WIKIPEDIA

Kurtosis

In probability theory and statistics, **kurtosis** (from Greek: κυρτός, *kyrtos* or *kurtos*, meaning "curved, arching") is a measure of the "tailedness" of the probability distribution of a real-valued random variable. In a similar way to the concept of skewness, kurtosis is a descriptor of the shape of a probability distribution and, just as for skewness, there are different ways of quantifying it for a theoretical distribution and corresponding ways of estimating it from a sample from a population. Depending on the particular measure of kurtosis that is used, there are various interpretations of kurtosis, and of how particular measures should be interpreted.

The standard measure of kurtosis, originating with Karl Pearson, is based on a scaled version of the fourth moment of the data or population. This number is related to the tails of the distribution, not its peak;[1] hence, the sometimes-seen characterization as "peakedness" is mistaken. For this measure, higher kurtosis is the result of infrequent extreme deviations (or outliers), as opposed to frequent modestly sized deviations.

The kurtosis of any univariate normal distribution is 3. It is common to compare the kurtosis of a distribution to this value. Distributions with kurtosis less than 3 are said to be *platykurtic*, although this does not imply the distribution is "flat-topped" as sometimes reported. Rather, it means the distribution produces fewer and less extreme outliers than does the normal distribution. An example of a platykurtic distribution is the uniform distribution, which does not produce outliers. Distributions with kurtosis greater than 3 are said to be *leptokurtic*. An example of a leptokurtic distribution is the Laplace distribution, which has tails that asymptotically approach zero more slowly than a Gaussian, and therefore produces more outliers than the normal distribution. It is also common practice to use an adjusted version of Pearson's kurtosis, the excess kurtosis, which is the kurtosis minus 3, to provide the comparison to the normal distribution. Some authors use "kurtosis" by itself to refer to the excess kurtosis. For the reason of clarity and generality, however, this article follows the non-excess convention and explicitly indicates where excess kurtosis is meant.

Alternative measures of kurtosis are: the L-kurtosis, which is a scaled version of the fourth L-moment; measures based on four population or sample quantiles.[2] These are analogous to the alternative measures of skewness that are not based on ordinary moments.[2]

Contents

Pearson moments

The kurtosis is the fourth standardized moment, defined as

$$\operatorname{Kurt}[X] = \operatorname{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] = \frac{\mu_4}{\sigma^4} = \frac{\operatorname{E}[(X-\mu)^4]}{(\operatorname{E}[(X-\mu)^2])^2},$$

where μ_4 is the fourth central moment and σ is the standard deviation. Several letters are used in the literature to denote the kurtosis. A very common choice is κ, which is fine as long as it is clear that it does not refer to a cumulant. Other choices include γ_2, to be similar to the notation for skewness, although sometimes this is instead reserved for the excess kurtosis.

The kurtosis is bounded below by the squared skewness plus 1:[3]

$$\frac{\mu_4}{\sigma^4} \geq \left(\frac{\mu_3}{\sigma^3}\right)^2 + 1,$$

where μ_3 is the third central moment. The lower bound is realized by the Bernoulli distribution. There is no upper limit to the excess kurtosis of a general probability distribution, and it may be infinite.

A reason why some authors favor the excess kurtosis is that cumulants are extensive. Formulas related to the extensive property are more naturally expressed in terms of the excess kurtosis. For example, let $X_1, ..., X_n$ be independent random variables for which the fourth moment exists, and let Y be the random variable defined by the sum of the X_i. The excess kurtosis of Y is

$$\operatorname{Kurt}[Y]-3=\frac{1}{\left(\sum_{j=1}^{n}\sigma_j^{\,2}\right)^2}\sum_{i=1}^{n}\sigma_i^{\,4}\cdot\left(\operatorname{Kurt}\left[X_i\right]-3\right),$$

where σ_i is the standard deviation of X_i. In particular if all of the X_i have the same variance, then this simplifies to

$$\operatorname{Kurt}[Y]-3=\frac{1}{n^2}\sum_{i=1}^{n}\left(\operatorname{Kurt}\left[X_i\right]-3\right).$$

The reason not to subtract off 3 is that the bare fourth moment better generalizes to multivariate distributions, especially when independence is not assumed. The cokurtosis between pairs of variables is an order four tensor. For a bivariate normal distribution, the cokurtosis tensor has off-diagonal terms that are neither 0 nor 3 in general, so attempting to "correct" for an excess becomes confusing. It is true, however, that the joint cumulants of degree greater than two for any multivariate normal distribution are zero.

