1. Problem's on Pearson's Coefficient of Skewness

Welcome to the series of E-learning module on problems on moments, skewness and kurtosis.

At the end of this session, you will be able to understand the calculation of:

- Skewness and its various measures
- Moments and
- kurtosis

Let us start with an introduction:
Skew is the indicator of lack of symmetry in a distribution. When the mean, median and mode of the distribution do not have the same value in a distribution we call it as a skewed distribution.
Kurtosis is a greek word which means "bulginess". In statistics we use kurtosis to help in identifying the degree of flatness or peakedness of frequency curve in the region about the mode. The degree of kurtosis of a distribution is measured relative to the peakedness of normal curve. This concept of kurtosis is rarely used in elementary statistical analysis.
The word Moment in statistics is used to describe the characteristics of a frequency distribution like central tendency, variation, skeweness and kurtosis. The utility of the moment lies in the sense that they indicate the different aspects of a given distribution.
Problem 1
From the following data, calculate the measure of skewness using the mean, median and standard deviation:

Figure 1

X	10-20	20-30	30-40	40-50	50-60	60-70	70-80
f	18	30	40	55	38	20	16

First row represents class interval (x) and second row represents frequency of the data (f).

Solution:

Figure 2

x	mv	d	f	fd	fd²	cf
10-20	15	-3	18	-54	162	18
20-30	25	-2	30	-60	120	48
30-40	35	-1	40	-40	40	88
40-50	45	0	55	0	0	143
50-60	65	1	38	38	38	181
60-70	75	2	20	40	80	201
70 - 80	85	3	16	48	144	217
Total			217	-28	584	

In this table, we have the mid value of the class interval in column 2.
Column 3 is the deviation of the mid value from the arbitrary mean.
Column 4 is the frequency of the data.
In column 5, we have multiplied the frequency with the deviation.
In column six, we have squared the deviation and then multiplied with the frequency and in column 7, we have calculated the cumulative frequency.

Now let us calculate the mean and the median using the formulas.
Mean is equal to arbitrary mean plus sigma of the product of frequency and deviation by N multiplied by the width of the class interval which is equal to 45 plus minus 28 divided by 217 into 10 is equal to 43 point 71.

The median class is equal to N by 2 is equal to 217 by 2 is equal to 108 point 5th item. Thus, the median class is equal to 40 to 50. let us use the formula to interpolate the median from the median class. Calculate the median as follows
Median is equal to lower limit plus 'N' by 2 minus 'cf' divided by 'f' into 'c'

Where,
L one = lower limit of the median class.
cf = cumulative frequency of the class preceding the median class.
f= simple frequency of the median class.
c=magnitude of the median class.

Substituting the values, median is equal to 40 plus 108 point 5 minus 88 divided by 55 into 10 is equal to 43 point 72.

Standard deviation is equal to the square root of sigma of the frequency deviation square divided by the total of frequency minus sigma of the frequency deviation divided by the total of the frequency whole square multiplied by width of the class interval.
Substituting the values, we get square root of 584 divided by 217 minus 0f minus 28 divided by 217 whole square multiplied by 10 is equal to 16 point 4.

Skewness is equal to thrice the mean minus median is equal to 3 into 43 point 71minus 43 point 72 is equal to 3 into minus 0 point 01 is equal to minus 0 point 03.
The coefficient of skewness is equal to the skewness divided by the standard deviation is equal to minus 0 point 03 divided by 16 point 4 is equal to minus 0 point 001.
The result shows that the distribution is negatively skewed, but the extent of the skewness is extremely negligible.

2. Problem's on Bowley's & Kelly's Coefficient of Skewness

Problem 2:
From the following data, calculate an appropriate measure of skewness.

Figure 3

Value in Rs.	<50	50-100	100-150	150-200	200 & above
f	40	80	130	60	30

First row represents the values in rupees and second row represents the frequency of the value.

Solution:

Figure 4

Value in Rs.	<50	50-100	100-150	150-200	200 & above
f	40	80	130	60	30
Cumulative Frequency	40	40+80 =120	120+130 =250	250+60 =310	310+30 =340

It should be noted that the given series is an open ended series and hence the Bowley's coefficient of skewness based on quartiles is the appropriate measure of skewness.
Since we have to calculate the quartiles we should first calculate the cumulative frequency as shown in the table.

Calculation of the quartiles are as shown below:
The lower quartile is found in the class interval 50 to 100 as the total frequency divided by four is equal to 85.
Now using the formula Q one is equal to 'lower limit plus 'N' divided by 4 minus the cumulative frequency of the preceding class divided by the frequency of the quartile class into the class width.
By substituting the value, we get 78 point 125. Similarly calculation is done for the median and the upper quartile.

