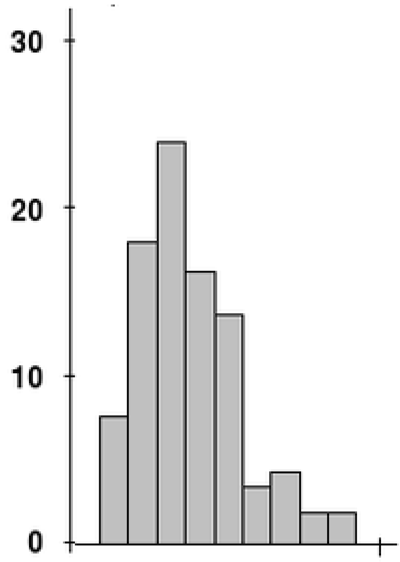
WIKIPEDIA

Skewness

In probability theory and statistics, **skewness** is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive or negative, or undefined.



Example distribution with non-zero (positive) skewness. These data are from experiments on wheat grass growth.

For a unimodal distribution, negative skew commonly indicates that the *tail* is on the left side of the distribution, and positive skew indicates that the tail is on the right. In cases where one tail is long but the other tail is fat, skewness does not obey a simple rule. For example, a zero value means that the tails on both sides of the mean balance out overall; this is the case for a symmetric distribution, but can also be true for an asymmetric distribution where one tail is long and thin, and the other is short but fat.

Some popular intuitions about skewness are not correct. As a 2005 journal article points out[1]:

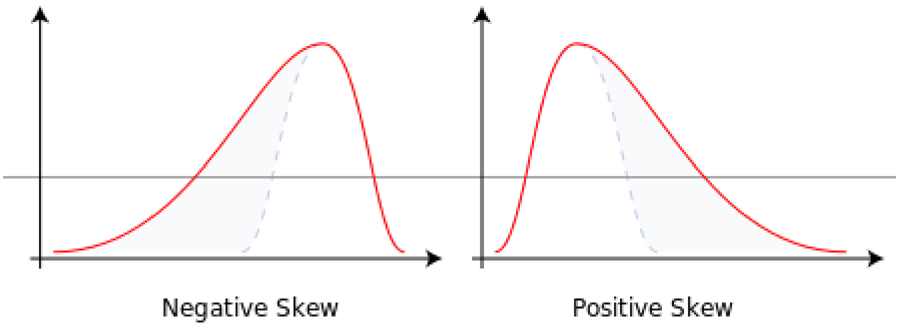
> Many textbooks teach a rule of thumb stating that the mean is right of the median under right skew, and left of the median under left skew. This rule fails with surprising frequency. It can fail in multimodal distributions, or in distributions where one tail is long but the other is heavy. Most commonly, though, the rule fails in discrete distributions where the areas to the left and right of the median are not equal.

Contents

Introduction

Consider the two distributions in the figure just below. Within each graph, the values on the right side of the distribution taper differently from the values on the left side. These tapering sides are called *tails*, and they provide a visual means to determine which of the two kinds of skewness a distribution has:

1. *negative skew*: The left tail is longer; the mass of the distribution is concentrated on the right of the figure. The distribution is said to be *left-skewed*, *left-tailed*, or *skewed to the left*, despite the fact that the curve itself appears to be skewed or leaning to the right; *left* instead refers to the left tail being drawn out and, often, the mean being skewed to the left of a typical center of the data. A left-skewed distribution usually appears as a *right-leaning* curve.[2]
2. *positive skew*: The right tail is longer; the mass of the distribution is concentrated on the left of the figure. The distribution is said to be *right-skewed*, *right-tailed*, or *skewed to the right*, *despite* the fact that the curve itself appears to be skewed or leaning to the left; *right* instead refers to the right tail being drawn out and, often, the mean being skewed to the right of a typical center of the data. A right-skewed distribution usually appears as a *left-leaning* curve.[2]



Skewness in a data series may sometimes be observed not only graphically but by simple inspection of the values. For instance, consider the numeric sequence (49, 50, 51), whose values are evenly distributed around a central value of 50. We can transform this sequence into a negatively skewed distribution by adding a value far below the mean, e.g. (40, 49, 50, 51). Similarly, we can make the sequence positively skewed by adding a value far above the mean, e.g. (49, 50, 51, 60).