For two random variables, X and Y, not necessarily independent, the kurtosis of the sum, $X + Y$, is

$$\begin{aligned}\operatorname{Kurt}[X+Y]=\frac{1}{\sigma_{X+Y}^4}\big(&\sigma_X^4\operatorname{Kurt}[X]+4\sigma_X^3\sigma_Y\operatorname{Cokurt}[X,X,X,Y]\\&+6\sigma_X^2\sigma_Y^2\operatorname{Cokurt}[X,X,Y,Y]\\&+4\sigma_X\sigma_Y^3\operatorname{Cokurt}[X,Y,Y,Y]+\sigma_Y^4\operatorname{Kurt}[Y]\big).\end{aligned}$$

Note that the binomial coefficients appear in the above equation.

Interpretation

The exact interpretation of the Pearson measure of kurtosis (or excess kurtosis) used to be disputed, but is now settled. As Westfall (2014)[4] notes, "...its only unambiguous interpretation is in terms of tail extremity; i.e., either existing outliers (for the sample kurtosis) or propensity to produce outliers (for the kurtosis of a probability distribution)." The logic is simple: Kurtosis is the average (or expected value) of the standardized data raised to the fourth power. Any standardized values that are less than 1 (i.e., data within one standard deviation of the mean, where the "peak" would be), contribute virtually nothing to kurtosis, since raising a number that is less than 1 to the fourth power makes it closer to zero. The only data values (observed or observable) that contribute to kurtosis in any meaningful way are those outside the region of the peak; i.e., the outliers. Therefore, kurtosis measures outliers only; it measures nothing about the "peak".

Many incorrect interpretations of kurtosis that involve notions of peakedness have been given. One is that kurtosis measures both the "peakedness" of the distribution and the heaviness of its tail.[5] Various other incorrect interpretations have been suggested, such as "lack of shoulders" (where the "shoulder" is defined vaguely as the area between the peak and the tail, or more specifically as the area about one standard deviation from the mean) or "bimodality".[6] Balanda and MacGillivray assert that the standard definition of kurtosis "is a poor measure of the kurtosis, peakedness, or tail weight of a distribution"[7] and instead propose to "define kurtosis vaguely as the location- and scale-free movement of probability mass from the shoulders of a distribution into its center and tails".[5]

Moors' interpretation

In 1986 Moors gave an interpretation of kurtosis.[8] Let

$$Z = \frac{X - \mu}{\sigma},$$

where X is a random variable, μ is the mean and σ is the standard deviation.

Now by definition of the kurtosis κ, and by the well-known identity $E[V^2] = \operatorname{var}[V] + [E[V]]^2$,

$$\kappa = E[Z^4] = \operatorname{var}[Z^2] + [E[Z^2]]^2 = \operatorname{var}[Z^2] + [\operatorname{var}[Z]]^2 = \operatorname{var}[Z^2] + 1\ .$$

The kurtosis can now be seen as a measure of the dispersion of Z^2 around its expectation. Alternatively it can be seen to be a measure of the dispersion of Z around +1 and −1. κ attains its minimal value in a symmetric two-point distribution. In terms of the original variable X, the kurtosis is a measure of the dispersion of X around the two values $\mu \pm \sigma$.

High values of κ arise in two circumstances:

- where the probability mass is concentrated around the mean and the data-generating process produces occasional values far from the mean,
- where the probability mass is concentrated in the tails of the distribution.

