Statistics Solutions Manual

Problem 1: NHL Salary Mean vs. Median

According to Bloomberg News, in 2011 the average annual salary for a player in the National Hockey League (NHL) can be described by the numbers $760,000 and $2,000,000. Which one of these numbers is the mean and which is the median? Explain your reasoning.

- ○ The mean is $760,000 and the median is $2,000,000 since the mean is always smaller than the median.
- ○ The mean is $2,000,000 and the median is $760,000 since the mean is always larger than the median.
- ○ The mean is $760,000 and the median is $2,000,000 since there are a small number of hockey players with exceptionally small salaries which decrease the mean but have little effect on the median.
- ◉ The mean is $2,000,000 and the median is $760,000 since there are a small number of hockey players with exceptionally large salaries which increase the mean but have little effect on the median.

✓

Analysis

This problem deals with **skewed distributions**. In financial data (like salaries), distributions are often "skewed right," meaning a few individuals earn exponentially more than the majority.

- The **Mean** is sensitive to outliers (extremely high values pull the average up).
- The **Median** is resistant to outliers (it represents the physical middle of the dataset).

Since $2,000,000 is significantly higher than $760,000, the higher number represents the arithmetic average (mean) pulled up by superstars, while the lower number represents the typical player (median).

Solution

Correct Choice: The mean is $2,000,000 and the median is $760,000 since there are a small number of hockey players with exceptionally large salaries which increase the mean but have little effect on the median.

Problem 2: Frequency Table Statistics

The table below shows the scores of a group of students on a 10-point quiz.

Test Score	Frequency
3	1
4	2
5	1
6	2
7	3
8	3
9	3
10	1

The mean score on this test is: 6.88

The median score on this test is: 7

1. Calculate the Mean ($\bar{x}$)

Formula: $\bar{x} = \frac{\sum(x \cdot f)}{\sum f}$

Score (x)	Frequency (f)	$x \cdot f$
3	1	3
4	2	8
5	1	5
6	2	12
7	3	21
8	3	24
9	3	27
10	1	10
Total	**16**	**110**

$$\text{Mean} = \frac{110}{16} = 6.875 \approx \mathbf{6.88}$$

2. Calculate the Median

The total frequency (n) is 16. The median is the average of the 8th and 9th values in the ordered set.

- Scores 3–6 account for $1 + 2 + 1 + 2 = 6$ students.
- The next score is 7, which has 3 students (occupying positions 7, 8, and 9).

Both the 8th and 9th values are 7.

$$\text{Median} = \mathbf{7}$$

Problem 3: Median of a List

Find the median for this list of numbers

67	92
40	75
85	81
87	29
79	7
82	96
13	76
35	

Median = 76

Data List: {67, 92, 40, 75, 85, 81, 87, 29, 79, 7, 82, 96, 13, 76, 35}

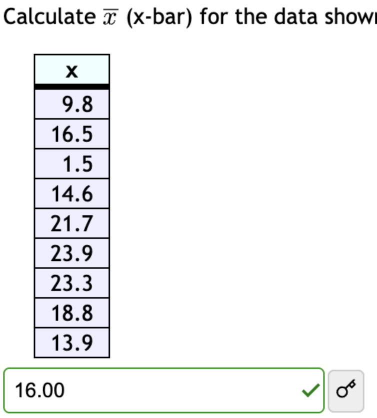
Solution

First, sort the data in ascending order ($n = 15$):

$$7, 13, 29, 35, 40, 67, 75, \mathbf{76}, 79, 81, 82, 85, 87, 92, 96$$

The median is the middle position: $\frac{n+1}{2} = \frac{16}{2} = 8$th value.
Median = 76

Problem 4: Calculate Mean ($\bar{x}$)



Calculate $\overline{x}$ (x-bar) for the data shown, to two decimal places

x
9.8
16.5
1.5
14.6
21.7
23.9
23.3
18.8
13.9

16.00

Data: $\{9.8, 16.5, 1.5, 14.6, 21.7, 23.9, 23.3, 18.8, 13.9\}$ $(n = 9)$

Solution

$$\bar{x} = \frac{\sum x}{n} = \frac{144}{9} = \mathbf{16.00}$$

Problem 5: Mean of a List

Find the mean for this list of numbers

17	39
63	56
43	1
84	8
13	91
49	57
67	92
100	

Mean = 52

Data: $\{17, 39, 63, 56, 43, 1, 84, 8, 13, 91, 49, 57, 67, 92, 100\}$ $(n = 15)$

Solution

$$\bar{x} = \frac{\sum x}{n} = \frac{780}{15} = \mathbf{52}$$

Problem 6: Mode of a List

Find the mode for this list of numbers

45	9
84	32
18	66
50	68
87	93
62	73
14	50
86	

Mode = 50

Data: {45, 9, 84, 32, 18, 66, **50**, 68, 87, 93, 62, 73, 14, **50**, 86}

Solution

The mode is the most frequently occurring number. The number **50** appears twice; all others appear once.

