LESSON NO. 2: STATISTICS

2001

7 (c) Consider the numbers $1, k, 3k-2, 9$ where $k \in \mathbf{Z}$. The mean of these numbers is $\bar{x}$. The standard deviation is σ.

(i) Express $\bar{x}$ in terms of k.

(ii) Given that $\sigma = \sqrt{20}$, find the value of k.

SOLUTION

7 (c) (i)

WEIGHTED MEAN $\bar{w}$ $\boxed{\bar{w} = \dfrac{\sum wx}{\sum w}}$ 5

$$\bar{x} = \frac{1+k+3k-2+9}{4} = \frac{4k+8}{4} = k+2$$

7 (c) (ii)

$\boxed{\sigma = \sqrt{\dfrac{\text{Sum of (Deviations)}^2}{\text{Number of numbers}}} = \sqrt{\dfrac{\sum d^2}{N}}}$ 6

where deviation $d = (x - \bar{x}) = (\text{Number} - \text{Mean})$

x	d	d^2
1	$k+1$	$(k+1)^2$
k	2	4
$3k-2$	$2k-4$	$(2k-4)^2$
9	$k-7$	$(k-7)^2$

$$\sigma = \sqrt{\frac{(k+1)^2 + 4 + (2k-4)^2 + (k-7)^2}{4}} = \sqrt{20}$$

$$\Rightarrow k^2 + 2k + 1 + 4 + 4k^2 - 16k + 16 + k^2 - 14k + 49 = 80$$

$$\Rightarrow 6k^2 - 28k - 10 = 0 \Rightarrow 3k^2 - 14k - 5 = 0$$

$$\Rightarrow (3k+1)(k-5) = 0 \Rightarrow k = 5 \ (k \in \mathbf{Z})$$

2002

7 (b) The table below shows the prices of various commodities in the year 2000, as a percentage of their prices in 1999. These are called *price relatives*. (For example, the price relative for *Food, Drink & Other Goods* is 105, indicating that the cost of these items was 5% greater in 2000 than in 1999.)

The table also shows the weight assigned to each commodity. The weight represents the importance of the commodity to the average consumer.

Commodity	Weight	Price in 2000 as % of price in 1999
Housing	8	110
Fuel and Transport	19	108
Tobacco	5	116
Services	16	105
Clothing & Durable Goods	10	97
Food, Drink & Other Goods	42	105

(i) Calculate the weighted mean of the price relatives in the table.

(ii) Calculate, correct to two decimal places, the change in the weighted mean if *Tobacco* is removed from consideration.

SOLUTION

7 (b) (i)

WEIGHTED MEAN $\overline{w}$ $\boxed{\overline{w}=\dfrac{\sum wx}{\sum w}}$ 5

Commodity	x	w	wx
Housing	110	8	880
Fuel and Transport	108	19	2052
Tobacco	116	5	580
Services	105	16	1680
Clothing & Durable Goods	97	10	970
Food, Drink & Other Goods	105	42	4410
		100	10572

$$\overline{w}=\frac{\sum wx}{\sum w}=\frac{10572}{100}=105\cdot 72$$

7 (b) (ii)

Commodity	x	w	wx
Housing	110	8	880
Fuel and Transport	108	19	2052
Services	105	16	1680
Clothing & Durable Goods	97	10	970
Food, Drink & Other Goods	105	42	4410
		95	9992

New weighted mean: $\overline{w}_{new} = \dfrac{\sum wx}{\sum w} = \dfrac{9992}{95} = 105\cdot 18$

Change in weighted mean $= 105\cdot 72 - 105\cdot 18 = 0\cdot 54$

2003

7 (c) The mean of the real numbers a and b is $\overline{x}$. The standard deviation is σ.

(i) Express σ in terms of a, b and $\overline{x}$.
(ii) Hence, express σ in terms of a and b only.
(iii) Show that $\overline{x}^2 - \sigma^2 = ab$.

SOLUTION

7 (c) (i)

$$\boxed{\overline{x} = \frac{x_1 + x_2 + \ldots\ldots\ldots\ldots + x_N}{N} = \frac{\text{Sum of the Numbers}}{\text{Number of Numbers}} = \frac{\sum x}{N}} \quad \ldots\ldots\ 3$$

$$\boxed{\sigma = \sqrt{\frac{\text{Sum of (Deviations)}^2}{\text{Number of numbers}}} = \sqrt{\frac{\sum d^2}{N}}} \quad \ldots\ldots\ 6$$

where deviation $d = (x - \overline{x}) = (\text{Number} - \text{Mean})$

x	d	d^2
a	$a - \overline{x}$	$(a - \overline{x})^2$
b	$b - \overline{x}$	$(b - \overline{x})^2$

$$\sigma = \sqrt{\frac{(a - \overline{x})^2 + (b - \overline{x})^2}{2}}$$

7 (c) (ii)

$$\overline{x} = \frac{a+b}{2} \Rightarrow \sigma = \sqrt{\frac{(a - \frac{a+b}{2})^2 + (b - \frac{a+b}{2})^2}{2}} = \sqrt{\frac{(\frac{a-b}{2})^2 + (\frac{b-a}{2})^2}{2}}$$

$$= \sqrt{\frac{(a-b)^2 + (b-a)^2}{8}} = \sqrt{\frac{2(a-b)^2}{8}} = \sqrt{\frac{(a-b)^2}{4}} = \frac{a-b}{2}$$

7 (c) (iii)

$$\overline{x}^2 - \sigma^2 = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2 = \frac{a^2 + 2ab + b^2 - a^2 + 2ab - b^2}{4} = \frac{4ab}{4} = ab$$

2004

7 (c) The mean of the real numbers p, q and r is $\bar{x}$ and the standard deviation is σ.