Excess kurtosis

The *excess kurtosis* is defined as kurtosis minus 3. There are 3 distinct regimes as described below.

Mesokurtic

Distributions with zero excess kurtosis are called **mesokurtic**, or mesokurtotic. The most prominent example of a mesokurtic distribution is the normal distribution family, regardless of the values of its parameters. A few other well-known distributions can be mesokurtic, depending on parameter values: for example, the binomial distribution is mesokurtic for $p = 1/2 \pm \sqrt{1/12}$.

Leptokurtic

A distribution with positive excess kurtosis is called **leptokurtic**, or leptokurtotic. "Lepto-" means "slender".[9] In terms of shape, a leptokurtic distribution has *fatter tails*. Examples of leptokurtic distributions include the Student's t-distribution, Rayleigh distribution, Laplace distribution, exponential distribution, Poisson distribution and the logistic distribution. Such distributions are sometimes termed *super-Gaussian*.[10]

Platykurtic

A distribution with negative excess kurtosis is called **platykurtic**, or platykurtotic. "Platy-" means "broad".[11] In terms of shape, a platykurtic distribution has *thinner tails*. Examples of platykurtic distributions include the continuous and discrete uniform distributions, and the raised cosine distribution. The most platykurtic distribution of all is the Bernoulli distribution with $p = 1/2$ (for example the number of times one obtains "heads" when flipping a coin once, a coin toss), for which the excess kurtosis is −2. Such distributions are sometimes termed *sub-Gaussian*.[12]

The coin toss is the most platykurtic distribution

Graphical examples

The Pearson type VII family

The effects of kurtosis are illustrated using a parametric family of distributions whose kurtosis can be adjusted while their lower-order moments and cumulants remain constant. Consider the Pearson type VII family, which is a special case of the Pearson type IV family restricted to symmetric densities. The probability density function is given by

$$f(x;a,m)=\frac{\Gamma(m)}{a\,\sqrt{\pi}\,\Gamma(m-1/2)}\left[1+\left(\frac{x}{a}\right)^2\right]^{-m},$$

where a is a scale parameter and m is a shape parameter.

All densities in this family are symmetric. The kth moment exists provided $m > (k + 1)/2$. For the kurtosis to exist, we require $m > 5/2$. Then the mean and skewness exist and are both identically zero. Setting $a^2 = 2m - 3$ makes the variance equal to unity. Then the only free parameter is m, which controls the fourth moment (and cumulant) and hence the kurtosis. One can reparameterize with $m = 5/2 + 3/\gamma_2$, where γ_2 is the excess kurtosis as defined above. This yields a one-parameter leptokurtic family with zero mean, unit variance, zero skewness, and arbitrary non-negative excess kurtosis. The reparameterized density is

$$g(x;\gamma_2)=f\left(x;\ a=\sqrt{2+\frac{6}{\gamma_2}},\ m=\frac{5}{2}+\frac{3}{\gamma_2}\right).$$

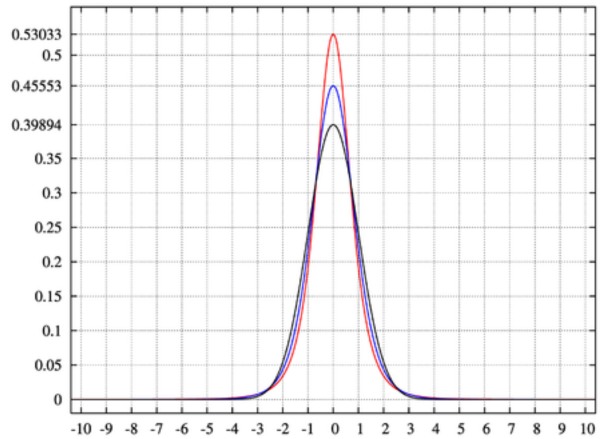
In the limit as $\gamma_2\to\infty$ one obtains the density

