Iraqi Journal of Statistical Sciences

https://stats.uomosul.edu.iq/

Skewness Measures – Article Review

Sarya Mazen Anas[1] , Hyllaa Anas Abdulmjeed[2]

[1]Department of Mathematics, College of Education for Pure Science, University of Mosul, Mosul, Iraq

[2]Department of Statistics and Informatics, College of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq

Article information

Article history:

Received January 13,2025

Revised: April 20,2025

Accepted: June 13,2025

Available: December 1, 2025

Keywords:

Skewness Measures

Karl-Pearson measure Bowly's me

Kelly's measure,The Z-criterion

Correspondence:

Sarya Mazen Anas

Hyllaa Anas Abdulmjeed

sar.aljabal1995@gmail.com

hyllaa.77@uomosul.edu.iq

Abstract

The research aims to define skewness to know the type of skewness, by applying skewness measures such as the Karl-Pearson measure in its cases (median, median and moments) and Kelly's measure and Bowly's measure on real data for the variable (x, y), where the variable (x) represents the data for cement expansion taken from Badoush Cement Plant n=80 for the period (2008-2010) and the variable (y) represents data showing the times of failure resulting from various reasons for the life test device for n=30 devices .The Z-criterion was used to measure skewness for the data of variable X and for the data of variable Y. The results for both variables were outside the acceptable period for this criterion (1.96, -1.96).

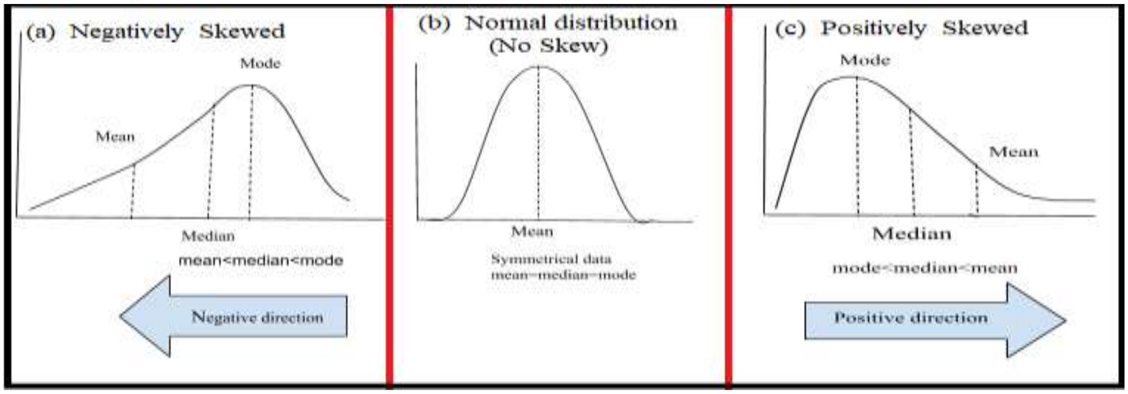
Doi:10.33899/iqjoss.v22i2.54085 ,

1. Introduction

Some data do not follow a symmetrical shape in their description. The data may include outliers values that extend the distribution from one end, which leads to the skewness of the curve, i.e. knowing the degree of data distribution around the center point in it. The nature of any distribution can be known by simply looking at the distribution curve. When representing the data of the phenomenon in the form of a frequency curve, skewness measures are used, which are used to know the skewness of the frequency distribution. The frequency distribution is positively skewed if it is skewed to the right and negatively skewed if the distribution is skewed to the left. (David, 2011)

2. Skewness

A frequency distribution of a set of data that are not symmetrical about the mean is called asymmetrical or skewed, or we can say that skewness is a departure from symmetry. In a skewed distribution, the outliers values in the data set move towards the upper or right tail, and the distribution is positively skewed. When these values move towards the lower or left tail, the distribution is negatively skewed. Symmetrical and asymmetrical curves are shown in the following graphs in figure (1) : (David, 2011) (ZEENALABIDEN, 2023) (Adebayo, 2024) (,2024)



Figure(1) : Symmetrical and asymmetrical curves are shown in graphs

In figure (b) we see that the curve is symmetrical about the mean, i.e. there is no deviation. In figures (a) and (c) we see that the curves are not symmetrical about the mean, i.e. they are skewed. In figure (a) the deviation is negative, while in figure (c) the deviation is positive. For a positively skewed distribution Mean >Median >Mode , and for a negatively skewed distribution Mean<Median<Mode .