The upper quartile is found in the class interval 150 to 200 as the total frequency divided by three fourth is equal to 255.
Now using the formula 'q' three is equal to lower limit plus the total frequency divided by three

fourth minus the cumulative frequency of the preceding class divided by the frequency of the quartile class into the class width.
By substituting the value, we get 154 point 15.

Bowleys coefficient of skewness is equal to upper quartile plus the lower quartile minus twice the median divided by the difference between the upper quartile and the lower quartile, when the values are substituted we get the coefficient as minus 0 point 081.
This shows that there is a negative skewness, which is very negligible magnitude.

Problem 3
From the following data, use kelly's measure of skewness to calculate.

Figure 5

Class interval	10-20	20-30	30-40	40-50	50-60	60-70	70-80
f	18	30	40	55	38	20	16

First row represents the class interval and second row represents the frequency of the data.

Solution:

Figure 6

Class interval	10-20	20-30	30-40	40-50	50-60	60-70	70-80
f	18	30	40	55	38	20	16
Cf	18	48	88	143	181	201	217

The Kelly's coefficient of skewness based on percentiles and Deciles to measure the skewness.
Since we have to calculate the percentiles and deciles we should first calculate the cumulative frequency by adding the frequencies as shown in the table.

Calculation of skewness from percentiles are as shown in figure:

The tenth percentile is found in the class interval 20 to 30 as the total frequency divided by ten is equal to 21 point 7. Now using the formula 'p' ten is equal to lower limit plus the total frequency divided by ten minus the cumulative frequency of the preceding class divided by the frequency of the percentile class into the class width.
By substituting the value, we get 21 point 23. Similarly calculation is done for fifty percentile

and ninety percentile.

Calculation of kellys coefficient of skewness for percentile is equal to ninety percentile plus tenth percentile minus twice fifty percentile divided by the difference between the ninety percentile and the tenth percentile.
By substituting the values, we get the skewness is equal to 0 point 02.
This shows that the series is positively skewed though the extent of skewness is extremely negligible.
Zero skewness gives a symmetrical distribution.

3. Problem's on Kurtosis

Problem 4
Calculation of the coefficient of skewness and kurtosis based on the following data.

Figure 7

X	4.5	14.5	24.5	34.5	44.5	54.5	64.5	74.5	84.5	94.5
F	1	5	12	22	17	9	4	3	1	1

First row represents the values distribution 'x' and second row represents the frequency.

Solution:
As we are going to calculate the coefficient of skewness and kurtosis from the moments, let us calculate the moments for the given data.
First let us prepare the various measures needed for calculation.

Figure 8

X	f	d	fd	fd^2	fd^3	fd^4
4.5	1	-4	-4	16	-64	256
14.5	5	-3	-15	45	-135	405
24.5	12	-2	-24	48	-96	192
34.5	22	-1	-22	22	-22	22
44.5	17	0	0	0	0	0
54.5	9	1	9	9	9	9
64.5	4	2	8	16	32	64
74.5	3	3	9	27	81	243
84.5	1	4	4	16	64	256
94.5	1	5	5	25	125	625
Total	**75**		**-30**	**224**	**-6**	**2072**

In this table the first column indicates the variable.
The second column indicates the frequency of the data and its total is taken as 75.
The third column indicates the deviation of the variable from the mean divided by 10 the width of the interval.
Fourth Column indicates the product of the deviation with the frequency and the total is minus 30.
Fifth Column indicates the product of the frequency with the square of the deviations having a total of 224.
In the Sixth column, we calculate the product of the cube of the deviation with the frequency and have a total of minus 6.
Finally the last column is the product of the fourth power of the deviation and the frequency

with a total of 2 thousand 72.

Calculation of the moments from the arbitrary mean, we get mu one dash is equal to sigma of frequency into deviation divided by N is equal to minus 30 divided by 75 is equal to minus 0 point 4.
Similarly, we calculate the mu two dash as 2 point 99, mu three dash as minus 0 point 08 and mu four dash as 27 point 63.

Calculation of the central moments, mu one is equal to zero, then we need the second, third and fourth moment.
mu two is equal to mu two dash minus mu one dash square is equal to 2 point 99 minus of minus 0 point 4 square is equal to 2 point 99 minus 0 point 16 is equal to 2 point 83.
similarly when we calculate the third and the fourth moment we get the third moment as 3 point 38 and the fourth moment is equal to 30 point 295.

