2.5) Coding

Worked example

Prove that if all values of x are multiplied by 5, the variance increases by a scale factor of 25

Your turn

Prove that if all values of x are multiplied by 3, the variance increases by a scale factor of 9

$$\sigma^2 = \frac{\Sigma(3x)^2}{n} - \left(\frac{\Sigma(3x)}{n}\right)^2$$

$$= \frac{\Sigma 9x^2}{n} - \left(\frac{3\Sigma x}{n}\right)^2$$

$$= \frac{9 \cdot \Sigma x^2}{n} - 9\left(\frac{\Sigma x}{n}\right)^2$$

$$= 9\left(\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2\right)$$

Worked example

Scores, x:

$$2090, 2080, 2070, 2060, 2050$$

a) Use the coding $y = x - 2000$ to code this data
b) Calculate the mean and standard deviation of the coded data
c) Use your answer to b) to calculate the mean and standard deviation of the original data

Your turn

Scores, x:

$$1010, 1020, 1030, 1040, 1050$$

a) Use the coding $y = x - 1000$ to code this data
b) Calculate the mean and standard deviation of the coded data
c) Use your answer to b) to calculate the mean and standard deviation of the original data

a) y: $10, 20, 30, 40, 50$
b) $\bar{y} = 30$, $\sigma_y = 14.1$ (3 sf)
c) $\bar{x} = 1030$, $\sigma_x = 14.1$ (3 sf)

Worked example

Scores, x:

$$2090, 2080, 2070, 2060, 2050$$

a) Use the coding $y = 2x$ to code this data

b) Calculate the mean and standard deviation of the coded data

c) Use your answer to b) to calculate the mean and standard deviation of the original data

Your turn

Scores, x:

$$1010, 1020, 1030, 1040, 1050$$

a) Use the coding $y = 3x$ to code this data

b) Calculate the mean and standard deviation of the coded data

c) Use your answer to b) to calculate the mean and standard deviation of the original data

a) y: $3030, 3060, 3090, 3120, 3150$

b) $\bar{y} = 3090$, $\sigma_y = 42.4$ (3 sf)

c) $\bar{x} = 1030$, $\sigma_x = 14.1$ (3 sf)

Worked example

Scores, x:

$$2090, 2080, 2070, 2060, 2050$$

a) Use the coding $y = \frac{x}{5}$ to code this data

b) Calculate the mean and standard deviation of the coded data

c) Use your answer to b) to calculate the mean and standard deviation of the original data

Your turn

Scores, x:

$$1010, 1020, 1030, 1040, 1050$$

a) Use the coding $y = \frac{x}{10}$ to code this data

b) Calculate the mean and standard deviation of the coded data

c) Use your answer to b) to calculate the mean and standard deviation of the original data

a) y: $101, 102, 103, 104, 105$

b) $\bar{y} = 103$, $\sigma_y = 1.41$ (3 sf)

c) $\bar{x} = 1030$, $\sigma_x = 14.1$ (3 sf)

Worked example

Scores, x:

$$2090, 2080, 2070, 2060, 2050$$

a) Use the coding $y = \frac{x-2000}{10}$ to code this data

b) Calculate the mean and standard deviation of the coded data

c) Use your answer to b) to calculate the mean and standard deviation of the original data

Your turn

Scores, x:

$$1010, 1020, 1030, 1040, 1050$$

a) Use the coding $y = \frac{x-1000}{10}$ to code this data

b) Calculate the mean and standard deviation of the coded data

c) Use your answer to b) to calculate the mean and standard deviation of the original data

a) y: $1, 2, 3, 4, 5$

b) $\bar{y} = 3$, $\sigma_y = 1.41$ (3 sf)

c) $\bar{x} = 1030$, $\sigma_x = 14.1$ (3 sf)

Worked example

Scores, x:

$$2090, 2080, 2070, 2060, 2050$$

a) Use the coding $y = \frac{x}{10} - 200$ to code this data

b) Calculate the mean and standard deviation of the coded data

c) Use your answer to b) to calculate the mean and standard deviation of the original data

Your turn

Scores, x:

$$1010, 1020, 1030, 1040, 1050$$

a) Use the coding $y = \frac{x}{10} - 100$ to code this data

b) Calculate the mean and standard deviation of the coded data

c) Use your answer to b) to calculate the mean and standard deviation of the original data

a) y: $1, 2, 3, 4, 5$

b) $\bar{y} = 3$, $\sigma_y = 1.41$ (3 sf)

c) $\bar{x} = 1030$, $\sigma_x = 14.1$ (3 sf)

Worked example

Scores, x, of 20 people were recorded.

The data was coded using $y = \frac{x-10}{5}$ and the following summations were obtained:

$$\sum y = 23 \, , \sum y^2 = 147.6$$

Calculate the standard deviation of the actual scores.

Your turn

Scores, x, of 40 people were recorded.

The data was coded using $y = \frac{x-5}{10}$ and the following summations were obtained:

$$\sum y = 32 \, , \sum y^2 = 764.1$$

Calculate the standard deviation of the actual scores.

$$\sigma_x = 42.97 \text{ (2 dp)}$$

Worked example

Scores, x, of 20 people were recorded.
The data was coded using $y = 5x - 10$ and the following summations were obtained:

$$\sum y = 23 \, , \sum y^2 = 147.6$$

Calculate the standard deviation of the actual scores.

Your turn

Scores, x, of 40 people were recorded.
The data was coded using $y = 10x - 5$ and the following summations were obtained:

$$\sum y = 32 \, , \sum y^2 = 764.1$$

Calculate the standard deviation of the actual scores.

$$\sigma_x = 0.4297 \, (4 \text{ dp})$$

Worked example

A teacher standardises scores, x, of his class by adding 10 to each score and then reducing the score by 8%.

The following summary statistics are calculated for the standardised scores, y:

$$n = 30, \bar{y} = 23.4, S_{yy} = 5.6$$

Calculate the mean and standard deviation of the original scores

Your turn

A teacher standardises scores, x, of his class by adding 8 to each score and then reducing the score by 10%.

The following summary statistics are calculated for the standardised scores, y:

$$n = 25, \bar{y} = 43.2, S_{yy} = 6.5$$

Calculate the mean and standard deviation of the original scores

Mean $= \bar{x} = 40$

Standard deviation $= \sigma_x = 0.567$ (3 sf)

Worked example

A person was recording times, x, in a race.
They then realised the stopwatch was slow and all times were actually 3 minutes more than the recorded times.
Explain the effect on:

a) The mean
b) The standard deviation
c) The median
d) The range
e) The lower quartile
f) The interquartile range

Your turn

A person was recording times, x, in a race.
They then realised the stopwatch was slow and all times were actually 5 minutes less than the recorded times.
Explain the effect on:

a) The mean
b) The standard deviation
c) The median
d) The range
e) The lower quartile
f) The interquartile range

a) Decreases by 5
b) No effect
c) Decreases by 5
d) No effect
e) Decreases by 5
f) No effect