Relationship of mean and median

The skewness is not directly related to the relationship between the mean and median: a distribution with negative skew can have its mean greater than or less than the median, and likewise for positive skew.[3]

In the older notion of nonparametric skew, defined as $(\mu - \nu)/\sigma$, where μ is the mean, ν is the median, and σ is the standard deviation, the skewness is defined in terms of this relationship: positive/right nonparametric skew means the mean is greater than (to the right of) the median, while negative/left nonparametric skew means the mean is less than (to the left of) the median. However, the modern definition of skewness and the traditional nonparametric definition do not in general have the same sign: while they agree for some families of distributions, they differ in general, and conflating them is misleading.

If the distribution is symmetric, then the mean is equal to the median, and the distribution has zero skewness.[4] If the distribution is both symmetric and unimodal, then the mean = median = mode. This is the case of a coin toss or the series 1,2,3,4,... Note, however, that the converse is not true in general, i.e. zero skewness does not imply that the mean is equal to the median.

A 2005 journal article points out[5]:

> Many textbooks, teach a rule of thumb stating that the mean is right of the median under right skew, and left of the median under left skew. This rule fails with surprising frequency. It can fail in multimodal distributions, or in distributions where one tail is long but the other is heavy. Most commonly, though, the rule fails in discrete distributions where the areas to the left and right of the median are not equal. Such distributions not only contradict the textbook relationship between mean, median, and skew, they also contradict the textbook interpretation of the median.

Definition

Pearson's moment coefficient of skewness

The skewness of a random variable X is the third standardized moment γ_1, defined as:[6][7]

$$\gamma_1 = \operatorname{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{\operatorname{E}\left[(X-\mu)^3\right]}{\left(\operatorname{E}\left[(X-\mu)^2\right]\right)^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}}$$

where μ is the mean, σ is the standard deviation, E is the expectation operator, μ_3 is the third central moment, and κ_t are the t$^{\text{th}}$ cumulants. It is sometimes referred to as **Pearson's moment coefficient of skewness**,[7] or simply the **moment coefficient of skewness**,[6] but should not be confused with Pearson's other skewness statistics (see below). The last equality expresses skewness in terms of the ratio of the third cumulant κ_3 to the 1.5th power of the second cumulant κ_2. This is analogous to the definition of kurtosis as the fourth cumulant normalized by the square of the second cumulant. The skewness is also sometimes denoted Skew[X].

Skewness can be expressed in terms of the non-central moment E[X^3] by expanding the previous formula,

$$\begin{aligned}
\gamma_1 &= \operatorname{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] \\
&= \frac{\operatorname{E}[X^3] - 3\mu\operatorname{E}[X^2] + 3\mu^2\operatorname{E}[X] - \mu^3}{\sigma^3} \\
&= \frac{\operatorname{E}[X^3] - 3\mu(\operatorname{E}[X^2] - \mu\operatorname{E}[X]) - \mu^3}{\sigma^3} \\
&= \frac{\operatorname{E}[X^3] - 3\mu\sigma^2 - \mu^3}{\sigma^3}.
\end{aligned}$$

Examples

Skewness can be infinite, as when

$$\Pr\left[X > x\right] = x^{-2} \text{ for } x > 1,\ \Pr[X < 1] = 0$$

where the third cumulants are infinite, or as when

$$\Pr[X < x] = (1-x)^{-3}/2 \text{ for negative } x \text{ and } \Pr[X > x] = (1+x)^{-3}/2 \text{ for positive } x.$$

where the third cumulant is undefined.

Properties

Starting from a standard cumulant expansion around a normal distribution, one can show that

$$\text{skewness} = \frac{6\,(\text{mean} - \text{median})}{\text{standard deviation}\,(1 + \frac{\text{kurtosis}}{8})} + O\,(\text{skewness}^2).$$

If Y is the sum of n independent and identically distributed random variables, all with the distribution of X, then the third cumulant of Y is n times that of X and the second cumulant of Y is n times that of X, so $\operatorname{Skew}[Y] = \operatorname{Skew}[X]/\sqrt{n}$. This shows that the skewness of the sum is smaller, as it approaches a Gaussian distribution in accordance with the central limit theorem. Note that the assumption that the variables be independent for the above formula is very important because it is possible even for the sum of two Gaussian variables to have a skewed distribution (see this example).

