

STA291

Fall 2009



LECTURE 11
THURSDAY, 1 OCTORBER

Homework and Suggested Study Material

2

- Suggested problems from the textbook:
6.9, 6.10, 6.15, 6.16, 6.19, 6.20
- You will get back Exam 1 at Lab on Tue.

Percentiles

3

- The p^{th} percentile is a number such that $p\%$ of the observations take values below it, and $(100-p)\%$ take values above it
- 50th percentile = median
- 25th percentile = lower quartile = Q_1
- 75th percentile = upper quartile = Q_3

In general,

- $L_p = (n+1)p/100^{\text{th}}$ is the spot in the ordered list of observations to find the p^{th} percentile

Quartiles

4

- 25th percentile
 - = lower quartile
 - = *approximately* median of the observations below the median

- 75th percentile
 - = upper quartile
 - = *approximately* median of the observations above the median

Median and Quartiles can be found from a stem and leaf plot

5

- Example: Murder Rate Data (w/o DC—key: 20|3 = 20.3)

Stem	Leaf	#
20	3	1
19		
18		
17		
16		
15		
14		
13	135	3
12	7	1
11	334469	6
10	2234	4
9	08	2
8	03469	5
7	5	1
6	034689	6
5	0238	4
4	46	2
3	0144468999	10
2	039	3
1	67	2

A quarter of the states
has murder rate above...

The median murder rate is...

A quarter of the states
has murder rate below...

-----+-----+-----+-----+

Five-Number Summary

6

- Maximum, Upper Quartile, Median, Lower Quartile, Minimum
- Statistical Software SAS output (Murder Rate Data)

Quantile	Estimate
100% Max	20.30
75% Q3	10.30
50% Median	6.70
25% Q1	3.90
0% Min	1.60

Note the distance from the median to the maximum compared to the median to the minimum.

Interquartile Range

7

- The Interquartile Range (IQR) is the difference between upper and lower quartile
- $IQR = Q_3 - Q_1$
- IQR = Range of values that contains the middle 50% of the data
- IQR increases as variability increases

Box Plot (AKA Box-and-Whiskers Plot)

8

- A box plot is basically a graphical version of the five-number summary (unless there are outliers)
- It consists of a ***box*** that contains the central 50% of the distribution (from lower quartile to upper quartile),
- A ***line*** within the box that marks the median,
- And ***whiskers*** that extend to the maximum and minimum values, unless there are outliers

Outliers

9

- An observation is an outlier if it falls
 - more than 1.5 IQR above the upper quartile or
 - more than 1.5 IQR below the lower quartile
- Example: Murder Rate Data w/o DC
 - upper quartile $Q3 = 10.3$
 - $IQR = 6.4$
 - $Q3 + 1.5 IQR = \underline{\hspace{2cm}}$
 - Any outliers?

Five-Number Summary/Box Plot

10

- On right-skewed distributions, minimum, Q_1 , and median will be “bunched up”, while Q_3 and the maximum will be farther away.
- For left-skewed distributions, the “mirror” is true: the maximum, Q_3 , and the median will be relatively close compared to the corresponding distances to Q_1 and the minimum.
- Guess on symmetric distributions?

Attendance Survey Question 11

11

- ***On a your index card:***
 - Please write down your name and section number
 - Today's Question: