Standard Deviation: σ
Variance: σ^2

INTERNATIONAL BACCALAUREATE

Mathematics: analysis and approaches

MAA

EXERCISES [MAA(HL) 4.1-4.3]

STATISTICS – BASIC CONCEPTS

Compiled by Christos Nikolaidis

O. Practice questions

1. [Maximum mark: 9] ***[without GDC]***

Consider the set of data 1, 2, 3, 4, 5.

(a) Find the mean. [2]

(b) Find the variance (by using both the formulas for the variance) [4]

(c) Write down the standard deviation. [1]

(d) Find the interquartile range. [2]

(a) Media $= \frac{5+4+3+2+1}{2} = 7.5$

(b) $\sigma^2 = \frac{(1-3)^2+(2-3)^2+(3-3)^2+(4-3)^2+(5-3)^2}{5} = \frac{10}{5} = 2$

OR

Let 3 "the median" = y

$\sigma^2 = \frac{(1-y)^2+(2-y)^2+(3-y)^2+(4-y)^2+(5-y)^2}{5}$

$\sigma^2 = \frac{1-2y+y^2+4-4y+y^2+9-6y+y^2+16-8y+y^2+25-10y+y^2}{5}$

$\sigma^2 = \frac{55-30y+5y^2}{5} = \frac{(y^2-6y+11)\cancel{5}}{\cancel{5}} = y^2-6y+11$

$9-18+11=2$

$\sigma^2 = \frac{1^2+2^2+3^2+4^2+5^2}{5} - (3^2) = \frac{55}{5} - 9 = 2,$ Not sure how to break down

(c) Standard Deviation $= \sqrt{\sigma^2} = \sigma = \sqrt{2}$

(d) $Q_1 = \frac{1+2}{2} = 1.5,\ Q_3 = \frac{4+5}{2} = 4.5,\ IQR = 4.5-1.5 = 3$

2. [Maximum mark: 10] ***[without GDC]***

Consider the set of data

10, 10, 20, 20, 20, 20, 20, 30, 30, 30

x	frequency
10	2
20	5
30	3

(a) Find the mean. [2]

(b) Find the variance (by using both the formulas for the variance) [5]

(c) Write down the standard deviation. [1]

(d) Find the interquartile range. [2]

(a) $\frac{20+100+90}{10}=\frac{210}{10}=21$

(b) $P=\frac{10}{2}=5$

Median $=20$

$\sigma^2=\frac{2(10-20)^2+3(30-20)^2}{10}=\frac{200+300}{10}=50,\ \sigma=\sqrt{50}$

OR

$\sigma^2=\frac{2(10^2)+5(20^2)+3(30^2)}{10}-(20^2)=\frac{200+2000+2700}{10}-400$

$\sigma^2=490-400=90$

$\sigma=\sqrt{90}$

Variance formulas use mean, not median.

$\sigma^2=\frac{2(10-21)^2+5(20-21)^2+3(30-21)^2}{10}=\frac{242+5+243}{10}=\frac{490}{10}=49$

OR

$\sigma^2=\frac{2(10^2)+5(20^2)+3(30^2)}{10}-(21)^2=\frac{200+2000+2700}{10}-441=490-441=49$

(c) $\sigma=\sqrt{49}=7$

(d) $Q_3=30,\ Q_1=20,\ IQR=30-20=10$

3. [Maximum mark: 9] ***[without GDC]*** No puede asumirse que sea aritmético

Consider the data x_1, x_2, x_3, ..., x_n, with mean $\bar{x}$ and standard deviation s.