Mode = 50

Problem 7: Normal Distribution (Acme Widgets)

The Acme Company manufactures widgets. The distribution of widget weights is bell-shaped. The widget weights have a mean of 47 ounces and a standard deviation of 6 ounces.

Use the Empirical Rule.

Suggestion: sketch the distribution in order to answer these questions.

a) 99.7% of the widget weights lie between 29 and 65

b) What percentage of the widget weights lie between 41 and 65 ounces? 83.85 %

c) What percentage of the widget weights lie below 59 ? 97.5 %

Given: Mean (μ) = 47, Standard Deviation (σ) = 6.

a) 99.7% of weights lie between...

According to the Empirical Rule, 99.7% of data lies within 3 standard deviations ($\pm 3\sigma$).

$$\text{Lower Bound} = \mu - 3\sigma = 47 - 18 = \mathbf{29}$$
$$\text{Upper Bound} = \mu + 3\sigma = 47 + 18 = \mathbf{65}$$

b) Percentage between 41 and 65?

- 41 is $47 - 6$ (-1σ). Area from Mean to -1σ is 34%.
- 65 is $47 + 18$ ($+3\sigma$). Area from Mean to $+3\sigma$ is 49.85% (half of 99.7%).

$$\text{Total Area} = 34 + 49.85 = \mathbf{83.85}\%$$

c) Percentage below 59?

- 59 is $47 + 12$ ($+2\sigma$).
- Area below Mean is 50%.
- Area from Mean to $+2\sigma$ is 47.5% (half of 95%).

$$\text{Total Area} = 50 + 47.5 = \mathbf{97.5}\%$$

Problem 8: Variation (Gap vs. Old Navy)

Suppose that we are studying the amount of time customers wait in line at the checkout at the Gap and Old Navy. The average wait time at both stores is five minutes. At the Gap, the standard deviation for the wait time is 2 minutes; at Old Navy the standard deviation for the wait time is 5 minutes.

Because Old Navy has a higher standard deviation, we know that there is more variation in the wait times at Old Navy. Overall, wait times at Old Navy are more spread out from the average; wait times at the Gap are more concentrated near the average.

Solution

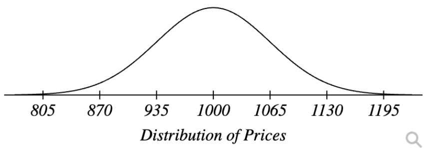
Standard deviation measures the spread of data.

- Gap $\sigma = 2$ (Low spread)
- Old Navy $\sigma = 5$ (High spread)

Because Old Navy has a higher standard deviation, we know that there is **more variation** in the wait times at Old Navy. Overall, wait times at Old Navy are **more spread out** from the average; wait times at the Gap are **more concentrated** near the average.

Problem 9: Normal Distribution Graph

The graph illustrates a normal distribution for the prices paid for a particular model of HD television. The mean price paid is $1000 and the standard deviation is $65.



What is the approximate percentage of buyers who paid less than $805?

0.15 %

What is the approximate percentage of buyers who paid between $1000 and $1195?

49.85 %

What is the approximate percentage of buyers who paid between $935 and $1000?

34 %

What is the approximate percentage of buyers who paid between $935 and $1065?

68 %

What is the approximate percentage of buyers who paid between $870 and $1000?

47.5 %

What is the approximate percentage of buyers who paid more than $1130?

2.5 %

Given: $\mu = \$1000, \sigma = \$65$.

Solutions

1. **Less than $805:** This is below -3σ. The tail area is $(100 - 99.7)/2 = \mathbf{0.15}\%$.
2. **Between $1000 and $1195:** This is from Mean to $+3\sigma$. $99.7/2 = \mathbf{49.85}\%$.
3. **Between $935 and $1000:** This is from -1σ to Mean. $68/2 = \mathbf{34}\%$.
4. **Between $935 and $1065:** This is from -1σ to $+1\sigma$. Empirical rule $= \mathbf{68}\%$.
5. **Between $870 and $1000:** This is from -2σ to Mean. $95/2 = \mathbf{47.5}\%$.
6. **More than $1130:** This is above $+2\sigma$. Tail area is $(100 - 95)/2 = \mathbf{2.5}\%$.

Problem 10: Sample vs. Population Statistics

This data is from a sample. Calculate the mean, standard deviation, and variance.

x	31.8	39.1	25.7	18.7	32.9	26	19.7	45.6

Download CSV

Using technology, calculate the following. Please round the following answers to 2 decimal places.

