Section 1.6: Measures of Centre

Friday, January 11, 2019 3:38 PM

numerical measures - another way to summarize a data set

(graphs → visual summary)

- particularly of use when comparing two data sets whose values are very similar

numerical measures

→ called parameters for populations

→ called statistics for samples

we'll look at:

measures of centre

measures of dispersion (spread / variability)

---

measures of centre:

mean (average): add all data points, divide by the number of data points

population mean:

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

μ: Greek letter "mu"

where x_i is the i^{th}

measurement

n is the size of the population

sample mean:

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

"x-bar"

one problem: greatly affected by "outliers"

↑

data points far from the majority of points

---

median - if you write your data as an ordered list, it's the middle value

- if you have an even number of data points, it's the average of the two middle points

---

example: starting salary for Bridge graduates

$60 000
$50 000
$58 000
$59 500
$1 500 000

$1 500 000

find the mean and median for this sample.

mean = $\bar{x}$ = $345 500

median = $59 500

by the way, there is no universally recognized symbol for the median. some use $\tilde{x}$ (x-tilde)

note: if there are lots of data points, how do you determine which is the middle one?

$$\text{position} = \tfrac{1}{2}(n+1)$$

for 75 points, the median is the 38th point

76 → 38.5th

take average of 38th and 39th points

mode - most common value in a data set
(rarely used in science/engineering because most data is continuous)

one last thing about the mean:
suppose you have summarized a data set in a

One last thing about the mean:

Suppose you have summarized a data set in a table:

x	f
1	5
2	2
3	4

called a "frequency table"

↑ values of x

↑ number of times they occur

$$\bar{x} = \frac{\sum x_i \cdot f_i}{n} = \frac{1 \cdot 5 + 2 \cdot 2 + 3 \cdot 4}{11}$$

$= \sum f_i$