$$g(x)=3\left(2+x^2\right)^{-\frac{5}{2}},$$

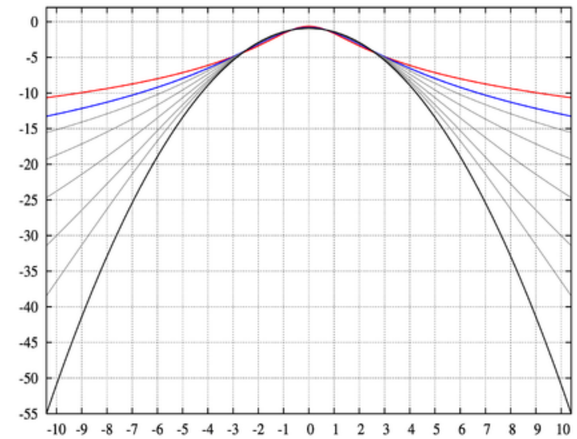
which is shown as the red curve in the images on the right.

In the other direction as $\gamma_2\to 0$ one obtains the standard normal density as the limiting distribution, shown as the black curve.

In the images on the right, the blue curve represents the density $x\mapsto g(x;2)$ with excess kurtosis of 2. The top image shows that leptokurtic densities in this family have a higher peak than the mesokurtic normal density, although this conclusion is only valid for this select family of distributions. The comparatively fatter tails of the leptokurtic densities are illustrated in the second image, which plots the natural logarithm of the Pearson type VII densities: the black curve is the logarithm of the standard normal density, which is a parabola. One can see that the normal density allocates little probability mass to the regions far from the mean ("has thin tails"), compared with the blue curve of the leptokurtic Pearson type VII density with excess kurtosis of 2. Between the blue curve and the black are other Pearson type VII densities with γ_2 = 1, 1/2, 1/4, 1/8, and 1/16. The red curve again shows the upper limit of the Pearson type VII family, with $\gamma_2=\infty$ (which, strictly speaking, means that the fourth moment does not exist). The red curve decreases the slowest as one moves outward from the origin ("has fat tails").

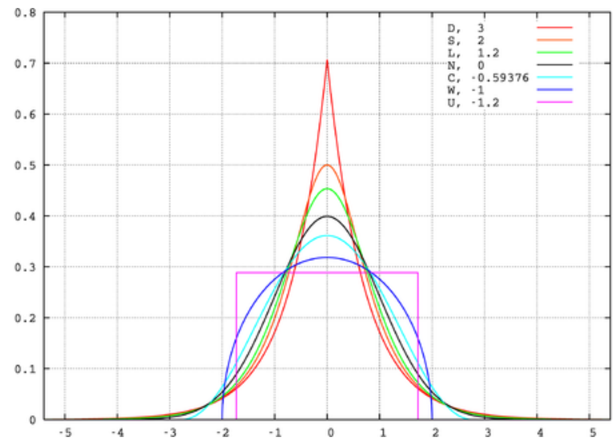


pdf for the Pearson type VII distribution with excess kurtosis of infinity (red); 2 (blue); and 0 (black)

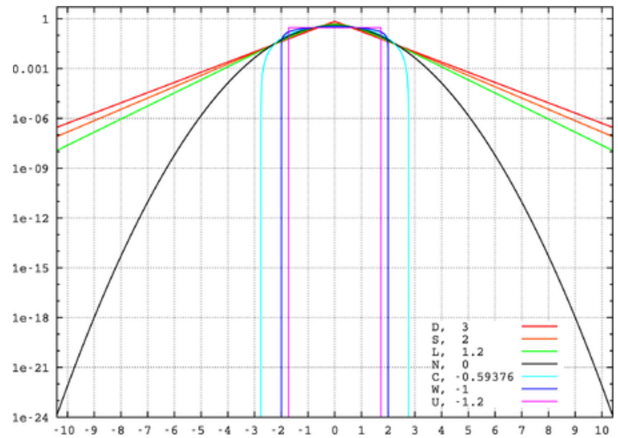


log-pdf for the Pearson type VII distribution with excess kurtosis of infinity (red); 2 (blue); 1, 1/2, 1/4, 1/8, and 1/16 (gray); and 0 (black)