3-Skewness measures

The most important skewness measures are:

3-1) Karl – Pearson

This method is often used to measure skewness. The formula for measuring the skewness coefficient is given: **a)** Using the mode, the first case is represented: (Al-Rawi, 1980) (Mahmoud, 2020) (Mukhlif,2024)

$$S_k = \frac{\text{Mean} - \text{Mode}}{\text{Standard deviation}} = \frac{\bar{x} - \text{Mo}}{s} \qquad (1)$$

Where S_k: represents the skewness coefficient, $\bar{x}$: represents the mean , Mo: represents the mode , s : represents the standard deviation.

b)Using the mediator, the second case is represented :

$\text{Mean} - \text{Mode} = 3(\text{ Mean} - \text{Median})$ or

$$\text{Mode} = 3\text{Median} - 2\text{Mean} \qquad (2)$$

When equation (2) is substituted into the first equation, we get :

$$S_k = \frac{3(\bar{x} - \text{Med})}{s} \qquad (3)$$

Where Med: represents the median.

c) Using the moments (the second moment and the third moment), the third case is represented as in the following formula :

$$S_k = \sqrt{B_1} = \sqrt{\frac{\mu_3^2}{\mu_2^3}} = \frac{\mu_3}{\sqrt{\mu_2^3}} \qquad (4)$$

Where μ_2: represents the second moment (variance), μ_3: represents the third moment.

Properties of Karl Pearson's skewness coefficient:

* $-1 \le S_k \le 1$
* $S_k = 0$ Symmetric distribution about the mean
* $S_k > 0$ Skewed to the right .
* $S_k < 0$ Skewed to the left .

3-2) Bowley's Coefficient

This method depends on the quartiles. It is preferable to use this measure when the data contains outliers or extreme values because this measure is not sensitive to these values. The formula for calculating the skewness coefficient is given:

$$S_k = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{(Q_3 - 2Q_2 + Q_1)}{Q_3 - Q_1} \quad (5)$$

Where Q_1, Q_2 and Q_3: represent the first, second and third quartiles, respectively, after arranging the data (ascending or descending) and according to the following formula Qj is the value whose order is $= \frac{j.n}{4}$, j=1,2,3 and n: represents the number of data. (Chandima, 2019) (Mahmoud, 2020) (Al-Safawi, 2008)

3-3) Kelly's Coefficient

The relative measures of skewness proposed by professor Kelly are based on percentages and decimals, and are considered more sensitive to outliers than the Bowley measure. Therefore, if the data contains outliers or extreme values, the value resulting from this measure may be more affected by these values. (Bhadouria, 2024)

$$S_k = \frac{D_1 + D_9 - 2D_5}{D_9 - D_1} \quad (6)$$

Where D_1, D_5, D_9: represent the first, fifth and ninth decimals respectively, after arranging the data (ascending or descending) and according to the following formula: D_k is the value whose order is =((k/10)*(1+n)), k= 1, 5, 9 and n represents the number of data

4- The Z-Score

Using the skewness and standard error of the skewness to find the Z-score:

$$Z_{Sk} = \frac{S_k - \bar{x}}{S_{ke}} \quad (7)$$

Where S_k: represents the skewness coefficient , $\bar{x}$: represents the mean , S_{ke}: represents standard diviation error of the skewness . Use Z-score to test the approximation of the sample to the normal distribution. This value should be between -1.96 and +1.96 for the significance level of 0.05. If this Z-score does not fall within this range, indicates that the sample is outside the normal state of skewness, i.e. it does not fall within the acceptable interval. Therefore, the sample should be adjusted and re-distributed or we should use a non-parametric statistical test. (2009, ANKARALI) (Ghasemi, 2012)

5- Applying the Skewness Measures to the Data as Follows

A- The variable (x) represents the data for cement expansion taken from Badoush Cement Plant n=80 for the period (2008-2010).