Sample skewness

For a sample of n values, a natural method of moments estimator of the population skewness is[8]

$$b_1 = \frac{m_3}{s^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}^{\,3}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2\right]^{3/2}},$$

where $\overline{x}$ is the sample mean, s is the sample standard deviation, and the numerator m_3 is the sample third central moment.

Another common definition of the *sample skewness* is[8][9]

$$\begin{aligned} G_1 &= \frac{k_3}{k_2^{3/2}} = \frac{n^2}{(n-1)(n-2)}\,\frac{m_3}{s^3} \\ &= \frac{\sqrt{n\,(n-1)}}{n-2}\,\frac{m_3}{m_2^{\frac{3}{2}}} = \frac{\sqrt{n\,(n-1)}}{n-2}\left[\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{\frac{3}{2}}}\right], \end{aligned}$$

where k_3 is the unique symmetric unbiased estimator of the third cumulant and $k_2 = s^2$ is the symmetric unbiased estimator of the second cumulant (i.e. the variance).

In general, the ratios b_1 and G_1 are both biased estimators of the population skewness γ_1; their expected values can even have the opposite sign from the true skewness. (For instance, a mixed distribution consisting of very thin Gaussians centred at −99, 0.5, and 2 with weights 0.01, 0.66, and 0.33 has a skewness of about −9.77, but in a sample of 3, G_1 has an expected value of about 0.32, since usually all three samples are in the positive-valued part of the distribution, which is skewed the other way.) Nevertheless, b_1 and G_1 each have obviously the correct expected value of zero for any symmetric distribution with a finite third moment, including a normal distribution.

Under the assumption that the underlying random variable X is normally distributed, it can be shown that $\sqrt{n}b_1 \xrightarrow{d} N(0,6)$, i.e., its distribution converges to a Normal distribution with mean 0 and variance 6. The variance of the skewness of a random sample of size n from a normal distribution is[10][11]

$$\operatorname{var}(G_1) = \frac{6n(n-1)}{(n-2)(n+1)(n+3)}.$$

An approximate alternative is $6/n$, but this is inaccurate for small samples.

In normal samples, b_1 has the smaller variance of the two estimators, with

$$\operatorname{var}(b_1) < \operatorname{var}\left(\frac{m_3}{m_2^{3/2}}\right) < \operatorname{var}(G_1),$$

where m_2 in the denominator is the (biased) sample second central moment.[8]

The adjusted Fisher–Pearson standardized moment coefficient G_1 is the version found in Excel and several statistical packages including Minitab, SAS and SPSS.[12]

Applications

Skewness is a descriptive statistic that can be used on conjunction with the histogram and the normal quantile plot to characterize the data or distribution.

Skewness indicates which direction and a relative magnitude of how far a distribution deviates from normal.

With pronounced skewness, standard statistical inference procedures such as a confidence interval for a mean will be not only incorrect, in the sense of having true coverage level unequal to the nominal (e.g., 95%) level, but also with unequal error probabilities on each side.

Skewness can be used to obtain approximate probabilities and quantiles of distributions (such as the value at risk in finance) via the Cornish-Fisher expansion.

Many models assume normal distribution; i.e., data are symmetric about the mean. The normal distribution has a skewness of zero. But in reality, data points may not be perfectly symmetric. So, an understanding of the skewness of the dataset indicates whether deviations from the mean are going to be positive or negative.

D'Agostino's K-squared test is a goodness-of-fit normality test based on sample skewness and sample kurtosis.