Sample Mean = 29.94

Sample Standard deviation = 9.30

Sample Variance = 86.58

Oops - now you discover that the data was actually from a population. Which of the values above do you *not* need to recalculate?

You don't need to recalcuate the mean .

Data: $\{31.8, 39.1, 25.7, 18.7, 32.9, 26, 19.7, 45.6\}$ $(n = 8)$

Calculations

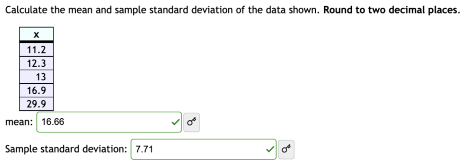
$$\text{Sample Mean } (\bar{x}) = \frac{239.5}{8} = \mathbf{29.94}$$
$$\text{Sample Variance } (s^2) = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{606.038}{7} = \mathbf{86.58}$$
$$\text{Sample Std Dev } (s) = \sqrt{86.58} = \mathbf{9.30}$$

Conceptual Question

When switching from sample to population data, the formulas for Variance and Standard Deviation change (dividing by N instead of $n-1$). However, the formula for the Mean ($\frac{\sum x}{n}$) remains the same.

Answer: You don't need to recalculate the **mean**.

Problem 11: Mean and Sample Standard Deviation (5 values)


Calculate the mean and sample standard deviation of the data shown. **Round to two decimal places.**

x
11.2
12.3
13
16.9
29.9

mean: 16.66

Sample standard deviation: 7.71

Data: $\{11.2, 12.3, 13, 16.9, 29.9\}$ $(n = 5)$

Calculations

1. Mean ($\bar{x}$):

$$\bar{x} = \frac{11.2 + 12.3 + 13 + 16.9 + 29.9}{5} = \frac{83.3}{5} = \mathbf{16.66}$$

2. Sample Standard Deviation (s): First, calculate the sum of squared deviations from the mean (SS):

$$\begin{aligned}(11.2 - 16.66)^2 &= (-5.46)^2 = 29.8116 \\ (12.3 - 16.66)^2 &= (-4.36)^2 = 19.0096 \\ (13.0 - 16.66)^2 &= (-3.66)^2 = 13.3956 \\ (16.9 - 16.66)^2 &= (0.24)^2 = 0.0576 \\ (29.9 - 16.66)^2 &= (13.24)^2 = 175.2976 \\ \text{Sum} &= 237.572\end{aligned}$$

Now, divide by $n - 1$ and take the square root:

$$s = \sqrt{\frac{237.572}{5-1}} = \sqrt{59.393} \approx 7.706 \to \mathbf{7.71}$$

Problem 12: Highway Mileage Statistics

The highway mileage (mpg) for a sample of 8 different models of a car company can be found below. Find the mean, median, mode, and standard deviation. Round to one decimal place as needed.

19, 22, 25, 27, 30, 32, 35, 35

Mean = 28.1

Median = 28.5

Mode = 35

Standard Deviation = 5.9

Data: $\{19, 22, 25, 27, 30, 32, 35, 35\}$ $(n = 8)$

Solutions

Mean:

$$\bar{x} = \frac{225}{8} = 28.125 \to \mathbf{28.1}$$

Median: The middle two numbers (4th and 5th positions) are 27 and 30.

$$\text{Median} = \frac{27 + 30}{2} = \mathbf{28.5}$$

Mode: The number **35** appears most frequently (twice).
Standard Deviation: Using the statistical calculator method or variance formula:

$$s \approx 5.914 \to \mathbf{5.9}$$

Problem 13: TV Watching Hours (Large Sample)

The table below gives the number of hours spent watching TV last week by a sample of 24 children. Find the minimum, maximum, range, mean, and standard deviation of the following data using using Excel, a calculator or other technology. The minimum, maximum, range should be entered as exact values. Round the mean and standard deviation to two decimals places.

41	58	36	57	31	46
86	28	56	52	36	35
27	47	90	55	52	26
85	76	32	54	96	18

Min = 18 ✓ Max = 96 ✓ Range = 78 ✓

Mean = 50.83 ✓ Standard Deviation = 21.98 ✓

Data: Sample of 24 children ($n = 24$).

Calculations

Min/Max/Range: Scanning the table:

- **Min** = **18**
- **Max** = **96**
- **Range** = $96 - 18 = \mathbf{78}$

Mean:

$$\text{Sum of all values} = 1220$$

$$\text{Mean} = \frac{1220}{24} = 50.8333... \rightarrow \mathbf{50.83}$$

Standard Deviation: Using technology (Excel or calculator):

$$s \approx 21.983... \rightarrow \mathbf{21.98}$$