Calculation of the skewness and the kurtosis:
Skewness is equal to beta one is equal to minus mu three square divided by mu two cube is equal to 3 point 38 square divided by 2 point 83 square is equal to 11 point 424 divided by 22 point 665 is equal to 0 point 504.
Kurtosis is equal to beta two is equal to mu four divided by mu two square is equal to 30 point 295 divided by 2 point 83 square is equal to 30 point 295 divided by 8 point 01 is equal to 3 point 782.

Skewness is 0 point 504 and hence the distribution is asymmetrical and the curve is leptokurtic as the kurtosis is above 3.

Problem 5
The first four central moments of distribution are 0, 2 point 5, 0 point 7 and 18 point 75. Comment on the skewness and kurtosis of the distribution.

Solution:
We are given that mu one is zero, mu two is 2 point 5, mu three is 0 point 7 and mu four is 18 point 75.
Testing the skewness & Kurtosis:
We know that the skewness & kurtosis is measured with the beta one coefficient & beta two coefficient respectively,
Beta one is equal to mu three square divided by the mu two cube which is equal to 0 point 7 square divided by 2 point 5 cube. Therefore beta one is equal to 0 point 031
Beta two is equal to mu four divided by mu two square which is equal to 18 point 75 divided by 2 point 5 square is equal to 3.
Since beta one is equal to 0 point 031, the distribution is slightly skewed that is it is not perfectly symmetrical.
Since beta two is equal to 3 the distribution is mesokurtic.

4. Problem’s on Moments: Part-1

Problem 6
Find the first, second third and fourth moments for the set of numbers 2,3,4,5 and 6.

Solution:
To calculate the first moment that is mean, we take the total of the variables and divided it by the total number of data.
So in this example we find the mean, that is ‘x’ bar is equal to two plus three plus four plus five plus six divided by five is equal to twenty divided by five is equal to four.

The second moment is equal to two square plus three square plus four square plus five square plus six square divided by five is equal to ninety by five is equal to eighteen.
The third moment is equal to cube of two plus cube of three plus cube of four plus cube of five plus cube of six divided by five is equal to four hundred forty by five is equal to eighty eight.

The fourth moment is equal to the fourth power of two plus the fourth power of three plus the fourth power of four plus the fourth power of five plus the fourth power of six divided by five is equal to two thousand two hundred and seventy four by five is equal to four hundred fifty four point eight.

Figure 9

(X)	$(\overline{X})$	$(X-\overline{X})$
2	4	2-4 = -2
3	4	3-4 = -1
4	4	4-4 = 0
5	4	5-4 = 1
6	4	6-4 = 2
N=5		$\Sigma(X-\overline{X}) = 0$

Using the same data let us calculate the first, second, third and fourth moments about the mean.
The first moment about the mean mu one is calculated by subtracting the first moment from the variable that is in this case we will find the total of sigma x minus mean, that is (two minus four equal to minus two) plus (three minus four equal to minus one) plus (four minus four equal to zero) plus (five minus four equal to one) plus (six minus four equal to two), which is equal to zero divided by five is equal to zero.
If you note as mentioned earlier, we see that the first moment about the mean is equal to zero

Figure 10

(X)	$(\overline{X})$	$(X-\overline{X})$	$(X-\overline{X})^2$
2	4	2-4 = -2	$(-2)^2 = 4$
3	4	3-4 = -1	$(-1)^2 = 1$
4	4	4-4 = 0	$(0)^2 = 0$
5	4	5-4 = 1	$(1)^2 = 1$
6	4	6-4 = 2	$(2)^2 = 4$
N=5			$\Sigma(\overline{X}-X)^2 = 10$

Let us now calculate the second moment about the mean here mu two is equal to the square of total of Sigma x minus mean which is calculated by taking (two minus four square) plus (three minus four square) plus (four minus four square) plus (five minus four square) plus (six minus four square) divided by five which is equal to (minus two square) plus (minus one square) plus (zero) plus (one square) plus (two square) divided by five is equal to four plus one plus zero plus one plus four divided by five is equal to ten divided by five is equal to two. This indicates the variance of the data.

Figure 11

(X)	(X)	$(X-X)$	$(X-X)^3$
2	**4**	**2-4 = -2**	$(-2)^3 = -8$
3	**4**	**3-4 = -1**	$(-1)^3 = -1$
4	**4**	**4-4 = 0**	$(0)^3 = 0$
5	**4**	**5-4 = 1**	$(1)^3 = 1$
6	**4**	**6-4 = 2**	$(2)^3 = 8$
N=5			$\Sigma(X-X)^3 = 0$

The calculation of the third moment about the mean mu three is obtained by taking the (cube value of two minus four) plus (cube of three minus four) plus (cube of four minus four) plus (cube of five minus four) plus (cube of six minus four) divided by five is equal to (cube of minus two) plus (cube of minus one) plus (zero) plus (cube of one) plus (cube of two) divided by five is equal to (minus eight) plus (minus one) plus (zero) plus (one) plus (eight) divided by five is equal to zero. This indicates the skeweness of the data.

