Lecture 6 (Chapter 4 Continued ...)

Change Of Origin and Scale method:

The following formula is used to calculate the standard deviation by change of origin and scale or short method.

$$S = h\sqrt{\frac{\sum f_i u_i^2}{n} - \left(\frac{\sum f u_i}{n}\right)^2}$$

Example:

Find the standard deviation using short method.

x_i	f_i	u_i	$f_i u_i$	$f_i u_i^2$
74.5	9	-2	-18	36
94.5	10	-1	-10	10
114.5	17	0	0	0
134.5	10	1	10	10
154.5	5	2	10	20
174.5	4	3	12	36
194.5	5	4	20	80
$\Sigma =$	60		24	192

$$u_i = \frac{x_i - a}{h} \qquad a = 114.5$$

$$u_i = \frac{74.5 - 114.5}{20} \qquad h = 20$$

Therefore

$$S = h\sqrt{\frac{\sum f_i u_i^2}{n} - \left(\frac{\sum f u_i}{n}\right)^2}$$

$$= 20\sqrt{\frac{192}{60} - \left(\frac{24}{60}\right)^2}$$

$$= 20\sqrt{3.20 - 0.16}$$

$$= 20\sqrt{3.04} = 20 \times 1.7436 = 34.87 \text{ grams}$$

Co-efficient Of Variation:

Karl Pearson introduced a relative measure of variation, Known as the co-efficient of variation, abbreviated as C.V, which expresses the Standard deviation as a percentage of the arithmetic mean of a data set.

Symbolically

$$C.V = \frac{S}{\bar{x}} \times 100, \text{ for Sample of data}$$

$$= \frac{\sigma}{\mu} \times 100, \text{ for population data}$$

Example:

Using the co-efficient of variation, determine whether or not there is a greater variation among the prices of certain similar commodities given, than among the life in hours under test

Price in Rupees	8	13	18	23	30
Life in hours	130	150	180	250	345

Price in Rupees X	X^2	Life in hours Y	Y^2
8	64	130	16900
13	169	150	22500
18	324	180	32400
23	529	250	62500
30	900	345	119025
92	1986	1055	253325

$$\bar{X} = \frac{92}{5} = 18.4 \qquad \bar{Y} = \frac{1055}{5} = 211$$

$$S_X = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{\frac{1986}{5} - \left(\frac{92}{5}\right)^2}$$

$$= \sqrt{397.2 - 338.56}$$

$$= \sqrt{58.44}$$

$$= 7.66$$

$$C.V = \frac{S}{\bar{X}} \times 100$$

$$= \frac{7.66}{18.4} \times 100$$

$$= 41.63\%$$

$$S_Y = \sqrt{\frac{\sum Y^2}{n} - \left(\frac{\sum Y}{n}\right)^2}$$

$$= \sqrt{\frac{253325}{5} - \left(\frac{1055}{5}\right)^2}$$

$$= \sqrt{50665 - 44521}$$

$$= \sqrt{6144}$$

$$= 78.38$$

$$C.V = \frac{S}{\bar{Y}} \times 100$$

$$= \frac{78.38}{211} \times 100$$

$$= 37.15\%$$

So 41.63% is larger than 37.15% So the prices of certain commodities are showing greater variation than that among the life in hours.

Example: Goals scored by two teams in a football season are as follows

No of goals	Number of matches	
x_i	A	B
0	27	17
1	9	9
2	8	6
3	5	5
4	4	3

By calculating the co-efficient of variation in each case, find which team is more consistent.

No. of Goals	Team A			Team B		
x_i	f_i	f_ix_i	$f_ix_i^2$	f_i	f_ix_i	$f_ix_i^2$
0	27	0	0	17	0	0
1	9	9	9	9	9	9
2	8	16	32	6	12	24
3	5	15	45	5	15	45
4	4	16	64	3	12	48
Total	53	56	150	40	48	126

Team A:

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{56}{53} = 1.06$$

$$S = \sqrt{\frac{\sum f_ix_i^2}{n} - \left(\frac{\sum f_ix_i}{n}\right)^2}$$

$$= \sqrt{\frac{150}{53} - \left(\frac{56}{53}\right)^2}$$

$$= \sqrt{1.7138}$$

$$= 1.308$$

$$C.V = \frac{S}{\bar{x}} \times 100$$

$$= \frac{1.308}{1.06} \times 100$$

$$= 123.4\%$$

Team B:

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{48}{40} = 1.20$$

$$S = \sqrt{\frac{\sum f_ix_i^2}{n} - \left(\frac{\sum f_ix_i}{n}\right)^2}$$

$$= \sqrt{\frac{126}{40} - \left(\frac{48}{40}\right)^2}$$

$$= \sqrt{1.71}$$

$$= 1.308$$

$$C.V = \frac{S}{\bar{x}} \times 100$$

$$= \frac{1.308}{1.20} \times 100$$

$$= 109.0\%$$

Hence Team B is more consistent.

Moments

A moment designates the power to which deviations are raised before averaging them. The quality

$$\frac{\sum(x_i - \mu)}{N}$$

is called the first population moment and is denoted by μ_1.

Similarly the quantity $\frac{\sum(x_i - \mu)^2}{N}$ is called the second population moment and is denoted by μ_2.

The corresponding sample moments are denoted by m_1 and m_2.

Generally moments are defined as

$$\mu_r = \frac{1}{N}\sum(x_i - \mu)^r, \text{ for population data}$$

$$m_r = \frac{1}{n}\sum(x_i - \bar{x})^r, \text{ for sample data}$$

These are also called central moments or the mean moments, and are used to describe a set of data.

Example: Calculate the first four moments about the mean for the following data:

45, 32, 37, 46, 39, 36, 41, 48 & 36

(P.T.O)

SQ :=

x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$(x_i-\bar{x})^3$	$(x_i-\bar{x})^4$
32	-8	64	-512	4096
36	-4	16	-64	256
36	-4	16	-64	256
37	-3	9	-27	81
39	-1	1	-1	1
41	1	1	1	1
45	5	25	125	625
46	6	36	216	1296
48	8	64	512	4096
360	0	232	186	10708

$$\bar{x} = \frac{\sum x}{n} = \frac{360}{9} = 40$$

$$m_1 = \frac{\sum(x_i-\bar{x})}{n} = 0$$

$$m_2 = \frac{\sum(x_i-\bar{x})^2}{n} = \frac{232}{9} = 25.78$$

$$m_3 = \frac{\sum(x_i-\bar{x})^3}{n} = \frac{186}{9} = 20.67$$

$$m_4 = \frac{\sum(x_i-\bar{x})^4}{n} = \frac{10708}{9} = 1189.78$$