Variance and standard deviation (ungrouped data)

Introduction

In this leaflet we introduce variance and standard deviation as measures of spread. We can evaluate the variance of a set of data from the mean that is, how far the observations deviate from the mean. This deviation can be both positive and negative, so we need to square these values to ensure positive and negative values do not simply cancel each other out when we add up all the deviations.

Variance

The **variance** of a set of values, which we denote by σ^2, is defined as

$$\sigma^2 = \frac{\sum(x-\bar{x})^2}{n}$$

where $\bar{x}$ is the mean, n is the number of data values, and x stands for each data value in turn. Recall that $\sum x$, for example, means add up all the values of x. Similarly, $\sum(x-\bar{x})^2$ means subtract the mean from each data value, square, and finally add up the resulting values. (If necessary revise the leaflet *Sigma Notation*).

An alternative, yet equivalent formula, which is often easier to use is

$$\sigma^2 = \frac{\sum x^2}{n} - \bar{x}^2$$

Worked example

Find the variance of $6, 7, 10, 11, 11, 13, 16, 18, 25$.

Firstly we find the mean, $\bar{x} = \frac{\sum x}{n} = \frac{117}{9} = 13$.

Method 1:

$$\sigma^2 = \frac{\sum(x-\bar{x})^2}{n}$$

It is helpful to show the calculation in a table:

x	6	7	10	11	11	13	16	18	25	Total
$x-\bar{x}$	-7	-6	-3	-2	-2	0	3	5	12	
$(x-\bar{x})^2$	49	36	9	4	4	0	9	25	144	280

$$\begin{aligned}
\sum(x-\bar{x})^2 &= 280 \\
\sigma^2 &= \frac{\sum(x-\bar{x})^2}{n} \\
&= \frac{280}{9} \\
&= 31.11 \quad \text{(2dp)}
\end{aligned}$$

Method 2:

$$\sigma^2 = \frac{\sum x^2}{n} - \bar{x}^2$$

x	6	7	10	11	11	13	16	18	25	Total
x^2	36	49	100	121	121	169	256	324	625	1801

$$\begin{aligned}\sigma^2 &= \frac{\sum x^2}{n} - \bar{x}^2 \\ &= \frac{1801}{9} - 13^2 \\ &= 200.11 - 169 \\ &= 31.11 \quad \text{(2dp)}\end{aligned}$$

Standard Deviation (σ)

Since the variance is measured in terms of x^2, we often wish to use the standard deviation where $\sigma = \sqrt{\text{variance}}$. The standard deviation, unlike the variance, will be measured in the same units as the original data.

In the above example $\sigma = \sqrt{31.11} = 5.58 \quad \text{(2 dp)}$

Exercises

Find the variance and standard deviation of the following correct to 2 decimal places:

1. a) $10, 16, 12, 15, 9, 16, 10, 17, 12, 15$

   b) $74, 72, 83, 96, 64, 79, 88, 69$

   c) £326, £438, £375, £366, £419, £424

Answers

1. a) 7.76, 2.79 b) 97.36, 9.87 c) £^{2}1 531.22, £39.13