

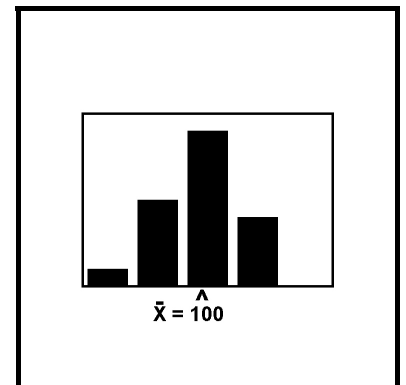
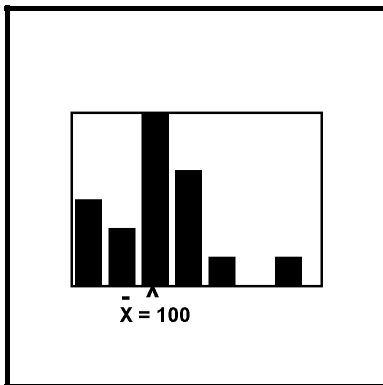
Descriptive Statistics

Lesson 3

Measuring Variation

3.1- What is Data Variation?

Data variation is a numeric value which measures the spread of data from the mean \bar{x} . For example, the two sets of numbers 20, 22, 25, 30 and 5, 22, 25, 45 both have the same mean $\bar{x} = 24.25$, but the spreads from 24.25 are different since the first group of data is not as varied as the second group of data. The following two histograms graphically demonstrate two sets of data both having a mean $\bar{x} = 100$ but different variations.



The following are three common methods for representing the variation of data.

The Range

The range of data is the difference between the largest and smallest numbers in the data.

Example: Assume the data is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Since 10 and 1 are the largest and smallest numbers respectively, the range is $10 - 1 = 9$.

The Absolute Mean Variation (AMV)

The following table shows how the absolute mean variation is computed:

Data x	$ x - \bar{x} $
1	$ 1 - 5.5 = 4.5$
2	$ 2 - 5.5 = 3.5$
3	$ 3 - 5.5 = 2.5$
4	$ 4 - 5.5 = 1.5$
5	$ 5 - 5.5 = 0.5$
6	$ 6 - 5.5 = 0.5$
7	$ 7 - 5.5 = 1.5$
8	$ 8 - 5.5 = 2.5$
9	$ 9 - 5.5 = 3.5$
10	$ 10 - 5.5 = 4.5$
$\bar{x} = 5.5$ AMV = $25/10 = 2.5$	

The Standard Deviation (s)

The following table shows how the standard deviation is computed:

(1) Data x	(2) $x - \bar{x}$	(3) $(x - \bar{x})^2$
1	-4.5	$(-4.5)^2 = 20.25$
2	-3.5	$(-3.5)^2 = 12.25$
3	-2.5	$(-2.5)^2 = 6.25$
4	-1.5	$(-1.5)^2 = 2.25$
5	-0.5	$(-0.5)^2 = 0.25$
6	0.5	$(0.5)^2 = 0.25$
7	1.5	$(1.5)^2 = 2.25$
8	2.5	$(2.5)^2 = 6.25$
9	3.5	$(3.5)^2 = 12.25$
10	4.5	$(4.5)^2 = 20.25$
$\bar{x} = 5.5$		$s = \sqrt{\frac{82.5}{10}} = \sqrt{8.25} \approx 2.87$

The following are the rules for computing the standard deviation:

Rule 1: Compute \bar{x} from the data in column 1.

Rule 2: The numbers in column 2 are computed using the formula $x - \bar{x}$.

Rule 3: The numbers in column 3 are computed using the formula $(x - \bar{x})^2$.

Rule 4: To compute the standard deviation:

- sum the numbers in column 3. In the table, this sum is 82.5.
- Divide this sum by the total number of values in column 3. This gives 8.25.
- The standard deviation s is the square root of the value computed in b. This gives $s = 2.87$.

Variance

The variance of a set of data is defined as the square of the standard deviation s^2 .

For the above example, the variance is $s^2 = 2.87^2 = 8.25$.

Of these methods used to compute the variation of a set of data, the variance and standard deviation are the most frequently used.

