Standard Deviation

Lecture 18
Section 5.3.4

Robb T. Koether

Hampden-Sydney College

Mon, Feb 20, 2012

Outline

1. Variability
2. The Standard Deviation
   - Examples
   - Alternate Formula
3. TI-83 Standard Deviations
4. Interpreting the Standard Deviation
5. Assignment

Outline

1. Variability
2. The Standard Deviation
   - Examples
   - Alternate Formula
3. TI-83 Standard Deviations
4. Interpreting the Standard Deviation
5. Assignment

Variability

- Our ability to estimate a parameter accurately depends on the variability of the estimator.

Variability

- Our ability to estimate a parameter accurately depends on the variability of the estimator.
- The more variable the estimator, the harder it is to estimate the parameter.

Variability

- Our ability to estimate a parameter accurately depends on the variability of the estimator.
- The more variable the estimator, the harder it is to estimate the parameter.
- The variability of the estimator depends on the variability that is inherent in the population.

Variability

- Our ability to estimate a parameter accurately depends on the variability of the estimator.
- The more variable the estimator, the harder it is to estimate the parameter.
- The variability of the estimator depends on the variability that is inherent in the population.
- The more variability in the population, the more variable the estimator.

Variability

- Our ability to estimate a parameter accurately depends on the variability of the estimator.
- The more variable the estimator, the harder it is to estimate the parameter.
- The variability of the estimator depends on the variability that is inherent in the population.
- The more variability in the population, the more variable the estimator.
- What do we mean by “variability” in the population?

Variability

- Our ability to estimate a parameter accurately depends on the variability of the estimator.
- The more variable the estimator, the harder it is to estimate the parameter.
- The variability of the estimator depends on the variability that is inherent in the population.
- The more variability in the population, the more variable the estimator.
- What do we mean by "variability" in the population?
- How do we measure it?

Variability

- Our ability to estimate a parameter accurately depends on the variability of the estimator.
- The more variable the estimator, the harder it is to estimate the parameter.
- The variability of the estimator depends on the variability that is inherent in the population.
- The more variability in the population, the more variable the estimator.
- What do we mean by “variability” in the population?
- How do we measure it?
- The IQR is one way.

Variability

- Our ability to estimate a parameter accurately depends on the variability of the estimator.
- The more variable the estimator, the harder it is to estimate the parameter.
- The variability of the estimator depends on the variability that is inherent in the population.
- The more variability in the population, the more variable the estimator.
- What do we mean by "variability" in the population?
- How do we measure it?
- The IQR is one way.
- We will learn another way.

An Example

- A person offers you \$100 if you can predict the high temperature on March 15, 2012 or on July 15, 2012 to within 5°, your choice of dates.

An Example

- A person offers you $100 if you can predict the high temperature on March 15, 2012 or on July 15, 2012 to within 5°, your choice of dates.
- For which date should you choose to predict the high temperature?

An Example

- A person offers you $100 if you can predict the high temperature on March 15, 2012 or on July 15, 2012 to within 5°, your choice of dates.
- For which date should you choose to predict the high temperature?
- On which date is the high temperature less variable?

An Example

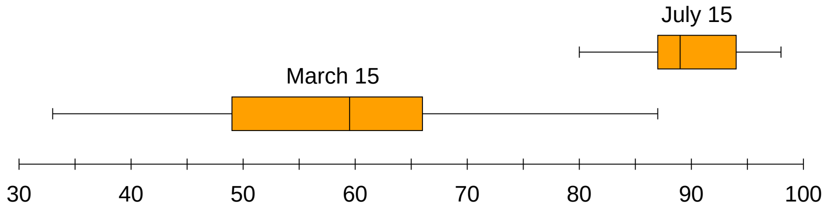
- A person offers you $100 if you can predict the high temperature on March 15, 2012 or on July 15, 2012 to within 5°, your choice of dates.
- For which date should you choose to predict the high temperature?
- On which date is the high temperature less variable?
- Naturally, you should choose the date with less variability.

An Example

- Here are boxplots for the daily highs on March 15 and July 15 for the past 30 years.

