Worksheet 3 Solution

Fred Azizi

2023-09-26

1. You've applied to be an operations manager for Coca-Cola Consolidated. In the interview, they ask you we have two different systems of filling a can or a bottle. Which process is working better (in the sense of consistency)?

   i. System 1: A sample of 50 12 oz bottles of Coke which are filled to a number between $11 - 13$ oz. The average of this sample is 12 and the standard deviation is 0.33.

   ii. System 2: A sample of 120 8 oz cans of Coke filled to a number between $7.3 - 8.7$ oz. The average of the sample is 8 and the standard deviation is 0.23.

   Answer:

   While system 2 has a lower standard deviation, comparing standard deviations can be misleading since these are two different type of bottles (8 oz and 12 oz) and standard deviation is affected by the scale of the data. A better option is CV.

$$\text{CV}_{\text{system I}} = \frac{s_I}{\bar{x}_I} = \frac{0.33}{12} = 0.0275$$
$$\text{CV}_{\text{system II}} = \frac{s_{II}}{\bar{x}_{II}} = \frac{0.23}{8} = 0.02875$$

   In this case $\text{CV}_{\text{system I}} < \text{CV}_{\text{system II}}$. Therefore, System I is more consistent.

2. A random sample of 10 students were asked how far they live from UMBC. These were their answers, rounded to the nearest half-mile:

$$1, 2.5, 1.5, 2, 0, 5, 3, 3.5, 1, 0.5$$

   We calculated the sample mean (2) in the previous worksheet.

   a. Compute the sample standard deviation.

   b. Compute the coefficient of variation.

   Answer:

a)

$$\begin{aligned} Var = \frac{1}{n-1}\sum_{i=1}^{10}(x_i - \bar{x})^2 &= \frac{1}{10-1}[(0-2)^2 + (0.5-2)^2 + (1-2)^2 + (1-2)^2 + (1.5-2)^2 + \\ &\quad + (2-2)^2 + (2.5-2)^2 + (3-2)^2 + (3.5-2)^2 + (5-2)^2] \\ &= \frac{1}{9}(4 + 2.25 + 1 + 1 + 0.25 + 0 + 0.25 + 1 + 2.25 + 9) \\ &= \frac{21}{9} \end{aligned}$$

Since sd=$\sqrt{\text{var}}$ $\rightarrow$ sd=$\sqrt{\frac{21}{9}} \approx 1.53$.

b) $CV = \frac{SD}{\bar{x}} \times 100\% = \frac{1.53}{2} \times 100 = 76.5\%$

3. A small sample of height (in cm) is collected as follow:

$$164, 148, 137, 157, 173, 156, 177, 172.$$

a. Determine the location and values of the first, second and third quartiles.

b. Calculate the IQR.

c. Determine the location and value of the 60th percentile. What does the value signify?

Answer:

a. $n = 8$. From slides, we know $L_P = (n+1)\frac{P}{100}$. For Q_1:

$$L_{25} = 9 \times \frac{25}{100} = 2.25$$

So the position of Q_1 is 2.25. Second (ordered) observation is 148. We need to find 0.25 of distance between the 2nd and 3rd ordered observation. $(156 - 148) \times 0.25 = 2$. Therefore:

$$Q_1 = 148 + 2 = 150$$

For Q_2:

$$L_{50} = 9 \times \frac{50}{100} = 4.5$$

Fourth ordered observation is 157. We need to find 0.5 of distance between ordered observation 4 and 5. $(164 - 157) \times 0.5 = 3.5$. Therefore:

$$Q_2 = \text{ Median } = 157 + 3.5 = 160.5$$

For Q_3:

$$L_{75} = 9 \times \frac{75}{100} = 6.75$$

Sixth ordered observation is 172. We need to find 0.75 of distance between ordered observation 6 and 7. $(173 - 172) \times 0.75 = 0.75$. Therefore:

$$Q_3 = 172 + 0.75 = 172.75$$

b. $IQR = Q_3 - Q_1 = 172.75 - 150 = 22.75$

c.

$$L_{60} = 9 \times \frac{60}{100} = 5.4$$

We need to find the .4 of distance between 5th ordered observation and 6th. $(172 - 164) \times 0.4 = 3.2$. Therefore, 60^{th} percentile is $164 + 3.2 = 167.2$.

Interpretation: 60 % of all values in our data is less than or egual lo 167.2.