3.1 - Example 1: Ms. Jones teaches a Latin class at a Senior center. The following data is the ages of her students: 74, 67, 65, 74, 67, 81, 65, 85, 67, 80.

Find the

- range.
- absolute mean variation.
- standard deviation.
- variance.

Solutions:

►(a).

The oldest and youngest ages are 85 and 65 respectively. Therefore, the range is $85 - 65 = 20$ years old.

►(b).

The following table shows how the absolute mean variation is computed.

Data x	$ x - \bar{x} $
74	$ 74 - 72.5 = 1.5$
67	$ 67 - 72.5 = 5.5$
65	$ 65 - 72.5 = 7.5$
74	$ 74 - 72.5 = 1.5$
67	$ 67 - 72.5 = 5.5$
81	$ 81 - 72.5 = 8.5$
65	$ 65 - 72.5 = 7.5$
85	$ 85 - 72.5 = 12.5$
67	$ 67 - 72.5 = 5.5$
80	$ 80 - 72.5 = 7.5$
$\bar{x} = 725/10 = 72.5$	AMV = $63/10 = 6.3$

►(c).

The following table shows how the standard deviation is computed:

(1) Data x	(2) $x - \bar{x}$.	(3) $(x - \bar{x})^2$
74	$74 - 72.5 = 1.5$	$(1.5)^2 = 2.25$
67	$67 - 72.5 = -5.5$	$(-5.5)^2 = 30.25$
65	$65 - 72.5 = -7.5$	$(-7.5)^2 = 56.25$
74	$74 - 72.5 = 1.5$	$(1.5)^2 = 2.25$
67	$67 - 72.5 = -5.5$	$(-5.5)^2 = 30.25$
81	$81 - 72.5 = 8.5$	$(8.5)^2 = 72.25$
65	$65 - 72.5 = -7.5$	$(-7.5)^2 = 56.25$
85	$85 - 72.5 = 12.5$	$(12.5)^2 = 156.25$
67	$67 - 72.5 = -5.5$	$(-5.5)^2 = 30.25$
80	$80 - 72.5 = 7.5$	$(7.5)^2 = 56.25$
$\bar{x} = 72.5$		$s = \sqrt{\frac{492.5}{10}} = \sqrt{49.25} \approx 7.02$

►(d).

The variance is $s^2 = 7.02^2 \approx 49.25$.

Solved Problems

3.1 - Solved Problem 1: Rick is a member of the All Star Bowling Team. Last week he bowled the following scores: 187, 167, 201, 185, 167, 210, 205, 167.

Find the

- (a). range.
- (b). absolute mean variation.
- (c). standard deviation.
- (d). variance.

Solutions:

►(a).

The highest and lowest scores are 210 and 167 respectively. Therefore the range is $210 - 167 = 43$ points.

►(b).

The following table shows how the absolute mean variation is computed.

Data x	$ x - \bar{x} $
187	$ 187 - 186.125 = 0.875$
167	$ 167 - 186.125 = 19.125$
201	$ 201 - 186.125 = 14.875$
185	$ 185 - 186.125 = 1.125$
167	$ 167 - 186.125 = 19.125$
210	$ 210 - 186.125 = 23.875$
205	$ 205 - 186.125 = 18.875$
167	$ 167 - 186.125 = 19.125$
$\bar{x} = 1489/8 = 186.125$	AMV = $117/8 = 14.625$

►(c).

The following table shows how the standard deviation is computed:

(1) Data x	(2) $x - \bar{x}$	(3) $(x - \bar{x})^2$
187	$187 - 186.125 = 0.875$	0.77
167	$167 - 186.125 = -19.125$	365.77
201	$201 - 186.125 = 14.875$	221.27
185	$185 - 186.125 = -1.125$	1.27
167	$167 - 186.125 = -19.125$	365.77
210	$210 - 186.125 = 23.875$	570.02
205	$205 - 186.125 = 18.875$	356.27
167	$167 - 186.125 = -19.125$	365.77
$\bar{x} = 186.13$		$s = \sqrt{\frac{2246.91}{8}} \approx 16.76$

►(d).

The variance is $s^2 = 16.76^2 \approx 280.90$.**Unsolved Problems with Answers****3.1 - Problem 1:** A die is tossed 20 times with the following outcomes:

4, 5, 6, 4, 2, 4, 2, 2, 4, 4, 2, 4, 1, 1, 3, 4, 5, 3, 3, 4.