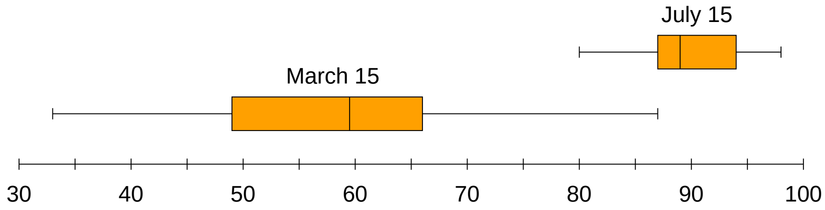
An Example

- Here are boxplots for the daily highs on March 15 and July 15 for the past 30 years.



An Example

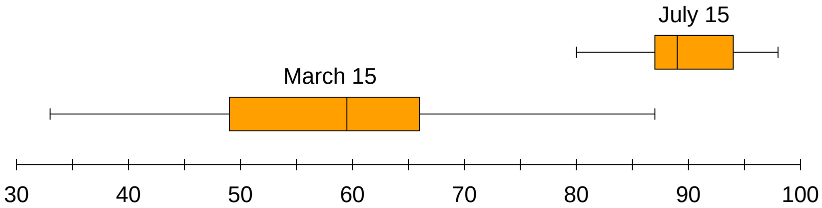
- Here are boxplots for the daily highs on March 15 and July 15 for the past 30 years.



- Which is more variable?

An Example

- Here are boxplots for the daily highs on March 15 and July 15 for the past 30 years.



- Which is more variable?
- How much more variable is it?

Measures of Variability

- The IQR is one measure of variability.
  - IQR for March 15 data $= 66 - 49 = 17$.
  - IQR for July 15 data $= 94 - 87 = 7$.
- By that measure, the July temperatures are about $2\frac{1}{2}$ times as variable as the March temperatures.
- We will also measure variability by using
  - the variance.
  - the standard deviation.

Deviations from the Mean

Definition (Deviation)

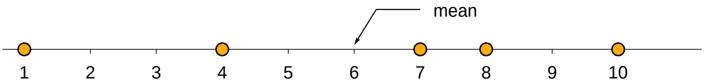
The deviation of an observation x is the difference between x and the sample mean $\overline{x}$.

$$\text{deviation of } x = x - \overline{x}.$$

For a member of the population, the deviation is measured from the population mean:

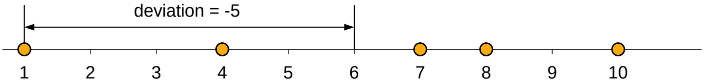
$$\text{deviation of } x = x - \mu.$$

Deviations from the Mean



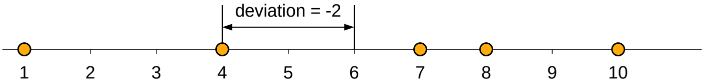
Deviations from the mean

Deviations from the Mean



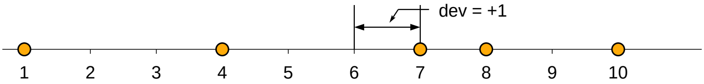
Deviations from the mean

Deviations from the Mean



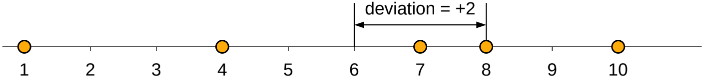
Deviations from the mean

Deviations from the Mean



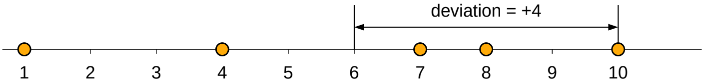
Deviations from the mean

Deviations from the Mean



Deviations from the mean

Deviations from the Mean



Deviations from the mean

Outline

1. Variability
2. The Standard Deviation
   - Examples
   - Alternate Formula
3. TI-83 Standard Deviations
4. Interpreting the Standard Deviation
5. Assignment

Deviations from the Mean

- How do we obtain *one* number that is representative of the whole set of individual deviations?
- Normally we use an average to summarize a set of numbers.
- Why will the average not work in this case?

Deviations from the Mean

- How do we obtain *one* number that is representative of the whole set of individual deviations?
- Normally we use an average to summarize a set of numbers.
- Why will the average not work in this case?
- It will not work because

$$\sum (x - \overline{x}) = 0.$$

Sum of Squared Deviations

- Rather than average the deviations, we will average their squares. That way, there will be no canceling.
- So we compute first the sum of the squared deviations.