Of well-known distributions



Probability density functions for selected distributions with mean 0, variance 1 and different excess kurtosis



Logarithms of probability density functions for selected distributions with mean 0, variance 1 and different excess kurtosis

Several well-known, unimodal and symmetric distributions from different parametric families are compared here. Each has a mean and skewness of zero. The parameters have been chosen to result in a variance equal to 1 in each case. The images on the right show curves for the following seven densities, on a linear scale and logarithmic scale:

- D: Laplace distribution, also known as the double exponential distribution, red curve (two straight lines in the log-scale plot), excess kurtosis = 3
- S: hyperbolic secant distribution, orange curve, excess kurtosis = 2
- L: logistic distribution, green curve, excess kurtosis = 1.2
- N: normal distribution, black curve (inverted parabola in the log-scale plot), excess kurtosis = 0
- C: raised cosine distribution, cyan curve, excess kurtosis = −0.593762...
- W: Wigner semicircle distribution, blue curve, excess kurtosis = −1
- U: uniform distribution, magenta curve (shown for clarity as a rectangle in both images), excess kurtosis = −1.2.

Note that in these cases the platykurtic densities have bounded support, whereas the densities with positive or zero excess kurtosis are supported on the whole real line.

There exist platykurtic densities with infinite support,

- e.g., exponential power distributions with sufficiently large shape parameter *b*

and there exist leptokurtic densities with finite support.

- e.g., a distribution that is uniform between −3 and −0.3, between −0.3 and 0.3, and between 0.3 and 3, with the same density in the (−3, −0.3) and (0.3, 3) intervals, but with 20 times more density in the (−0.3, 0.3) interval

Sample kurtosis

For a sample of n values the **sample excess kurtosis** is

$$g_2 = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{n}\sum_{i=1}^n (x_i - \overline{x})^4}{\left(\frac{1}{n}\sum_{i=1}^n (x_i - \overline{x})^2\right)^2} - 3$$

where m_4 is the fourth sample moment about the mean, m_2 is the second sample moment about the mean (that is, the sample variance), x_i is the i^{th} value, and $\overline{x}$ is the sample mean.

This formula has the simpler representation,

$$g_2 = \frac{1}{n} \sum_{i=1}^{n} z_i^4 - 3$$

where the z_i values are the standardized data values using the standard deviation defined using n rather than $n - 1$ in the denominator.

For example, suppose the data values are 0, 3, 4, 1, 2, 3, 0, 2, 1, 3, 2, 0, 2, 2, 3, 2, 5, 2, 3, 999.

Then the z_i values are −0.239, −0.225, −0.221, −0.234, −0.230, −0.225, −0.239, −0.230, −0.234, −0.225, −0.230, −0.239, −0.230, −0.230, −0.225, −0.230, −0.216, −0.230, −0.225, 4.359

and the z_i^4 values are 0.003, 0.003, 0.002, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.002, 0.003, 0.003, 360.976.

The average of these values is 18.05 and the excess kurtosis is thus 18.05 − 3 = 15.05. This example makes it clear that data near the "middle" or "peak" of the distribution do not contribute to the kurtosis statistic, hence kurtosis does not measure "peakedness". It is simply a measure of the outlier, 999 in this example.

Sampling variance under normality

The variance of the sample kurtosis of a sample of size n from the normal distribution is[13]

$$\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}$$

Stated differently, under the assumption that the underlying random variable X is normally distributed, it can be shown that $\sqrt{n}g_2 \xrightarrow{d} N(0, 24)$.[14]

Upper bound

An upper bound for the sample kurtosis of n ($n > 2$) real numbers is[15]

$$\frac{\mu_4}{\sigma^4} \leq \frac{1}{2} \frac{n-3}{n-2} \left(\frac{\mu_3}{\sigma^3}\right)^2 + \frac{n}{2}.$$