Figure 12

(X)	$(\overline{X})$	$(X-\overline{X})$	$(X-\overline{X})^4$
2	4	2-4 = -2	$(-2)^4 = 16$
3	4	3-4 = -1	$(-1)^4 = 1$
4	4	4-4 = 0	$(0)^4 = 0$
5	4	5-4 = 1	$(1)^4 = 1$
6	4	6-4 = 2	$(2)^4 = 16$
N=5			$\Sigma(X-X)^3 = 34$

The calculation of the fourth moment about the mean mu four is obtained by taking the (fourth root value of two minus four) plus (fourth root of three minus four) plus (fourth root of four minus four) plus (fourth root of five minus four) plus (fourth root of six minus four) divided by five is equal to (fourth root of minus two) plus (fourth root of minus one) plus (zero) plus (fourth root of one) plus (fourth root of two) divided by five is equal to sixteen plus one plus zero plus one plus sixteen which is equal to thirty four divided by five is equal to six point eight. This indicates the kurtosis of the data.

5. Problem's on Moments: Part-2

Problem 7:
From the following data calculate moments about the,
1. Assumed mean 25
2. Actual mean and
3. Moments about the zero

Figure 13

Variance	0-10	10-20	20-30	30-40
Frequency	1	3	4	2

In the table, first row represents the variance and second row represents the frequency.

Solution:

Figure 14

Variable	Mid value (m)	f	d	fd	fd^2	fd^3	fd^4
0-10	5	1	-2	-2	4	-8	16
10-20	15	3	-1	-3	3	-3	3
20-30	25	4	0	0	0	0	0
30-40	35	2	1	2	2	2	2
Total		N=10		-3	9	-9	21

In this table we will calculate 'd' which is obtained by the mid value minus the arbitrary mean 25, in the next column we will multiply the frequency with the 'd' value and obtain 'fd' which is the values for the first moment and its square, cube and fourth root in the subsequent columns.

First let us calculate the moments about the arbitrary mean.
mu one dash is equal to sigma of frequency and deviation that is minus three divided by the number of the data 'N' which is 10 into the width of the class interval 'c' that is 10 which is equal to minus three is the first moment.
Similarly the second moment is calculated by taking mu two dash is equal to sigma of frequency and square of the deviation which is 9 divided by the number of the data 'N' 10 into the square of the width of the class interval 'c' 10 which is equal to ninety.
The third moment is equal to minus 9 hundred.
The fourth moment is equal to 21 thousand.

Next let us calculate the moments about the actual mean.
As we know that, moment of the first order is equal to zero, therefore mu one is equal to zero. mu two is equal to mu two dash minus mu one dash square which is equal to ninety minus of minus three square is equal to eighty one.

mu three is equal to mu three dash minus three into mu one dash and mu two dash plus 2 into mu one dash cube is equal to minus 900 minus 3 into minus 3 into 90 plus two into minus three cube which is equal to minus one hundred forty four.
mu four is calculated as mu four dash minus four into mu one dash and mu three dash plus six into mu one dash square into mu two dash minus three into mu one dash to the power of four which is equal to 21 thousand minus ten thousand eight hundred plus four thousand eight sixty minus 243 is equal to fourteen thousand eight hundred seventeen.

Now let us calculate the moments about the zero.
The first moment about the zero is equal to arbitrary mean plus mu dash one is equal to 25 minus 3 is equal to 22.
The second moment about zero is equal to mu two plus square of the first moment about zero is equal to 81 plus 22 square is equal to 565.
The third moment about zero is equal to mu three plus thrice the moment of first and second about the zero minus twice the cube of first moment about zero is equal to 144 plus 3 into product of 22 & 565 minus 2 into 22 cube is equal to fifteen thousand eight hundred fifty.
The fourth moment about zero is equal to mu four plus four into first and third moment about zero minus 6 into square of first moment with the second moment about zero plus three into the fourth moment about zero to the power of four which is equal to fourteen thousand eight hundred seventeen plus thirteen lakh ninety four thousand eight hundred minus sixteen lakh forty thousand seven hundred sixty plus seven lakh two thousand seven hundred sixty eight which is equal to four lakh seventy one thousand six hundred twenty five.

Here’s a summary of our learning in this session, where we have understood the calculation of:

- Skewness and its various measures
- Moments and
- kurtosis