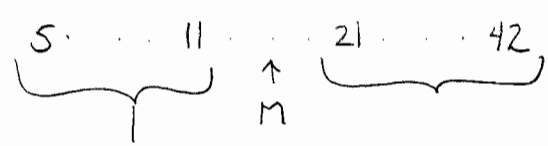
RANGE = MAX - MIN → POOR MEASURE OF SPREAD

QUARTILES: SPLIT DATA IN HALF

IF  $n$  IS ODD → PUT MEDIAN IN BOTH GROUPS\*

Q1 = "MEDIAN" OF LOW GROUP

Q3 = "MEDIAN" OF HIGH GROUP



### TUKEY HINGE METHOD

Q1 BASED ON  $n' = 2$  AND  $D(\frac{n'+1}{2}) = \frac{2+1}{2} = 1.5$

∴ AVERAGE 5 & 11 AND  $Q1 = 8$

Q3 HAS DEPTH OF  $1\frac{1}{2}$  FROM TOP DOWN → AVERAGE 21 & 42

∴  $Q3 = (21 + 42) / 2 = 31.5$

5 POINT SUMMARY = MIN, Q1, M, Q3, MAX

SILLY EXAMPLE: 5, 8, 16, 31.5, 42 DO YOU SEE WHY IT'S SILLY?

IQR =  $Q3 - Q1 = 31.5 - 8 = 23.5$

### REAL BOXPLOT\* - INTRODUCE OUTSIDE VALUE IS LECTURE; COVER PLOTTING IN LAB

- FIRST LOOK FOR OUTSIDE VALUES

OUTSIDE ON TOP IS ABOVE  $FENCE_{UPPER} = Q3 + (1.5)(IQR)$

OUTSIDE ON BOTTOM IS BELOW  $FENCE_{LOWER} = Q1 - (1.5)(IQR)$

EXAMPLE: {5, 11, 21, 42} HAS IQR = 23.5

$F_U = 31.5 + (1.5)(23.5) = 66.75$  (NONE OUTSIDE)

$F_L = 8 - (1.5)(23.5) = -27.25$  (NONE OUTSIDE)

EXAMPLE: {5, 18, 19, 20, 21, 22, 55}

HAS 5 POINT SUMMARY 5, 18.5, 20, 21.5, 55

IQR =  $21.5 - 18.5 = 3$

$F_U = 21.5 + (1.5)(3) = 26$  (55 IS ABOVE)

$F_L = 18.5 - (1.5)(3) = 14$  (5 IS BELOW)

↖ "OUTSIDE VALUES"

\* DIFFERS FROM HS 67 TEXT

### QUICK REVIEW OF STANDARD DEVIATION

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

"ROOT MEAN SQUARE"

{21, 42, 5, 11}

$$\bar{x} = 19.75$$

THIS SHOULD BE REVIEW FROM PRE-REQ. COURSE

	VALUE	DEVIATION	SQ DEVIATION	
	$x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	
1	21	$21 - 19.75 = 1.25$	1.5625	
2	42	$42 - 19.75 = 22.25$	495.0625	
3	5	-14.75	217.5625	
4	11	-8.75	76.5625	
			<u>790.75</u>	$\sum (x_i - \bar{x})^2$

"SUM OF SQUARES"

$$s = \sqrt{\frac{1}{4-1} \cdot 790.75}$$

$$= \sqrt{263.5833}$$

$$\approx 16.2352$$

ALWAY CHECK WITH CALCULATOR