Estimators of population kurtosis

Given a sub-set of samples from a population, the sample excess kurtosis above is a biased estimator of the population excess kurtosis. An alternative estimator of the population excess kurtosis is defined as follows:

$$\begin{aligned}
G_2 &= \frac{k_4}{k_2^2} \\
&= \frac{n^2\,((n+1)\,m_4 - 3\,(n-1)\,m_2^2)}{(n-1)\,(n-2)\,(n-3)}\;\frac{(n-1)^2}{n^2\,m_2^2} \\
&= \frac{n-1}{(n-2)\,(n-3)}\left((n+1)\,\frac{m_4}{m_2^2} - 3\,(n-1)\right) \\
&= \frac{n-1}{(n-2)\,(n-3)}\,((n+1)\,g_2 + 6) \\
&= \frac{(n+1)\,n\,(n-1)}{(n-2)\,(n-3)}\;\frac{\sum_{i=1}^n (x_i - \bar{x})^4}{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right)^2} - 3\,\frac{(n-1)^2}{(n-2)\,(n-3)} \\
&= \frac{(n+1)\,n}{(n-1)\,(n-2)\,(n-3)}\;\frac{\sum_{i=1}^n (x_i - \bar{x})^4}{k_2^2} - 3\,\frac{(n-1)^2}{(n-2)(n-3)}
\end{aligned}$$

where k_4 is the unique symmetric unbiased estimator of the fourth cumulant, k_2 is the unbiased estimate of the second cumulant (identical to the unbiased estimate of the sample variance), m_4 is the fourth sample moment about the mean, m_2 is the second sample moment about the mean, x_i is the i^{th} value, and $\bar{x}$ is the sample mean. Unfortunately, G_2 is itself generally biased. For the normal distribution it is unbiased.[2]

Applications

The sample kurtosis is a useful measure of whether there is a problem with outliers in a data set. Larger kurtosis indicates a more serious outlier problem, and may lead the researcher to choose alternative statistical methods.

D'Agostino's K-squared test is a goodness-of-fit normality test based on a combination of the sample skewness and sample kurtosis, as is the Jarque–Bera test for normality.

For non-normal samples, the variance of the sample variance depends on the kurtosis; for details, please see variance.

Pearson's definition of kurtosis is used as an indicator of intermittency in turbulence.[16]

Kurtosis convergence

Applying band-pass filters to digital images, kurtosis values tend to be uniform, independent of the range of the filter. This behavior, termed *kurtosis convergence*, can be used to detect image splicing in forensic analysis.[17]

Other measures

A different measure of "kurtosis" is provided by using L-moments instead of the ordinary moments.[18][19]

See also

- Kurtosis risk
- Maximum entropy probability distribution

References

1. Westfall, PH (2014). "Kurtosis as Peakedness, 1905 - 2014. *R.I.P.*" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4321753). *Am Stat*. **68** (3): 191–195. doi:10.1080/00031305.2014.917055 (https://doi.org/10.1080%2F00031305.2014.917055). PMC 4321753 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4321753). PMID 25678714 (https://www.ncbi.nlm.nih.gov/pubmed/25678714).
2. 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).
3. Pearson, K. (1929). "Editorial note to 'Inequalities for moments of frequency functions and for various statistical constants'". *Biometrika*. **21** (1–4): 361–375. doi:10.1093/biomet/21.1-4.361 (https://doi.org/10.1093%2Fbiomet%2F21.1-4.361).
4. Westfall, PH (2014). "Kurtosis as Peakedness, 1905 - 2014. *R.I.P.*" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4321753). *The American Statistician*. **68** (3): 191–195. doi:10.1080/00031305.2014.917055 (https://doi.org/10.1080%2F00031305.2014.917055). PMC 4321753 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4321753). PMID 25678714 (https://www.ncbi.nlm.nih.gov/pubmed/25678714).
5. Balanda, Kevin P.; MacGillivray, H.L. (1988). "Kurtosis: A Critical Review". *The American Statistician*. **42** (2): 111–119. doi:10.2307/2684482 (https://doi.org/10.2307%2F2684482). JSTOR 2684482 (https://www.jstor.org/stable/2684482).
6. Darlington, Richard B (1970). "Is Kurtosis Really 'Peakedness'?". *The American Statistician*. **24** (2): 19–22. doi:10.1080/00031305.1970.10478885 (https://doi.org/10.1080%2F00031305.1970.10478885).
7. Balanda and MacGillivray, p. 114.
8. Moors, JJA (1986b). "The meaning of kurtosis: Darlington reexamined". *The American Statistician*. **40** (4): 283–284. doi:10.1080/00031305.1986.10475415 (https://doi.org/10.1080%2F00031305.1986.10475415).
9. http://medical-dictionary.thefreedictionary.com/lepto-
10. Benveniste, A.; Goursat, M.; Ruget, G. (June 1980). "Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications". *IEEE Transactions on Automatic Control*. **25** (3): 385–399. doi:10.1109/tac.1980.1102343 (https://doi.org/10.1109%2Ftac.1980.1102343). ISSN 0018-9286 (https://www.worldcat.org/issn/0018-9286).
11. http://www.yourdictionary.com/platy-prefix