Other measures of skewness

Other measures of skewness have been used, including simpler calculations suggested by Karl Pearson[13] (not to be confused with Pearson's moment coefficient of skewness, see above). These other measures are:

Pearson's first skewness coefficient (mode skewness)

The Pearson mode skewness,[14] or first skewness coefficient, is defined as

$$\frac{\text{mean} - \text{mode}}{\text{standard deviation}}.$$

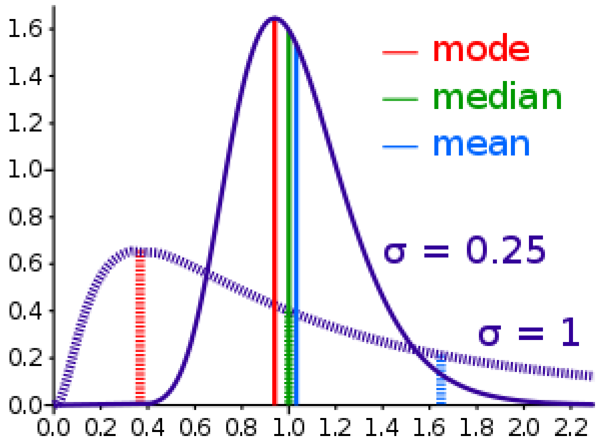
Pearson's second skewness coefficient (median skewness)

The Pearson median skewness, or second skewness coefficient,[15][16] is defined as

$$\frac{3\,(\text{mean} - \text{median})}{\text{standard deviation}}.$$

Which is a simple multiple of the nonparametric skew.

Quartile-based measures



Comparison of mean, median and mode of two log-normal distributions with different skewnesses.

Bowley's measure of skewness (from 1901),[17][18] also called Yule's coefficient (from 1912)[19][20] is defined as:

$$B_1 = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}.$$

When writing it as $\frac{\frac{Q_3+Q_1}{2} - Q_2}{\frac{Q_3-Q_1}{2}}$, it is easier to see that the numerator is the average of the upper and lower quartiles (a measure of location) minus the median while the denominator is (Q3-Q1)/2 which (for symmetric distributions) is the MAD measure of dispersion.

Other names for this measure are Galton's measure of skewness,[21] the Yule–Kendall index[22] and the quartile skewness,

A more general formulation of a skewness function was described by Groeneveld, R. A. and Meeden, G. (1984):[23][24][25]

$$\gamma(u) = \frac{F^{-1}(u) + F^{-1}(1-u) - 2F^{-1}(1/2)}{F^{-1}(u) - F^{-1}(1-u)}$$

where F is the cumulative distribution function. This leads to a corresponding overall measure of skewness[24] defined as the supremum of this over the range $1/2 \leq u < 1$. Another measure can be obtained by integrating the numerator and denominator of this expression.[23] The function $\gamma(u)$ satisfies $-1 \leq \gamma(u) \leq 1$ and is well defined without requiring the existence of any moments of the distribution.[23] Quantile-based skewness measures are at first glance easy to interpret, but they often show significantly larger sample variations, than moment-based methods. This means that often samples from a symmetric distribution (like the uniform distribution) have a large quantile-based skewness, just by chance.

Bowley's measure of skewness is $\gamma(u)$ evaluated at $u = 3/4$. Kelley's measure of skewness uses $u = 0.1$.[26]

Groeneveld & Meeden's coefficient

Groeneveld & Meeden have suggested, as an alternative measure of skewness,[23]

$$B_3 = \operatorname{skew}(X) = \frac{(\mu - \nu)}{E(|X - \nu|)},$$

where μ is the mean, ν is the median, $|...|$ is the absolute value, and $E()$ is the expectation operator. This is closely related in form to Pearson's second skewness coefficient.

L-moments

Use of L-moments in place of moments provides a measure of skewness known as the L-skewness.[27]

Distance skewness

A value of skewness equal to zero does not imply that the probability distribution is symmetric. Thus there is a need for another measure of asymmetry that has this property: such a measure was introduced in 2000.[28] It is called **distance skewness** and denoted by dSkew. If *X* is a random variable taking values in the *d*-dimensional Euclidean space, *X* has finite expectation, *X*' is an independent identically distributed copy of *X*, and $\|\cdot\|$ denotes the norm in the Euclidean space, then a simple *measure of asymmetry* with respect to location parameter θ is

$$\mathrm{dSkew}(X) := 1 - \frac{\mathrm{E}\,\|X - X'\|}{\mathrm{E}\,\|X + X' - 2\theta\|} \text{ if } \Pr(X = \theta) \neq 1$$

and dSkew(*X*) := 0 for *X* = θ (with probability 1). Distance skewness is always between 0 and 1, equals 0 if and only if *X* is diagonally symmetric with respect to θ (*X* and θ −*X* have the same probability distribution) and equals 1 if and only if X is a constant *c* ($c \neq \theta$) with probability one.[29] Thus there is a simple consistent statistical test of diagonal symmetry based on the **sample distance skewness**:

$$\mathrm{dSkew}_n(X) := 1 - \frac{\sum_{i,j} \|x_i - x_j\|}{\sum_{i,j} \|x_i + x_j - 2\theta\|}.$$