Draw the frequency distribution curve for the variable (x) as in figure (2) :

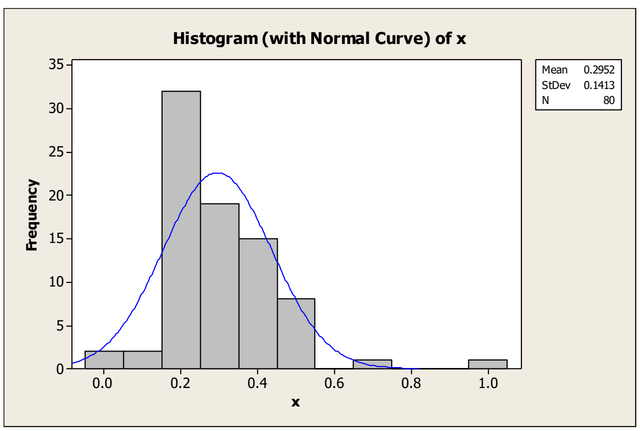


Figure (2): Represents the frequency curve of the variable x

We notice that the frequency distribution curve is skewed to the right. This indicates that (mean > median). To find the skewness measures for the variable x for the cement expansion data, as follows :

3-1) Karl-Pearson

a) In the case of the mode from equation (1) , we first find the value of the mean , mode , and standard deviation $\bar{x} = 0.29525$, $M_\circ = 0.23$, $S = 0.14127$, then

$$S_k = \frac{0.29525 - 0.23}{0.14127} = 0.4619$$

We notice that the value of the skewness scale is > 0, which gives a skew to the right .

b) In the case of the presence of the median from equation (3), we need to find the value of the mean, the median, and the standard deviation of the data. It was $\bar{x} = 0.29525$, Med=0.26 , S=0.14127, then

$$S_k = \frac{3(\,0.29525 - 0.26)}{0.14127} = 0.7486$$

Also the result of the scale > 0, so the frequency distribution curve is skewed to the right.

c) In the case of the moments in equation (4), we find the values of the second moment and the third moment ($M_3 = 0.0043$, $M_2 = 0.0\ \ 2$), then we apply equation (4) as follows :

$$B_1 = \frac{(0.0043)^2}{(0.02)^3} = 2.3463$$

$$S_K = \sqrt{B_1} = 1.5318$$

Since the result of the scale > 0, then the skewness of this data is on the right side.We note that the value of this scale is > 1. This indicates that the expansion data in the case of moments is outside the normal state of skewness because the slight skewness is -1 $\leq S_k \leq$1, while the large skewness (i.e. outside the normal state of skewness) is less than (-1) or greater than (1) .

3-2) Bowley's Coefficient

The value of the first, second and third quartiles must be found from the following formula Qj=(j.n)/4 , to find the sequence of the quartiles, then determine the value of the quartile from the data (after arranging them in ascending order) as follows:

($Q_1 = \frac{1*80}{4} = 20$, $Q_2 = \frac{2*80}{4} = 40$, $Q_3 = \frac{3*80}{4} = 60$) ,the values of quartiles ($Q_1 = 0.20$, $Q_2 = 0.26$, $Q_3 = 0.38$) then apply formula (5) as follows:

$$S_K = \frac{0.38 + 0.20 - 2(0.26)}{0.38 - 0.20} = 0.333$$

Also the data is skewed to the right .

3-3) Kelly's Coefficient

The value of the first, fifth and ninth decimals must be found from the following formula Dk=((k/10)*(1+n)) to find the sequence of decimals, then determine the value of the decimal from the data (after arranging it in ascending order) as follows:

$$D_1 = ((1/10) * (1 + 80) = 8.1 \cong 9$$
$$D_5 = ((5/10) * (1 + 80) = 40.5 \cong 41$$
$$D_9 = ((9/10) * (1 + 80)) = 72.9 \cong 73$$

The values of the decimals($D_1 = 0.18$, $D_5 = 0.26$, $D_9 = 0.47$) , we substitute D_1, D_5, D_9, in formula (6) as follows:

$$S_K = \frac{0.47 + 0.18 - 2(0.26)}{0.47 - 0.18} = 0.4483$$

Also the coefficient value is > 0

After finding the value of all the skewness measures, the value of the Z-score was taken for each measure according to formula (7), where as the value of the standard deviation error of the skewness Ske = 0.0158 obtained from the minitab program, the results were as follows in table (1):