Definition (Sum of squared deviations)

The sum of squared deviations, denoted SSX, of a set of numbers is the sum of the squares of their deviations from their mean.

$$\text{SSX} = \sum (x - \overline{x})^2.$$

Sum of Squared Deviations

Example (Calculating SSX)

- Let the sample be $\{1, 4, 7, 8, 10\}$.
- Then

$$\begin{aligned} \text{SSX} &= (1-6)^2 + (4-6)^2 + (7-6)^2 + (8-6)^2 + (10-6)^2 \\ &= (-5)^2 + (-2)^2 + 1^2 + 2^2 + 4^2 \\ &= 25 + 4 + 1 + 4 + 16 \\ &= 50. \end{aligned}$$

Sum of Squared Deviations

Practice

- Let the sample be $\{1, 5, 6, 9, 14\}$.
- Calculate
  - The sample mean.
  - The deviations.
  - The squared deviations.
  - The sum of the squared deviations.

Measuring Population Variability

Definition (Variance of a population)

The variance of a population, denoted σ^2, is the average of the squared deviations of the members of the population.

$$\sigma^2 = \frac{\sum (x - \mu)^2}{N}.$$

Definition (Standard deviation of a population)

The standard deviation of a population, denoted σ, is the square root of the population variance.

$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}.$$

Estimating Population Variability

Definition (Variance of a sample)

The variance of a sample, denoted s^2, is the sum of the squared deviations of the members of the sample, divided by 1 less than the sample size.

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

Definition (Standard deviation of a sample)

The standard deviation of a sample, denoted s, is the square root of the sample variance.

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

The Sample Variance

- Theory shows that if we divided $\sum(x - \overline{x})^2$ by n, then s^2 would systematically underestimate σ^2.
- Theory also shows that if we divided $\sum(x - \overline{x})^2$ by $n - 1$ instead of n, then s^2 would not systematically underestimate nor overestimate σ^2.
- Therefore, we divide by $n - 1$.

Outline

1. Variability
2. The Standard Deviation
   - Examples
   - Alternate Formula
3. TI-83 Standard Deviations
4. Interpreting the Standard Deviation
5. Assignment

Example

Example (Calculating s)

- For the sample $\{1, 4, 7, 8, 10\}$, we found that

$$\text{SSX} = 50.$$

- Therefore,

$$s^2 = \frac{50}{4} = 12.5$$

and so

$$s = \sqrt{12.5} = 3.54.$$

Sum of Squared Deviations

Practice

- Let the sample be $\{1, 5, 6, 9, 14\}$.
- Calculate s^2 and s.

Example

- How does s compare to the individual deviations?
- We will interpret s as being "representative" of the deviations in the sample.
- Does that seem reasonable for the previous examples?

Example

- It turns out that the standard deviation of the March 15 high temperatures is 12.7.
- The standard deviation of the July 15 high temperatures is 4.5.

Outline

1. Variability
2. The Standard Deviation
   - Examples
   - Alternate Formula
3. TI-83 Standard Deviations
4. Interpreting the Standard Deviation
5. Assignment

Alternate Formula for SSX

- An alternate formula for SSX is

$$\text{SSX} = \sum x^2 - \frac{\left(\sum x\right)^2}{n}.$$

- Then, as before

$$s^2 = \frac{\text{SSX}}{n-1}$$

and

$$s = \sqrt{\frac{\text{SSX}}{n-1}}.$$

- Often this formula is easier to use.

Example

Example (Alternate formula for SSX)

- Let the sample be $\{-8, -4, 1, 3, 10\}$.
- Then
$$\sum x = 2$$
and
$$\sum x^2 = 64 + 16 + 1 + 9 + 100 = 190.$$
- So
$$\begin{aligned} \text{SSX} &= 190 - \frac{2^2}{5} \\ &= 189.2 \end{aligned}$$
- Then $s = \sqrt{\frac{189.2}{4}} = \sqrt{47.3} = 6.877$.

Sum of Squared Deviations

Practice

- Let the sample be $\{1, 5, 6, 7, 10\}$.
- Find $\sum x$.
- Find $\sum x^2$.
- Use the alternate formula to find SSX, s^2, and s.