(i) Show that the mean of the four numbers of p, q, r and $\bar{x}$ is also $\bar{x}$.

(ii) The standard deviation of p, q, r and $\bar{x}$ is k. Show that $k:\sigma = \sqrt{3}:2$.

Solution

7 (c) (i)

$$\bar{x} = \frac{x_1 + x_2 + \dots\dots\dots\dots + x_N}{N} = \frac{\text{Sum of the Numbers}}{\text{Number of Numbers}} = \frac{\sum x}{N}$$ 3

Mean of three numbers $\bar{x} = \dfrac{p+q+r}{3}$

Mean of four numbers $= \dfrac{p+q+r+\bar{x}}{4} = \dfrac{p+q+r+(\frac{p+q+r}{3})}{4} \times \dfrac{3}{3} = \dfrac{3p+3q+3r+p+q+r}{12}$

$= \dfrac{4p+4q+4r}{12} = \dfrac{p+q+r}{3} = \bar{x}$

7 (c) (ii)

$$\sigma = \sqrt{\frac{\text{Sum of (Deviations)}^2}{\text{Number of numbers}}} = \sqrt{\frac{\sum d^2}{N}}$$ 6

where deviation $d = (x - \bar{x}) = (\text{Number} - \text{Mean})$

x	d	d^2
p	$p-\bar{x}$	$(p-\bar{x})^2$
q	$q-\bar{x}$	$(q-\bar{x})^2$
r	$r-\bar{x}$	$(r-\bar{x})^2$

$$\sigma = \sqrt{\frac{\sum d^2}{N}} = \sqrt{\frac{(p-\bar{x})^2 + (q-\bar{x})^2 + (r-\bar{x})^2}{3}}$$

x	d	d^2
p	$p-\bar{x}$	$(p-\bar{x})^2$
q	$q-\bar{x}$	$(q-\bar{x})^2$
r	$r-\bar{x}$	$(r-\bar{x})^2$
$\bar{x}$	$\bar{x}-\bar{x}$	0

$$k = \sqrt{\frac{\sum d^2}{N}} = \sqrt{\frac{(p-\bar{x})^2 + (q-\bar{x})^2 + (r-\bar{x})^2 + 0}{4}}$$

$$\frac{k}{\sigma} = \frac{\sqrt{\dfrac{(p-\bar{x})^2 + (q-\bar{x})^2 + (r-\bar{x})^2}{4}}}{\sqrt{\dfrac{(p-\bar{x})^2 + (q-\bar{x})^2 + (r-\bar{x})^2}{3}}} = \frac{\sqrt{3}}{2} \Rightarrow k:\sigma = \sqrt{3}:2$$

2005

7 (c) On 1st September 2003 the mean age of the first-year students in a school is 12·4 years and the standard deviation is 0·6 years. One year later all of these students have moved into second year and no other students have joined them.

(i) State the mean and the standard deviation of the ages of these students on 1st September 2004. Give a reason for each answer.

A new group of first-year students begins on 1st September 2004. This group has a similar age distribution and is of a similar size to the first-year group of September 2003.

(ii) State the mean age of the combined group of the first-year and second-year students on 1st September 2004.

(iii) State whether the standard deviation of the ages of this combined group is less than, equal to, or greater than 0·6 years. Give a reason for your answer.

SOLUTION

7 (c) (i)

Mean = 13·4 years

As all the students are one year older, the mean is increased by one.

Standard deviation = 0·6 years

The spread of ages about the mean is still the same.

7 (c) (ii)

Mean $\approx \dfrac{12\cdot 4+13\cdot 4}{2}=12\cdot 9$

7 (c) (iii)

Standard deviation > 0·6 years

There is a greater spread of ages in the combined group than in a single year group.

2006

7 (c) The mean of the integers form $-n$ to n, inclusive, is 0. Show that the standard deviation is $\sqrt{\frac{n(n+1)}{3}}$.

SOLUTION

7 (c)

Write out the integers as a list:

$-n, -n+1, -n+2, \ldots\ldots, -2, -1, 0, 1, 2, \ldots\ldots, n-2, n-1, n$

As you can see, the mean of these numbers is zero.

The number of numbers, $N = 2n + 1$

x	d	d^2
$-n$	$-n$	n^2
$-n+1$	$-n+1$	$(n-1)^2$
$-n+2$	$-n+2$	$(n-2)^2$
$\vdots$	$\vdots$	$\vdots$
-2	-2	4
-1	-1	1
0	0	0
1	1	1
2	2	4
$\vdots$	$\vdots$	$\vdots$
$n-2$	$n-2$	$(n-2)^2$
$n-1$	$n-1$	$(n-1)^2$
n	n	n^2
		$\sum d^2 = 2\sum_{r=1}^{n} r^2$

$$\boxed{\sigma = \sqrt{\frac{\text{Sum of (Deviations)}^2}{\text{Number of numbers}}} = \sqrt{\frac{\sum d^2}{N}}} \quad \ldots\ldots\ 6$$

where deviation $d = (x - \bar{x}) = (\text{Number} - \text{Mean})$

$$\sigma = \sqrt{\frac{2\sum_{r=1}^{n} r^2}{(2n+1)}}$$

$$\boxed{\sum_{r=1}^{n} r^2 = S_n = 1^2 + 2^2 + \ldots\ldots + n^2 = \tfrac{n}{6}(n+1)(2n+1)} \quad \ldots\ 8$$

$$\Rightarrow \sigma = \sqrt{\frac{2 \times \frac{n}{6}(2n+1)(n+1)}{(2n+1)}} = \sqrt{\frac{n(n+1)}{3}}$$