Objectives

★ To construct samples using the measures of center

Sep 26-7:44 PM

Constructing Samples

Arithmetic Mean

Definition

Suppose there are n observations in a data set consisting of the observations $x_1, x_2, x_3, \ldots, x_n$, then the **arithmetic mean** is defined to be

$$\frac{(x_1 + x_2 + \cdots + x_n)}{n}$$

The formula can also be represented as:

$$\frac{\sum x_i}{n}$$

where

x_i is the i^{th} data value in the data set, and

$\sum$ (pronounced "sigma") is a mathematical notation for adding values.

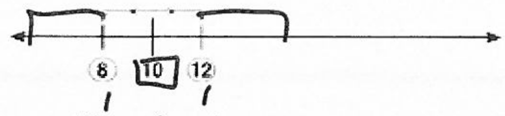
Sep 26-7:44 PM

Example 4.12: Construct a Sample

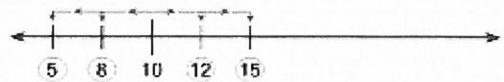
Construct a sample of 4 measurements whose mean is 10.

Solution

Choose two points that are equidistant from the mean.



Again, choose two more points that are equidistant from the mean.



Thus, the sample becomes 5, 8, 12, 15.

Sep 26-7:46 PM

Median

Definition

The **median** of a set of observations is the data value in the middle of an ordered array. The same number of data values is on either side of the median value.

If the number of data values is an even number, then the median is the mean of the middle two numbers.

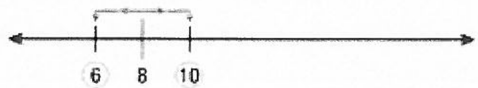
Sep 26-7:48 PM

Example 4.13: Construct a Sample

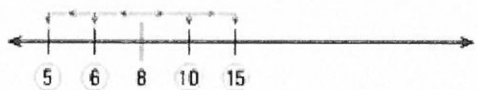
Construct a sample of 4 measurements whose median is 8.

Solution

Choose two points that are equidistant from the median, 8.



Choose two points on either side of the median. It does not matter if the two points are equidistant from the median.



Thus, the sample becomes 5, 6, 10, 15.

Sep 26-7:48 PM

Range and Mode

Definition

The **range** is the difference between the largest and the smallest data values.

Definition

The **mode** of a data set is the most frequently occurring value.

If all of the data values occur only once, or they each occur an equal number of times, the data set is considered to have **no mode**.

If only one value occurs the most, then the data set is said to be **unimodal**.

If exactly two values occur equally often, then the data set is said to be **bimodal**.

If more than two values occur equally often, the data set is **multimodal**.

Sep 26-7:48 PM

Example 4.14: Construct a Bimodal Sample

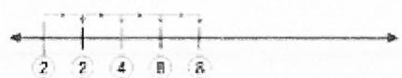
Construct a bimodal sample of 5 measurements whose range is 6.

Solution

First, choose a point.

Then choose a point which is at a distance equal to the range from the first point.

Choose the remaining sample values in between these two values. Since the sample is bimodal, choose the values such that two sample values are tied for the most frequent value.



Thus, the sample becomes **2, 2, 4, 8, 8**.

Sep 26-7:49 PM

Variance

Definition

The **variance** of a data set containing the sample data is given by

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

where

x_i is the i^{th} data value in the data set,
$\overline{x}$ is the sample mean, and
n is the size of the sample,

and is called the **sample variance**.

Sep 26-7:50 PM

Example 4.15: Adding a Sample Value

Given the following sample: $\{4, 6, 7, 9\}$, add one more sample value that will make the mean equal to 15.

Solution

Let x be the added value.

Now the sample becomes 4, 6, 7, 9, x.

The desired mean is 15.

$$\begin{aligned} \frac{(\text{sum of all sample values})}{n} &= 15 \\ \frac{(4+6+7+9+x)}{5} &= 15 \\ (26+x) &= 15 \cdot 5 \\ x &= 75-26 \\ x &= 49 \end{aligned}$$

Thus the new sample value to be added is 49.

Sep 26-7:54 PM

Example 4.16: Adding a Sample Value

Given the following sample: $\{3, 5, 6, 10\}$, add one more sample value that will not change the mean nor the range.

Solution

In order for the range to be constant, we must add a value in between the smallest and the largest values of the sample.

In order for the mean to be constant, we must add a value equal to the mean of the sample.

In this problem, the mean of the sample is $\frac{(3+5+6+10)}{4} = 6$.

Hence the value to be added is 6.

Sep 26-7:55 PM

Example 4.17: Adding a Sample Value

Given the following sample: $\{4, 7, 9, 12, 13\}$, add one more sample value that will not change the mean nor the variance.

Solution

In order for the mean to be constant, we must add a value equal to the mean itself.

Variance is equal to

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

where

x_i is the i^{th} data value in the data set,
$\bar{x}$ is the sample mean, and
n is the size of the sample.

If we add a new sample value equal to the mean, the numerator in the variance will not change but the denominator **increases by one. As a result, the variance will be decreased.**

Hence it is not possible to add a new sample value that will not change either the mean or the variance.

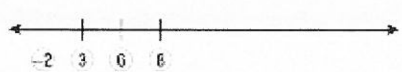
Sep 26-7:55 PM

Example 4.18: Adding a Sample Value

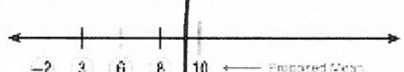
Construct a sample of 6 measurements whose mean is larger than $\frac{2}{3}$ of the measurements.

Solution

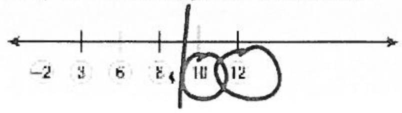
Since the mean is larger than $\frac{2}{3}$ of measurements, the mean must be larger than $\left(\frac{2}{3}\right)6 = 4$ of the measurements. Randomly select any four numbers. For example,



Now select a value for the mean which is larger than any of these numbers.



Now select one number larger than 10, say 12. Thus the sample now looks like this:



Sep 26-7:56 PM

Now, let x be the 6^{th} value.

Now the sample becomes $-2, 3, 6, 8, 12, x$.

The desired mean is 10.

$$\frac{(\text{sum of all sample values})}{n} = 10$$

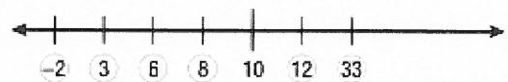
$$\frac{(-2+3+6+8+12+x)}{6} = 10$$

$$(27+x) = 10 \cdot 6$$

$$x = 60-27$$

$$x = 33$$

The 6^{th} value of our sample would be 33.



Thus the sample becomes $-2, 3, 6, 8, 12, 33$.

Sep 26-7:57 PM

Sep 30-9:42 AM