Outline

1. Variability
2. The Standard Deviation
   - Examples
   - Alternate Formula
3. TI-83 Standard Deviations
4. Interpreting the Standard Deviation
5. Assignment

TI-83 - Standard Deviations

TI-83 Standard Deviations

- Follow the procedure for computing the mean.
- The display shows $\mathtt{Sx}$ and $\sigma\mathtt{x}$.
- $\mathtt{Sx}$ is the sample standard deviation.
- $\sigma\mathtt{x}$ is the population standard deviation.

Example

Example

TI-83 Standard Deviations

- Let the sample be $\{-8, -4, 1, 3, 10\}$.
- We get
  - $\mathrm{Sx} = 6.877$.
  - $\sigma\mathrm{x} = 6.151$.

Example

Example

TI-83 Standard Deviations

- Let the sample be $\{-8, -4, 1, 3, 10\}$.
- We get
  - $\mathrm{Sx} = 6.877$.
  - $\sigma\mathrm{x} = 6.151$.
- Which one is *s*?

Example

Example

TI-83 Standard Deviations

- Let the sample be $\{-8, -4, 1, 3, 10\}$.
- We get
  - $\mathtt{Sx} = 6.877$.
  - $\sigma\mathtt{x} = 6.151$.
- Which one is s?
- Then what is $\sigma\mathtt{x}$?

Sum of Squared Deviations

Practice

- Let the sample be $\{1, 5, 6, 7, 10\}$.
- Use the TI-83 to find s and s^2.
- What are the values of $\sum x$ and $\sum x^2$?

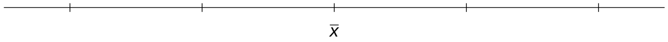
Outline

1. Variability
2. The Standard Deviation
   - Examples
   - Alternate Formula
3. TI-83 Standard Deviations
4. Interpreting the Standard Deviation
5. Assignment

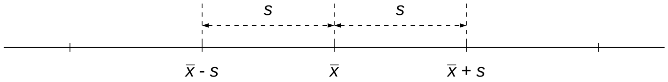
Interpreting the Standard Deviation

- Observations that deviate from $\overline{x}$ by much more than s are unusually far from the mean.
- Observations that deviate from $\overline{x}$ by much less than s are unusually close to the mean.

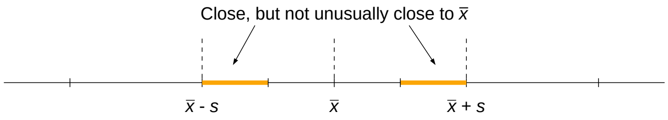
Interpreting the Standard Deviation



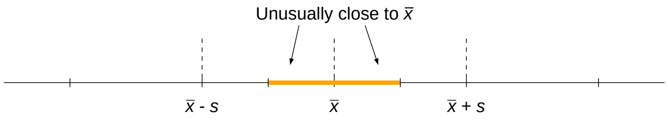
Interpreting the Standard Deviation



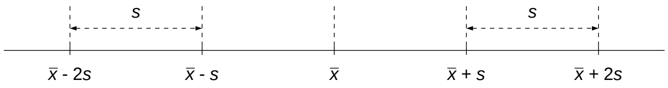
Interpreting the Standard Deviation



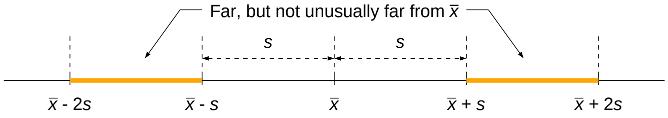
Interpreting the Standard Deviation



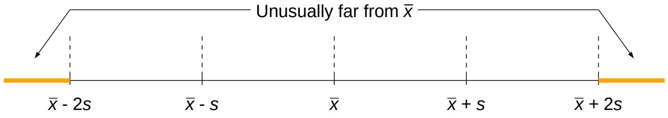
Interpreting the Standard Deviation



Interpreting the Standard Deviation



Interpreting the Standard Deviation



Outline

1. Variability
2. The Standard Deviation
   - Examples
   - Alternate Formula
3. TI-83 Standard Deviations
4. Interpreting the Standard Deviation
5. Assignment

Assignment

Homework

- Read Section 5.3.4, pages 326 - 333.
- Let's Do It! 5.13, 5.14, 5.15.
- Page 333, exercises 10, 11, 14, 16 - 18, 20, 21.
- Chapter 5 review, p. 345, exercises 29 - 32, 36 - 40, 42 - 44, 47, 52, 53, 55.