Name Answer Key

Statistics; Sec 3.2 – Measures of Variation

Answer all questions using complete sentences.

1. Which average – mean, median, or mode – is associated with the standard deviation?

   Mean is associated with standard deviation

2. What symbol is used for the standard deviation when it is a sample statistic? What symbol is used for the standard deviation when it is a population parameter?

   statistic: s parameter: σ

3. Consider the data set: 2 3 4 5 6
   a. Find the range.

   4

   b. Use the defining formula to compute the sample standard deviation s.

   $$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{4+1+0+1+4}{4}} \approx \boxed{1.58}$$

   c. Use the defining formula to compute the population standard deviation σ.

   $$\mu = \frac{\Sigma x}{N} = \frac{20}{5} = 4 \qquad \sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}} = \sqrt{\frac{4+1+0+1+4}{5}} \boxed{\approx 1.41}$$

4. In this problem, we explore the effect on the standard deviation of adding the same constant to each data value in a data set (a – c) and the effect on the standard deviation of multiplying each data value in a data set by the same constant (d – e). Consider the data set: 5 9 10 11 15
   a. Use the defining formula, the computation formula, or a calculator to compute s.

   $$\bar{x} = \frac{\Sigma x}{n} = \frac{50}{5} = 10 \qquad s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{25+1+0+1+25}{4}} \approx \boxed{3.61}$$

   b. Add 5 to each data value to get the new data set 10, 14, 15, 16, 20. Compute s.

   $$\bar{x} = \frac{\Sigma x}{n} = \frac{75}{5} = 15 \qquad s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{25+1+0+1+25}{4}} \approx \boxed{3.61}$$

   c. Compare the results of parts (a) and (b). In general, how do you think the standard deviation of a data set changes if the same constant is added to each data value?

   The results are the same. In general, the standard deviation won't change when a constant is added to all the data values.

d. Multiply each data value by 5 to obtain the new data set $25, 45, 50, 55, 75$. Compute s.

$$\bar{x} = \frac{\Sigma x}{n} = \frac{250}{5} = 50 \qquad s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{625+25+0+25+625}{4}} \approx \boxed{18.03}$$

e. Compare the results from parts (a) and (d). In general, how does the standard deviation change if each data value is multiplied by a constant c?

when each value is multiplied by a constant, the standard deviation is that constant times greater.

f. You recorded the weekly distances you bicycled in miles and computed the standard deviation to be $s = 3.1$ miles. Your friend wants to know the standard deviation in kilometers. Do you need to redo all the calculations? Given 1 mile ≈ 1.6 kilometers, what is the standard deviation in kilometers?

No, just take 3.1×1.6

$\boxed{s \approx 4.96 \text{ km}}$

5. Given the sample data

$$x: \quad 23 \quad 17 \quad 15 \quad 30 \quad 25$$

a. Find the range.

15

b. Verify that $\sum x = 110$ and $\sum x^2 = 2568$.

$\Sigma x = 15 + 17 + 23 + 25 + 30 = 110$ ✓

$\Sigma x^2 = 15^2 + 17^2 + 23^2 + 25^2 + 30^2 = 2568$ ✓

c. Use the results of part (b) and appropriate computation formulas to compute the sample variance s^2 and sample standard deviation s.

$$s^2 = \frac{\Sigma x^2 - (\Sigma x)^2/n}{n-1} = \frac{2568 - (110)^2/5}{4} = \boxed{37}$$

$$s = \sqrt{37} \approx \boxed{6.08}$$

d. Use the defining formulas to compute the sample variance s^2 and sample standard deviation s.

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{49+25+1+9+64}{4} = \boxed{37}$$

$$s = \sqrt{37} \approx \boxed{6.08}$$

e. Suppose the given data comprise the entire population of all x values. Compute the population variance σ^2 and population standard deviation σ.

$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N} = \frac{49+25+1+9+64}{5} = \boxed{29.6}$$

$$\sigma = \sqrt{29.6} \approx \boxed{5.44}$$

6. Consider population data with $\mu = 20$ and $\sigma = 2$.
   a. Compute the coefficient of variation.

$$C.V. = \frac{\sigma}{\mu} \times 100\% = \frac{2}{20} \times 100\% = \boxed{10\%}$$

   b. Compute an 88.9% Chebyshev interval around the population mean.

$\mu - 3\sigma$ to $\mu + 3\sigma$

$20-3(2)$ to $20+3(2)$

$\boxed{14 \text{ to } 26}$

7. Do bonds reduce the overall risk of an investment portfolio? Let x be a random variable representing annual percent return for Vanguard Total Stock Index (all stocks). Let y be a random variable representing annual return for Vanguard Balanced Index (60% stock and 40% bond). For the past several years, we have the following data.

x:	11	0	36	21	31	23	24	-11	-11	-21
y:	10	-2	29	14	22	18	14	-2	-3	-10

   a. Compute $\sum x, \sum x^2, \sum y, \sum y^2$.

$\sum x = 103$

$\sum y = 90$

$\sum x^2 = 4607$

$\sum y^2 = 2258$

   b. Use the results of part (a) to compute the sample mean, variance, and standard deviation for x and for y.

x

$\bar{x} = 10.3$

$s^2 = 394.01$

$s = 19.85$

y

$\bar{x} = 9$

$s^2 = 160.89$

$s = 12.68$

   c. Compute a 75% Chebyshev interval around the mean for x values and also for y values. Use the intervals to compare the two funds.

x

$\bar{x} \pm 2s$

$10.3 \pm 2(19.85)$

-29.4 to 50

y

$\bar{x} \pm 2s$

$9 \pm 2(12.68)$

-16.36 to 34.36

Vanguard Balanced Index has a smaller spread.

   d. Compute the coefficient of variation for each fund. Use the coefficients of variation to compare the two funds. If s represents risks and $\bar{x}$ represents expected return, then $s/\bar{x}$ can be thought of as a measure of risk per unit of expected return. In this case, why is a smaller CV better? Explain.

x

$$C.V. = \frac{19.85}{10.3} \times 100\%$$

$$= 192.7\%$$

y

$$C.V. = \frac{12.68}{9} \times 100\%$$

$$= 140.9\%$$

A smaller C.V. is better because it means a lower risk.