Find the

- (a). range.
- (b). absolute mean variation.
- (c). standard deviation.
- (d). variance.

Answers:

- (a). 5
- (b). 1.12
- (c). 1.31
- (d). 1.73



Refer back to 3.1 - Example 1 & 3.1 - Solved Problem 1.

3.2 - Computing the Variance and Standard Deviation for Frequency Distributions.

The following example of a frequency distribution demonstrates how to compute its standard deviation:

Class	(1) Mid-Value	(2) Frequency	(3) (1)x(2)	(4) [(1) - \bar{x}] ²	(5) (2)x(4)
[25,35)	30	5	150	(30 - 57.5) ² = 757.25	3781.25
[35,45)	40	7	280	(40 - 57.5) ² = 306.25	2143.75
[45,55)	50	10	500	(50 - 57.5) ² = 56.25	562.50
[55,65)	60	6	360	(60 - 57.5) ² = 6.25	37.50
[65,75)	70	4	280	(70 - 57.5) ² = 156.25	625.50
[75,85)	80	12	960	(80 - 57.5) ² = 506.25	6075.50
Total		44	2530		13225

$$\bar{x} = 2530/44 = 57.50$$

$$s^2 = 13225/44 = 300.57$$

$$s = \sqrt{300.57} \approx 17.34$$

3.2 - Example 1: A survey of hourly wages of fifty employees at a local fast food restaurant resulted in the following frequency distribution:

Hourly wage Classes	Number of Employees
[\$3.50,\$4.50)	4
[\$4.50,\$5.50)	14
[\$5.50,\$6.50)	16
[\$6.50,\$7.50)	12
[\$7.50,\$8.50)	3
[\$8.50,\$9.50)	1
Total	50

Compute the variance and standard deviation.

Solution:

The following table computes the standard deviation and variance:

Classes	(1) Mid-values	(2) frequency	(3) (1)x(2)	(4) $[(1) - \bar{x}]^2$	(5) (2)x(4)
[\$3.50,\$4.50)	\$4.00	4	\$16.00	$(4 - 5.98)^2 = 3.92$	15.68
[\$4.50,\$5.50)	\$5.00	14	\$70.00	$(5 - 5.98)^2 = 0.96$	13.44
[\$5.50,\$6.50)	\$6.00	16	\$96.00	$(6 - 5.98)^2 = 0.0004$	0.0064
[\$6.50,\$7.50)	\$7.00	12	\$84.00	$(7 - 5.98)^2 = 1.04$	12.48
[\$7.50,\$8.50)	\$8.00	3	\$24.00	$(8 - 5.98)^2 = 4.08$	12.24
[\$8.50,\$9.50)	\$9.00	1	\$ 9.00	$(9 - 5.98)^2 = 9.12$	9.12
Total		50	\$299.00		62.97

$$\bar{x} = \$299/50 = \$5.98$$

$$s^2 = 62.97/50 \approx \$1.26$$

$$s = \sqrt{1.26} \approx \$1.12$$

Solved Problems

3.2 - Solved Problem 1: The Frozen Foods Company recently developed a new pizza. To check the retail prices the supermarkets are charging for this pizza, a survey of the prices that are charged by 20 supermarkets is taken. The following frequency distribution was collected:

Price per Pizza	Number of Supermarkets
[\$4.50,\$4.70)	1
[\$4.70,\$4.90)	4
[\$4.90,\$5.10)	2
[\$5.10,\$5.30)	0
[\$5.30,\$5.50)	3
[\$5.50,\$5.70)	1
[\$5.70,\$5.90)	3
[\$5.90,\$6.10)	0
[\$6.10,\$6.30)	4
[\$6.30,\$6.50)	2
Total	20

Compute the variance and standard deviation.

