7.1 Coding of Mean and Standard Deviation

Example

The number of pages in six randomly chosen novels was recorded as being;

The Shipping News	Annie Proulx	337
Deadly Décisions	Kathy Reichs	333
Birdsong	Sebastian Faulks	305
My Legendary GirlFriend	Mike Gayle	353
Enduring Love	Ian McEwan	249
Black Notice	Patricia Cornwell	413

(i) Calculate the Mean and Standard Deviation of this original data, μ_o and σ_o

	Data, x	x^2
	337	113569
	333	110889
	305	93025
	353	124609
	249	62001
	413	170569
Σ	1990	674662

$$\mu = \frac{\Sigma x}{f}$$ (Use the exact value in subsequent calculations)

$$\sigma = \sqrt{\frac{\Sigma x^2}{f} - \mu^2}$$

The idea of coding is to find a transformation of the form,

$$X_t = \frac{x_o - a}{b}$$

that can be applied to the original data.
In this equation

- x_o is the original data
- a and b are suitable constants
- X_t is the transformed data

To be useful, the transformed data should be easier to manipulate and, in particular, finding the mean and standard deviation of the transformed data μ_t and σ_t should be a relatively easy task.
These are then decoded to give the mean and standard deviation of the original data.

(ii) Code the example data using the transformation

$$X_t = \frac{x_o - 301}{4}$$

x_o	X_t	X_t^2
337		
333		
305		
353		
249		
413		
Σ		

(iii) For the coded (transformed) data find the mean and the standard deviation. That is, find μ_t and σ_t

(iv) Use the value of the the mean of the coded data, and the fact that it was coded using the coding

$$X_t = \frac{x_o - 301}{4}$$

to find the mean, μ_o, of the original data.

(v) Use the value of the the standard deviation of the coded data, and the fact that it for coded using the coding

$$X_t = \frac{x_o - 301}{4}$$

to find the standard deviation, σ_o, of the original data.

7.2 Summary

For data that has been coding using a transformation of the form,

$$X_t = \frac{x_o - a}{b}$$

and given the mean and standard deviation of the coded (transformed) data, μ_t and σ_t the original mean and standard deviation is retrieved using the facts that,

$$u_o = b\,\mu_t + a$$

$$\sigma_o = b\,\sigma_t$$

7.3 Exercise

Question 1

Data is coded using the transformation

$$y = \frac{x - 250}{80}$$

The mean of the coded data is 4.5

The standard deviation of the coded data is 3.2

Find the mean and the standard deviation of the original data.

Question 2

The annual income, *I*, of 100 IT consultants was recorded.

The data were coded using

$$y = \frac{I - 400}{10\,000}$$

and the following summations were obtained;

$$\Sigma y = 887 \qquad \Sigma y^2 = 8065$$

(i) Calculate the mean and the standard deviation of the coded data.

(ii) Calculate the mean and the standard deviation of the original data.

(iii) Most incomes will be within two standard deviations of the mean.
Calculate $\mu \pm 2 \times \sigma$ and so state the values between which most incomes lie.

Question 3

Mrs Crump gathers information about the times taken by 29 girls to get from Emma Darwin Hall to Kingsland Hall for breakfast one morning.

Time minutes	N° girls f	Mid-Interval minutes t	Coding $y = (t - 7) / 3$ y	y^2	fy	fy^2
$2 \leqslant t < 6$	5					
$6 \leqslant t < 8$	14					
$8 \leqslant t < 12$	8					
$12 \leqslant t < 20$	2					
		Σ				

(i) Fill in the empty cells on the table above, where the data is coded using

$$y = \frac{t - 7}{3}$$

(ii) Use Σfy to find mean and Σfy^2 to find standard deviation of the coded data.

(iii) Find the mean and standard deviation of the original data.

Question 4

S1 Examination Question from January 2012 Q4

The marks, x, of 45 students randomly selected from those students who sat a mathematics examination are shown below.

36	39	39	40	41	42	42	43	44	45
46	46	46	48	50	52	53	53	54	54
55	55	56	57	57	59	60	60	60	60
61	63	64	64	64	65	65	66	67	68
69	71	72	73	73					

(a) Write down the modal mark of these students.

[1 mark]

(b) Find the values of the lower quartile, the median and the upper quartile.

[3 marks]

For these students $\Sigma x = 2497$ and $\Sigma x^2 = 143\,369$

(c) Find the mean and the standard deviation of the marks of these students.

[3 marks]

(d) Describe the skewness of the marks of these students, giving a reason for your answer.

[2 marks]

The mean and the standard deviation of the marks of all the students who sat the examination where 55 and 10 respectively. The examiners decided that the total mark of each student should be scaled by subtracting 5 marks and then reducing the mark by a further 10 %.

(e) Find the mean and the standard deviation of the scaled marks of all the students.

[4 marks]