(a) If each number is increased by k,

(i) show that the new mean is $\bar{x}+k$ (i.e. it is also increased by k)

(ii) show that the new standard deviation is s (i.e. it remains the same) [4]

(b) If each number is multiplied by k

(i) show that the new mean is $k\bar{x}$ (i.e. it is also multiplied by k)

(ii) show that the new standard deviation is ks (i.e. it is also multiplied by k)

(iii) write down the relation between the original and the new variance. [5]

(a.i) $S_n = \frac{n}{2}(x_1+x_n)$

$\bar{x} = \dfrac{\frac{n}{2}(x_1+x_n)}{n}$

$\bar{x}' = \dfrac{\frac{n}{2}([x_1+k]+[x_n+k])}{n}$

$\bar{x} = \frac{\cancel{n}}{2}(x_1+x_n)\cdot\frac{1}{\cancel{n}}$

$\bar{x}' = \frac{\cancel{n}}{2}(x_1+x_n+2k)\cdot\frac{1}{\cancel{n}}$

$\bar{x} = \frac{1}{2}(x_1+x_n)$

$\bar{x}' = \frac{1}{2}(x_1+x_n+2k)$

$\bar{x}' = \underbrace{\frac{1}{2}(x_1+x_n)}_{\bar{x}}+k$

$\bar{x}' = \bar{x}+k$

(a.ii) $\frac{n}{2}([x_1+k]^2+[x_n+k]^2)$ = Sum of Squares

$\sigma^2 = \dfrac{\frac{\cancel{n}}{2}([x_1+k]^2+[x_n+k]^2)}{\cancel{n}} - \left(\frac{1}{2}(x_1+x_n)+k\right)^2$

$\sigma^2 = \frac{1}{2}\left[(x_1+k)^2+(x_n+k)^2\right] - \frac{1}{4}(x_1^2+2x_1x_n+x_n^2)-k^2$

$\sigma^2 = \frac{1}{2}(x_1^2+x_n^2+2k(x_1+x_n)+2k^2) - \frac{1}{4}x_1^2-\frac{1}{2}x_1x_n-\frac{1}{4}x_n^2-k^2$

$\sigma^2 = \frac{1}{2}x_1^2+\frac{1}{2}x_n^2+kx_1+kx_n+\cancel{k^2}-\frac{1}{4}x_1^2-\frac{1}{2}x_1x_n-\frac{1}{4}x_n^2-\cancel{k^2}$

$\sigma^2 = \frac{1}{4}x_1^2+\frac{1}{4}x_n^2+kx_1+kx_n-\frac{1}{2}x_1x_n$

(a.ii)

A. Exam style questions (SHORT)

4. [Maximum mark: 6] ***[without GDC]***

A fair six-sided die, with sides numbered 1, 1, 2, 3, 4, 5 is thrown. Find the mean and variance of the score.

$\text{Mean} = \frac{16}{6} = \frac{8}{3}$

$\sigma^2 = \frac{2(1^2) + 2^2 + 3^2 + 4^2 + 5^2}{6} - \left(\frac{8}{3}\right)^2$

$\sigma^2 = \frac{2+4+9+16+25}{6} - \frac{64}{9}$

$\sigma^2 = \frac{56}{6} - \frac{64}{9}$

5. [Maximum mark: 6] ***[with GDC]***

Consider the six numbers 2, 3, 6, 9, a and b. The mean of the numbers is 6 and the variance is 10. Find the value of a and of b, if $a < b$.

$6 = \frac{2+3+6+9+a+b}{6}$

$36 = 20 + a + b$

$16 = a + b$ (A)

By GDC: $a = 5$, $b = 11$

$10 = \frac{2^2+3^2+6^2+9^2+a^2+b^2}{6} - 6^2$

$10 = \frac{1}{6}(130 + a^2 + b^2) - 36$

$46 = \frac{1}{6}(130 + a^2 + b^2)$

$276 = 130 + a^2 + b^2$

$146 = a^2 + b^2$ (B)

6. [Maximum mark: 6] ***[with GDC]***

A teacher drives to school. She records the time taken on each of 20 days. She finds that

$$\sum_{i=1}^{20} x_i = 626 \text{ and } \sum_{i=1}^{20} {x_i}^2 = 19780.8,$$

where x_i denotes the time, in minutes, taken on the i-th day.