Solution:

The following table computes the standard deviation and variance:

Classes	(1) Mid-values	(2) frequency	(3) (1)x(2)	(4) [(1)- \bar{x}] ²	(5) (2)x(4)
[\$4.50,\$4.70)	\$4.60	1	4.60	0.86	0.86
[\$4.70,\$4.90)	\$4.80	4	19.20	0.53	2.12
[\$4.90,\$5.10)	\$5.00	2	10.00	0.28	0.56
[\$5.10,\$5.30)	\$5.20	0	0.00	0.11	0.00
[\$5.30,\$5.50)	\$5.40	3	16.20	0.02	0.06
[\$5.50,\$5.70)	\$5.60	1	5.60	0.01	0.01
[\$5.70,\$5.90)	\$5.80	3	17.40	0.07	0.21
[\$5.90,\$6.10)	\$6.00	0	0.00	0.22	0.00
[\$6.10,\$6.30)	\$6.20	4	24.80	0.45	1.80
[\$6.30,\$6.50)	\$6.40	2	12.80	0.76	1.52
Total		20	\$110.60		\$7.14

$$\bar{x} = \$110.60/20 = \$5.53$$

$$s^2 = 7.14/20 \approx 0.36$$

$$s = \sqrt{0.36} \approx \$0.60$$

Unsolved Problems with Answers

3.2 - Problem 1: Ms. Clark is the manager of a weight reduction club of 294 women. For the members, she recorded their individual weights at the time they joined the club. The following frequency distribution represents this data:

Weight classes	Number of Members
[130,140)	15
[140,150)	45
[150,160)	55
[160,170)	42
[170,180)	32
[180,190)	48
[190,200)	25
[200,210)	12
[210,220)	9
[220,230)	9
[230,240)	2
Total	294

Find the variance and standard deviation.

Answers:

$$s^2 = 528.92$$

$$s = 23.00$$

↑↑ Refer back to 3.2 - Example 1 & 3.2 - Solved Problem 1.

3.3 - An Application for the Standard Deviation

In Statistics, we frequently are interested in the data that fall within a given number of standard deviations from the mean \bar{x} .

3.3 - Example 1: A sample of 30 families in New York City was recently taken. The following data, listed in ascending order, represents the number of children per family:

0, 1, 2, 2, 2, 2, 2, 2, 2, 2,
2, 2, 3, 3, 3, 3, 3, 4, 4, 4,
4, 4, 5, 5, 5, 5, 5, 6, 6, 7

Find

- the mean \bar{x} .
- the standard deviation.
- the numbers that are within one standard deviation of \bar{x} .
- the numbers that are within two standard deviations of \bar{x} .
- the numbers that are within three standard deviations of \bar{x} .
- the percent of numbers that are within two standard deviations of \bar{x} .

Solutions:

►(a).

To find the mean, add all the above numbers and divide by 30. This gives $\bar{x} \approx 3.33$.

►(b).

Following the rules in section 1 of this lesson, we find the standard deviation $s = 1.65$.

►(c).

To find the numbers that are within one standard deviation of \bar{x} , we select those numbers that are between

$3.33 - 1.65 = 1.68$ and $3.33 + 1.65 = 4.98$. This would include all numbers between 2 and 4 children:
 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
 3, 3, 3, 3, 3, 4, 4, 4, 4, 4.

►(d).

To find the numbers that are within two standard deviations of \bar{x} , we select those numbers that are between $3.33 - 2(1.65) = 0.03$ and $3.33 + 2(1.65) = 6.63$. This would include all numbers between 1 and 6 children:

1, 2, 2, 2, 2, 2, 2, 2, 2, 2,
 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5,
 5, 5, 5, 5, 6, 6.

►(e).

To find the numbers that are within three standard deviations of \bar{x} , we select those numbers that are between $3.33 - 3(1.65) = -1.62$ and $3.33 + 3(1.65) = 8.28$. This would include all 30 numbers.

►(f).

Since 28 out of 30 numbers are within 2 standard deviations of \bar{x} , the percent is $28/30 \approx 93\%$.

Solved Problems

3.3 - Solved Problem 1: A computer generates the following 25 random numbers ranging from 1 to 10:

1, 2, 2, 3, 3, 4, 5, 5, 5, 6,
 6, 6, 6, 6, 6, 6, 7, 7, 7, 8,
 9, 9, 9, 10, 10.

Find

- the mean \bar{x} .
- the standard deviation.
- the numbers that are within one standard deviation of \bar{x} .
- the numbers that are within two standard deviations of \bar{x} .
- the numbers that are within three standard deviations of \bar{x} .
- the percent of numbers that are within two standard deviations of \bar{x} .