Table (1): Shows the value of the Z criterion for the skewness measures

Metrics	Karl-Pearson			Bowley	Kelly
	in case the Mode	in case the Median	in case the Moment		
Z-score	10.5465	28.6913	78.2627	2.4103	9.6852

The table above shows that the Z-score value for both the Pearson skewness scale in all cases and the Bowley's and Kelly skewness scale does not fall within the acceptable range between (-1.96 and .961), meaning that all Z-score values for the scales fall outside the acceptable range and are greater than 1.96. This indicates that the frequency distribution curve is highly skewed outside the normal state.

B-The variable (y) represents data showing the failure times resulting from various reasons for the life test device for (n=30 devices).

To draw the frequency curve for the variable Y as in figure (3)

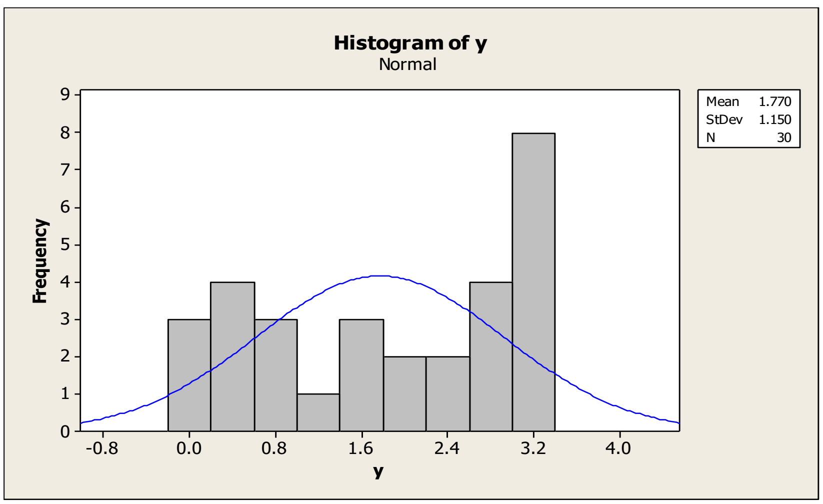


Figure (3): Represents the frequency curve of the variable y.

We notice that the frequency distribution curve is skewed to the left. This indicates that (mean < median). To find the measures of skewness for the variable y for the data showing the failure times of the life test device as follows:

1- Karl-Pearson:

* In the case of the mode from equation (1), we first find the value of the arithmetic mean, mode, and standard deviation, which were $\bar{x}$=1.7703, S=1.1499, $M_\circ$=3, then

$$S_k = \frac{1.7703-3}{1.1499} = -0.5079$$

We notice that the value of the skewness measure < 0, which gives a skew on the left side.

* In the case of the median from equation (3), we need to find the value of the arithmetic mean , median , and standard deviation of the data. It was $\bar{x}$ = 1.7703, Med = 1.965, S = 1.1499, then

$$S_k = \frac{3(1.7703-1.965)}{1.1499} = -1.0694$$

Also the scale result < 0 so the frequency distribution curve is skewed to the left to a large extent this indicates that the skewness is outside the normal state .

*In the case of the moments in equation (4) we find the values of the second moment and the third moment (M_3=-0.409219, M_2=1.3223), then we apply equation (4) as follows :

$$B_1 = \frac{(-0.409219)^2}{(1.3223)^3} = \quad 0.0724$$

$$S_K = \sqrt{B_1} = 0.2691$$

The value of the skewness measure for this case is > 0

2- Bowley's Coefficient

The value of the first, second and third quartiles must be found from the following formula Qj=(j.n)/4 , to find the sequence of the quartiles, then determine the value of the quartile from the data (after arranging them in ascending order) as follows:

$$(Q_1 = \frac{1*30}{4} = 7.5 \cong 8 \text{ , } Q_2 = \frac{2*30}{4} = 15 \text{ , } Q_3 = \frac{3*30}{4} = 22.5 \cong 23 \quad)$$