For this period, calculate

(a) the mean time μ taken to drive to school;

(b) the variance σ^2 of the time taken to drive to school.

(c) the sum $\sum_{x=1}^{20}(x_i-\mu)^2$.

(a) $\frac{626}{20} = 31.3$ minutes

(b) $\sigma^2 = \frac{19780.8}{20} - 31.3^2$

$\sigma^2 = 9.35$

(c) $\sum_{x=1}^{20} x_i^2 - \sum_{x=1}^{20} 2x_i\mu - \sum_{x=1}^{20}\mu^2$

$19780.8 - 2(31.3)\sum_{x=1}^{20} x_i - 20(31.3^2)$

$= -39000.6$

$\sum_{x=1}^{3}(x-1)^2 = 5$

$\sum_{x=1}^{3}(x^2-2x+1) = 5$

$\sum_{x=1}^{3} x^2 - \sum_{x=1}^{3} 2x + 3(1)$

$= 14 - 12 + 3 = 5$

Equivalent when split up!

✓ $\sum_{x=1}^{4} 2x = 2+4+6+8 = 20$

✗ $\sum_{x=1}^{4} 2 + \sum_{x=1}^{4} x = 4(2) + 1+2+3+4$
$= 18$

✓ $2\sum_{x=1}^{4} x = 2(1+2+3+4) = 2(10) = 20$

∘ Constant multiplies can be moved out to the front of summation notation.

7. [Maximum mark: 4] ***[without GDC]***

Ten numbers have mean 9 and standard deviation 2. Find the sum of their squares.

$\frac{\sum_{i=1}^{10} x_i}{10} = 9$ $\qquad \frac{\sum_{i=1}^{10} x_i^2}{10} - (9^2) = 2^2$

$\sum_{i=1}^{10} x_i = 90$ $\qquad \sum_{i=1}^{10} x_i^2 = (4+81)10 = 850$

8. [Maximum mark: 6] ***[with GDC]***

Consider the 10 data items $x_1, x_2, \ldots, x_{10}$. Given that $\sum_{i=1}^{10} x_i^{\ 2} = 1341$ and the standard deviation is 6.9, find the value of $\bar{x}$.

$6.9 = \frac{1341}{10} - \bar{x}$

$\bar{x} = 134.1 - 6.9 = 127.2$

$6.9 = \sigma$

$\sigma^2 = 6.9^2 = 47.61$

$47.61 = 134.1 - \bar{x}^2$

$\bar{x}^2 = 134.1 - 47.61 = 86.49$

$\bar{x} = \sqrt{86.49} = \pm 9.3$

9. [Maximum mark: 8] ***[with GDC]***

Twenty candidates sat an examination in French. The sum of their marks was 826 and the sum of the squares of their marks was 34 132. Two candidates sat the examination late and their marks were a and b. The new mean and variance were calculated, giving the following results:

mean = 42 and variance = 32.

Find a set of possible values of a and b.

$\sum_{i=1}^{20} x_i = 826 + a + b$

$\sum_{i=1}^{20} x_i^2 = 34132 + a^2 + b^2$

$42 = \frac{826 + a + b}{20}$

$a + b = 14\cancel{4}$

$32 = \frac{34132 + a^2 + b^2}{20} - 42^2$

$a^2 + b^2 = 1788\cancel{4}$

By GDC: $\sim -22.07, \sim 36.07$

10*. [Maximum mark: 10] ***[with GDC]***

Consider the set of data

x	**frequency**
2	10
5	20
7	30
a	15
b	25

where a and b are **integers**. The mean of the numbers is 7 and the variance is 5.7.

Find the value of a and of b.

Bes intersection

11*. [Maximum mark: 10] ***[with GDC]***

Consider the set of data

x	**frequency**
2	10
5	20
7	30
8	a
10	b

The mean of the numbers is 7 and the variance is 5.7. Find the value of a and of b.

Bes intersection