Medcouple

The medcouple is a scale-invariant robust measure of skewness, with a breakdown point of 25%.[30] It is the median of the values of the kernel function

$$h(x_i, x_j) = \frac{(x_i - x_m) - (x_m - x_j)}{x_i - x_j}$$

taken over all couples (x_i, x_j) such that $x_i \geq x_m \geq x_j$, where x_m is the median of the sample $\{x_1, x_2, \ldots, x_n\}$. It can be seen as the median of all possible quantile skewness measures.

See also

- Bragg peak
- Coskewness
- Shape parameters
- Skew normal distribution
- Skewness risk

Notes

1. von Hippel, Paul (2005). "Mean, Median, and Skew: Correcting a Textbook Rule" (https://ww2.amstat.org/publications/jse/v13n2/vonhippel.html). *Journal of Statistics Education*. **13** (2).
2. Susan Dean, Barbara Illowsky "Descriptive Statistics: Skewness and the Mean, Median, and Mode" (http://cnx.org/content/m17104/latest/), Connexions website
3. von Hippel, Paul T. (2005). "Mean, Median, and Skew: Correcting a Textbook Rule" (http://www.amstat.org/publications/jse/v13n2/vonhippel.html). *Journal of Statistics Education*. **13** (2).
4. "1.3.5.11. Measures of Skewness and Kurtosis" (http://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm). NIST. Retrieved 18 March 2012.
5. von Hippel, Paul T. (2005). "Mean, Median, and Skew: Correcting a Textbook Rule" (http://www.amstat.org/publications/jse/v13n2/vonhippel.html). *Journal of Statistics Education*. **13** (2).