,the values of quartiles ($Q_1 = 0.65$, $Q_2 = 1.81$, $Q_3 = 3$) then apply formula (5) as follows:

$$S_K = \frac{3 + 0.65 - 2(1.81)}{3 - 0.65} = 0.0127$$

Also the skewness value for this measure is > 0

3- Kelly's Coefficient

The value of the first, fifth and ninth decimals must be found from the following formula Dk=((k/10)*(1+n)) to find the sequence of decimals, then determine the value of the decimal from the data (after arranging it in ascending order) as follows:

$$D_1 = ((1/10) * (1 + 30) = 3.1 \cong 3$$

$$D_5 = ((5/10) * (1 + 30) = 15.5 \cong 16$$

$$D_9 = ((9/10) * (1 + 30)) = 27.9 \cong 28$$

The values of the decimals ($D_1 = 0.13$, $D_5 = 2.12$, $D_9 = 3$) , we substitute D_1, D_5, D_9, in formula (6) as follows:

$$S_K = \frac{3 + 0.13 - 2(2.12)}{3 - 0.13} = 0.386$$

Also the value of the scale is > 0

After finding the value of all the skewness scales, the value of the Z-score was taken for each scale according to formula (7),), where as the value of the standard deviation error of the skewness Ske = 0.21 obtained from the minitab program, the results were as follows in table (2):

Table (2): Shows the value of the Z criterion for the skewness measures

Metrics	Karl-Pearson			Bowley	Kelly
	in case the Moment	in case the Moment	in case the Moment		
Z-score	-10.8486	-13.5224	-7.1486	-8.3695	-10.2680

The table above shows that the Z-score value for both Karl Pearson's skewness measure in all cases and Bowley's and Kelly's skewness measure does not fall within the acceptable range between (-1.96 and .961), meaning that all Z-score values for the measures fall outside the acceptable range and are significantly less than -1.96. This indicates that the frequency distribution curve for this data is skewed significantly to the left, i.e. outside the normal state.

4. Conclusions

1- The values of the Karl Pearson skewness scales (in the case of mode and median), Bowley and Kelly are greater than zero and show that the skewness is on the right side, which is a slight skewness within the specified period of skewness, while the value of the Karl Pearson scale in the case of moments is greater than (1), which is considered a large skewness outside the normal state of skewness

2- The Z-score value for all skewness scales does not fall within the acceptable interval for the Z-score between(1.96 and -1.96)

3- The value of the Karl Pearson skewness scale in the case of the median is less than (-1) and is not within the specified period for the values of the skewness scale, which indicates a large skewness to the left, i.e. outside the normal state of skewness.

4- The values of the Karl Pearson skewness scale in the case of moments and the Bowley scale are greater than zero, while the data has a negative skewness to the left

5- The Z-score value for all skewness scales does not fall within the acceptable interval for the Z-score between (1.96 and 1.96) .