Solutions:

►(a).

To find the mean, add all the above numbers and divide by 25. This gives $\bar{x} \approx 5.92$

►(b).

Following the rules in section 1 of this lesson, we find the standard deviation $s \approx 2.48$.

►(c).

To find the numbers that are within one standard deviation of \bar{x} , we select those numbers that are between $5.92 - 2.48 = 3.44$ and $5.92 + 2.48 = 8.40$. This would include all numbers between 4 and 8:

4, 5, 5, 5, 6,
6, 6, 6, 6, 6,
6, 7, 7, 7, 8.

►(d).

To find the numbers that are within two standard deviations of \bar{x} , we select those numbers that are between $5.92 - 2(2.48) = 0.96$ and $5.92 + 2(2.48) = 10.88$. This would include all the numbers.

►(e).

To find the numbers that are within three standard deviations of \bar{x} , we select those numbers that are between $5.92 - 3(2.48) = -1.52$ and $5.92 + 3(2.48) = 13.36$. This would include all numbers 25 numbers.

►(f).

Since all of the 25 numbers are within 2 standard deviations of \bar{x} , the percent of these numbers is 100%.

Unsolved Problems with Answers

3.3 - Problem 1: Ms. Jones recently gave a final examination in Spanish. The thirty students in her class received the follow grades:

57.47, 60.05, 60.83, 61.78, 62.70,
62.73, 62.80, 63.16, 63.24, 63.27,
63.31, 63.94, 64.18, 64.68, 64.83,
64.97, 65.18, 65.26, 65.31, 65.51,
65.60, 65.88, 66.31, 66.46, 66.56,
67.21, 67.57, 67.64, 68.44, 69.61.

Find the

(a). mean \bar{x} .

(b). standard deviation.

(c). numbers that are within one standard deviation of \bar{x} .

(d). numbers that are within two standard deviations of \bar{x} .

(e). numbers that are within three standard deviations of \bar{x} .

(f). the percent of numbers that are within two standard deviations of \bar{x} .

Answers:

►(a). $\bar{x} = 64.55$

►(b). $s = 2.57$

►(c).

62.70, 62.73, 62.80, 63.16, 63.24,
63.27, 63.31, 63.94, 64.18, 64.68,
64.83, 64.97, 65.18, 65.26, 65.31,
65.51, 65.60, 65.88, 66.31, 66.46,
66.56

►(d).

60.05, 60.83, 61.78, 62.70, 62.73,
62.80, 63.16, 63.24, 63.27, 63.31,
63.94, 64.18, 64.68, 64.83, 64.97,
65.18, 65.26, 65.31, 65.51, 65.60,
65.88, 66.31, 66.46, 66.56, 67.21,
67.57, 67.64, 68.44, 69.61

►(e). All the numbers.

►(f). 96.67%



Refer back to 3.3 - Example 1 & 3.3 - Solved Problem 1.

Supplementary Problems

1. For the set of numbers 0,1,2,3,4,5,6,...,100, use the formulas

$$1 + 2 + 3 + 4 + \dots + n = \frac{n(n + 1)}{2} ,$$

$$1^2 + 2^2 + 3^2 + 4^2 + \dots + n^2 = \frac{n(n + 1)(2n + 1)}{6} .$$

a. to find the mean \bar{x} .

b. to find its standard deviation.

c. List all numbers that fall within one standard deviation of the mean.

d. List all numbers that fall within two standard deviations of the mean.

2. A statistician was hired by a professional basketball team to do a study on the performance of the team. The

following is a list, in numeric order, of the final scores of the team's last 50 games:

85 85 85 86 87 89 90 91 91 92 92 92 93
93 94 97 97 99 99 100 100 103 103 103 104 106
106 106 107 108 108 108 109 109 109 112 113 114 115
116 117 117 118 119 119 120 121 122 123 125.

- a. Find the standard deviation.
- b. For this data complete the following frequency distribution table:

Classes	Frequency
[80,90)	
[90,100)	
[100,110)	
[110,120)	
[120,130)	

- c. For the above distribution, compute the standard deviation.
-