6. "Measures of Shape: Skewness and Kurtosis" (http://brownmath.com/stat/shape.htm), 2008–2016 by Stan Brown, Oak Road Systems
7. Pearson's moment coefficient of skewness (http://www.fxsolver.com/browse/formulas/Pearson's+moment+coefficient+of+skewness), FXSolver.com
8. Joanes, D. N.; Gill, C. A. (1998). "Comparing measures of sample skewness and kurtosis". *Journal of the Royal Statistical Society, Series D*. **47** (1): 183–189. doi:10.1111/1467-9884.00122 (https://doi.org/10.1111%2F1467-9884.00122).
9. Doane, David P., and Lori E. Seward. "Measuring skewness: a forgotten statistic." (http://ww2.amstat.org/publications/jse/v19n2/doane.pdf) Journal of Statistics Education 19.2 (2011): 1-18. (Page 7)
10. Duncan Cramer (1997) Fundamental Statistics for Social Research. Routledge. ISBN 9780415172042 (p 85)
11. Kendall, M.G.; Stuart, A. (1969) *The Advanced Theory of Statistics, Volume 1: Distribution Theory, 3rd Edition*, Griffin. ISBN 0-85264-141-9 (Ex 12.9)
12. Doane DP, Seward LE (2011) J Stat Educ 19 (2)
13. "Archived copy" (https://web.archive.org/web/20100705025706/http://www.stat.upd.edu.ph/s114%20cnotes%20fcapistrano/Chapter%2010.pdf) (PDF). Archived from the original (http://www.stat.upd.edu.ph/s114%20cnotes%20fcapistrano/Chapter%2010.pdf) (PDF) on 5 July 2010. Retrieved 9 April 2010.
14. Weisstein, Eric W. "Pearson Mode Skewness" (http://mathworld.wolfram.com/PearsonModeSkewness.html). *MathWorld*.
15. Weisstein, Eric W. "Pearson's skewness coefficients" (http://mathworld.wolfram.com/PearsonsSkewnessCoefficients.html). *MathWorld*.
16. Doane, David P., and Lori E. Seward. "Measuring Skewness: A Forgotten Statistic? (http://www.amstat.org/publications/jse/v19n2/doane.pdf)" Journal of Statistics Education 19.2 (2011): 1-18.
17. Bowley, A. L. (1901). Elements of Statistics, P.S. King & Son, Laondon. Or in a later edition: BOWLEY, AL. "Elements of Statistics, 4th Edn (New York, Charles Scribner)."(1920).
18. Kenney JF and Keeping ES (1962) *Mathematics of Statistics, Pt. 1, 3rd ed.*, Van Nostrand, (page 102).
19. Yule, George Udny. An introduction to the theory of statistics. C. Griffin, limited, 1912.
20. Groeneveld, Richard A. "An influence function approach to describing the skewness of a distribution." The American Statistician 45.2 (1991): 97-102.
21. Johnson *et al* (1994) p 3, p 40
22. Wilks DS (1995) *Statistical Methods in the Atmospheric Sciences*, p 27. Academic Press. ISBN 0-12-751965-3
23. Groeneveld, R.A.; Meeden, G. (1984). "Measuring Skewness and Kurtosis". *The Statistician*. **33** (4): 391–399. doi:10.2307/2987742 (https://doi.org/10.2307%2F2987742). JSTOR 2987742 (https://www.jstor.org/stable/2987742).
24. MacGillivray (1992)
25. Hinkley DV (1975) "On power transformations to symmetry", *Biometrika, 62, 101–111*
26. A.W.L. Pubudu Thilan. "Applied Statistics I: Chapter 5: Measures of skewness" (http://www.math.ruh.ac.lk/~pubudu/app5.pdf) (PDF). *University of Ruhuna*. p. 21.
27. Hosking, J.R.M. (1992). "Moments or L moments? An example comparing two measures of distributional shape". *The American Statistician*. **46** (3): 186–189. doi:10.2307/2685210 (https://doi.org/10.2307%2F2685210). JSTOR 2685210 (https://www.jstor.org/stable/2685210).
28. Szekely, G.J. (2000). "Pre-limit and post-limit theorems for statistics", In: *Statistics for the 21st Century* (eds. C. R. Rao and G. J. Szekely), Dekker, New York, pp. 411–422.
29. Szekely, G. J. and Mori, T. F. (2001) "A characteristic measure of asymmetry and its application for testing diagonal symmetry", *Communications in Statistics – Theory and Methods* 30/8&9, 1633–1639.
30. G. Brys; M. Hubert; A. Struyf (November 2004). "A Robust Measure of Skewness". *Journal of Computational and Graphical Statistics*. **13** (4): 996–1017. doi:10.1198/106186004X12632 (https://doi.org/10.1198%2F106186004X12632).

- Premaratne, G., Bera, A. K. (2001).

Adjusting the Tests for Skewness and Kurtosis for Distributional Misspecifications. Working Paper Number 01-0116, University of Illinois. Forthcoming in Comm in Statistics, Simulation and Computation.2016 1-15

- Premaratne, G., Bera, A. K. (2000). Modeling Asymmetry and Excess Kurtosis in Stock Return Data. Office of Research Working Paper Number 00-0123, University of Illinois

References

- Johnson, NL, Kotz, S, Balakrishnan N (1994) *Continuous Univariate Distributions, Vol 1, 2nd Edition* Wiley ISBN 0-471-58495-9
- MacGillivray, HL (1992). "Shape properties of the g- and h- and Johnson families". *Comm. Statistics — Theory and Methods*. **21**: 1244–1250.
- Skewness Measures for the Weibull Distribution (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2590356)

External links

- Hazewinkel, Michiel, ed. (2001) [1994], "Asymmetry coefficient" (https://www.encyclopediaofmath.org/index.php?title=p/a013590), *Encyclopedia of Mathematics*, Springer Science+Business Media B.V. / Kluwer Academic Publishers, ISBN 978-1-55608-010-4
- An Asymmetry Coefficient for Multivariate Distributions (http://petitjeanmichel.free.fr/itoweb.petitjean.skewness.html) by Michel Petitjean
- On More Robust Estimation of Skewness and Kurtosis (http://repositories.cdlib.org/cgi/viewcontent.cgi?article=1017&context=ucsdecon) Comparison of skew estimators by Kim and White.
- Closed-skew Distributions — Simulation, Inversion and Parameter Estimation (http://dahoiv.net/master/index.html)