References

1. Adebayo. O. P, Ogunjimi . O. A, Ahmed. I ,(2024), "Application of MANOVA and Hotelling's T square on Academic Performance of University Students Based on Mode of Entry " , Iraqi Journal of statistical Sciences, Vol. 21, No. 2,2024, pp (1-8). DOI: 10.33899/iqjoss.2024.185231
2. Ahmad H. A. 1and Saieed H. A. (2024) , " Bayesian Estimation of TLR distribution under Sq and Elinex loss" , Iraqi Journal of Statistical Sciences, Vol. 21, No. 2,2024, pp (62-69) ,DOI: 10.33899/iqjoss.2024.185241
3. Al-Rawi, Dr. Khasha Mahmoud (1980),"Introduction to Statistics" University of Mosul, College of Agriculture.
4. Al-Safawi, Asst. Prof. Dr. Safaa Younis (2008), "Statistics", University of Mosul, College of Computers and Mathematics, Department of Statistics and Informatics
5. Al-Sarraf,Israa Najib Saeed(2021),"Robust Methods for Principal Component Regression:A Comparative Study with Simulation", Master's Thesis, Unpublished,Department of Statistics and Informatics, University of Mosul.
6. ALTINAY, Galip(2016)," A Simple Class of Measures of Skewness", Bandirma Onyedi Eylul University, Turkey, https://mpra.ub.uni-muenchen.de/72353/
7. ANKARALI , H., YAZICI ,A. C. and ANKARALI ,S. (2009)," A Bootstrap Confidence Interval for Skewness and Kurtosis and Properties of t-test in Small Samples from Normal Distribution", Trakya Univ Tip Fak Derg 2009;26(4):297-305. https://www.researchgate.net/publication/289621506_A_Bootstrap_Confidence_Interval_for_Skewness_and_Kurtosis_and_Properties_of_t-test_in_Small_Samples_from_Normal_Distributio
8. Bhadouria, S.S.(2024)" Business Statistics", Indira Gandhi National Tribal University , https://www.igntu.ac.in.
9. Chandima N.P.G. Arachchige_,Luke A. Prendergasty (2019)," Mean skewness measures",Department of Mathematics and Statistics, La Trobe University Melbourne, Victoria, Australia, 3086 , DOI : arXiv:1912.06996v2
10. David P. Doane and David P. Doane (2011) ,"Measuring Skewness: A Forgotten Statistic? " , Journal of Statistics Education Volume 19, Number 2(2011) , DOI:10.48550/arXiv.1912.06996
11. Ghasemi A., Zahediasl S.(2012)," Normality Tests for Statistical Analysis : A Guide for Non-Statisticians " , Int J Endocrinol Metab. 2012;10(2):486-489. DOI: 10.5812/ijem.3505
12. Mahmoud.A.Eltehiwy,1 Abu-Bakr A. Abdul-Motaal,(2020) ," A New coefficient of Skewness for grouped data", Biometrics & Biostatistics International Journal, Volume 9 Issue 2 – 2020 , 1Faculty of Politics and Economics, Department of Statistics, Beni Suef University, Egypt . DOI: 10.15406/bbij.2020.09.00300
13. Mukhlif .M.H.&,Saieed .H.A., (2024),"Study Of Alpha Power Weibull Distribution With Application", Iraqi Journal of statistical Sciences,Vol.21,No.1,2024,Pp(112-126) . DOI:10.33899/iqjoss.2024.183252
14. ZEENALABIDEN,N.A.and SARAÇOĞLU,B.,(2023),"Exponential Power – Chen Distribution and Its Some roperties" , Iraqi Journal of statistical Sciences,Vol.20,No.2,2023,Pp(134-4).DOI:10.33899/iqjoss.2023.181219
15. ZEENALABIDEN , N.A.and SARAÇOĞLU ,B., (2023) ," A New Family Of Distributions: Exponential Power-X Family Of Distributions And Its Some Properties" , Iraqi Journal of statistical Sciences,Vol.20,No.2,2023,Pp(235-284) . DOI: 10.33899/IQJOSS.2023.0181258

مقاييس الالتواء – مراجعة مقال

سارية مازن أنس[1] ، هيلاء أنس عبدالمجيد[2]

[1] قسم الرياضيات , كلية التربية للعلوم الصرفة , جامعة الموصل , العراق ،[2] قسم الاحصاء والمعلوماتية , كلية علوم الحاسوب والرياضيات , جامعة الموصل , العراق

الخلاصة: يهدف البحث الى تعريف الالتواء ومعرفة نوع الالتواء للبيانات , من تطبيق مقاييس الالتواء مثل مقياس Karl–Pearson في حالاته (الوسيط والوسيط والعزوم) ومقياس Kelly's ومقياس Bowly's على بيانات حقيقية للمتغير (x,y) حيث ان المتغير (x) يمثل البيانات لتمدد الاسمنت الماخوذة من معمل سمنت بادوش n=80 للفترة (2008–2010) (الصراف , 2021) والمتغير (y) يمثل بيانات تظهر اوقات العطل الناتجة عن اسباب متنوعة لجهاز اختبار الحياة لـ n=30 جهاز (ZEENALABIDEN, 2023) . واستخدام المعيار (Z) لمقاييس الالتواء لبيانات المتغير X و لبيانات المتغير Y فكانت النتائج لكلا المتغيرين هي خارج الفترة المقبولة لهذا المعيار (1.96- ,1.96) .

الكلمات المفتاحية : مقاييس الالتواء , مقياس Karl–Pearson, مقياس Kelly's , ومقياس Bowly's , المعيار (Z).