# Statistics - homework 2

1. (a) range: $8-1=7$

mean deviation:

mean = 4.5

$$M.D = \frac{2\cdot|1-4.5|+3\cdot|3-4.5|+|5-4.5|+|6-4.5|+|7-4.5|+2\cdot|8-4.5|}{10}$$

$= 2.3$

sample variance:

$$\sigma^2 = \frac{2\cdot(1-4.5)^2+3\cdot(3-4.5)^2+(5-4.5)^2+(6-4.5)^2+(7-4.5)^2+2\cdot(8-4.5)^2}{10}$$

$= 6.45$

standard deviation: $\sigma = \sqrt{\sigma^2} = \sqrt{6.45} \approx 2.54$

(b) range: $31-14=17$

Mean deviation:

Mean $\approx 21.9$

$$M.D \approx \frac{|14-21.9|+|15-21.9|+2\cdot|18-21.9|+|27-21.9|+|30-21.9|+|31-21.9|}{7}$$

$\approx 6.41$

sample variance:

$$\sigma^2 \approx \frac{(14-21.9)^2+(15-21.9)^2+2(18-21.9)^2+(27-21.9)^2+(30-21.9)^2+(31-21.9)^2}{7}$$

$\approx 44.98$

standard deviation: $\sigma = \sqrt{\sigma^2} = \sqrt{44.98} \approx 6.71$

4. let c represent the average of the remaining 9 observations.

since we know b and c are of equal size,

$\frac{b+c}{2} = a$ (proven in exercise 1)

$b+c = 2a$

Therefore, $c = 2a-b$

(c) ~~...~~ lets $z_1$ be the standardized score for $X_1$ and let $z_n$ be the

(1) standardized score for $X_n$ $\quad z_1 = \frac{x_1-\bar{x}}{\sigma} \quad z_n = \frac{X_n-\bar{X}}{\sigma}$

Since $x_1, \dots, x_n$ is ordered, $x_1 \leq \bar{x} \leq x_n$

(2) since $\sigma$ is always positive $z_1 \leq 0 \leq z_n$

(3) since $|z_1| = |z_n|$ and, from (2), $z_1 \neq z_n$ (unless $z_1 = z_n = 0$) $z_1 = -z_n$

From (1), we have

(4) $\bar{x} = x_1 - z_1\cdot\sigma = x_n - z_n\cdot\sigma$

(5) $2\bar{x} = x_1 - z_1\cdot\sigma + x_n - z_n\cdot\sigma$

Because of (3), $-2_1 \cdot \sigma$ and $-2n \cdot \sigma$ sum to 0.

(6) $2\bar{x} = x_1 + x_n$

(7) Therefore $\bar{x} = \frac{x_1 + x_n}{2}$

7 a (1) $\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$

(2) $\sigma^2 = \frac{\sum_{i=1}^{n} (x_i^2 - 2x_i\bar{x} + \bar{x}^2)}{n}$

(3) $\sigma^2 = \frac{\sum_{i=1}^{n} (x_i^2) - \sum_{i=1}^{n} (2x_i\bar{x}) + \sum_{i=1}^{n} (\bar{x}^2)}{n}$

(4) $\sigma^2 = \frac{\sum_{i=1}^{n} (x_i^2)}{n} - \frac{2\bar{x} \cdot \sum_{i=1}^{n} (x_i)}{n} + \frac{n \cdot \bar{x}^2}{n}$

(5) $\sigma^2 = \frac{\sum_{i=1}^{n} (x_i^2)}{n} - 2\bar{x} \cdot \bar{x} + \bar{x}^2$

(6) $\sigma^2 = \frac{\sum_{i=1}^{n} (x_i^2)}{n} - \bar{x}^2$

b $\sigma^2 = \frac{1^2 + 5^2 + 2^2 + 3^2 + m^2}{5} - \left(\frac{1+5+2+3+m}{5}\right)^2 = \frac{39+m^2}{5} - \left(\frac{11+m}{5}\right)^2 = 2$ (1)

$39 + m^2 - \frac{(11+m)^2}{5} = 10$ (2)

$39 + m^2 - \frac{m^2 + 22m + 121}{5} = 10$ (3)

$39 + \frac{4m^2}{5} - \frac{22m}{5} - \frac{121}{5} = 10$ (4)

$4m^2 - 22m + 24 = 0$ (5)

$m_1 = \frac{-(-22) + \sqrt{(-22)^2 - 4 \cdot 4 \cdot 24}}{2 \cdot 4} = 4$ (6)

$m_2 = \frac{-(-22) - \sqrt{(-22)^2 - 4 \cdot 4 \cdot 24}}{2 \cdot 4} = \frac{3}{2}$ (7)

ⓒ i. The median of $\{1,5,2,3,4\}$ is 3

ii. The median of $\{1,5,2,3,\frac{3}{2}\}$ is 2

D) ⓐ The maximal salary among the 7% of workers who earn the least is \$17,000.

ⓑ The inter-quartile range is $Q_3 - Q_1 = \$30,375 - \$21,000 = \$9,375$.

ⓒ The median is \$26,000 (also known as the 50th percentile or the 2nd quartile).

ⓓ if every worker receives a 10% increase and then a \$300 bonus.

i, The new average would be $1.1 \cdot \bar{X} + 300 = 1.1 \cdot 26064.20 + 300 = \$28,970.62$

ii, The new median would be $1.1 \cdot \bar{X} + 300 = 1.1 \cdot 26000 + 300 = \$28,900$