12. The original paper presenting sub-Gaussians Kahane, J. P. (1960). "Propriétés locales des fonctions à séries de Fourier aléatoires" (https://eudml.org/doc/216962) [Local properties of functions in terms of random Fourier series]. *Stud. Math.* **19** (1): 1–25. doi:10.4064/sm-19-1-1-25 (https://doi.org/10.4064%2Fsm-19-1-1-25). ISSN 0039-3223 (https://www.worldcat.org/issn/0039-3223). See also Buldygin, V. V.; Kozachenko, Y. V. (1980). "Sub-Gaussian random variables". *Ukrainian Mathematical Journal*. **32** (6): 483–489. doi:10.1007/BF01087176 (https://doi.org/10.1007%2FBF01087176).
13. Cramer, Duncan (1997). *Fundamental Statistics for Social Research*. Routledge. p. 89. ISBN 978-0-415-17204-2.
14. Kendall, M.G.; Stuart, A. (1969) *The Advanced Theory of Statistics, Volume 1: Distribution Theory, 3rd Edition*, Griffin. ISBN 0-85264-141-9
15. Sharma, R.; Bhandari, R. (11 Sep 2013). "Skewness, kurtosis and Newton's inequality". *Rocky Mountain Journal of Mathematics*. **45** (5): 1639–1643. arXiv:1309.2896 (https://arxiv.org/abs/1309.2896). Bibcode:2013arXiv1309.2896S (https://ui.adsabs.harvard.edu/abs/2013arXiv1309.2896S). doi:10.1216/RMJ-2015-45-5-1639 (https://doi.org/10.1216%2FRMJ-2015-45-5-1639). MR 3452232 (https://www.ams.org/mathscinet-getitem?mr=3452232). Zbl 1355.60028 (https://zbmath.org/?format=complete&q=an:1355.60028). Check date values in: `|year= / |date= mismatch` (help)
16. Sandborn, V. A. (1959). "Measurements of Intermittency of Turbulent Motion in a Boundary Layer". *Journal of Fluid Mechanics*. **6** (2): 221–240. Bibcode:1959JFM.....6..221S (https://ui.adsabs.harvard.edu/abs/1959JFM.....6..221S). doi:10.1017/S0022112059000581 (https://doi.org/10.1017%2FS0022112059000581).
17. "Archived copy" (https://web.archive.org/web/20140903073743/http://www.cs.albany.edu/~lsw/homepage/PUBLICATIONS_files/ICCP.pdf) (PDF). Archived from the original (http://www.cs.albany.edu/~lsw/homepage/PUBLICATIONS_files/ICCP.pdf) (PDF) on 2014-09-03. Retrieved 2015-04-17.
18. 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).
19. Hosking, J.R.M. (2006). "On the characterization of distributions by their L-moments". *Journal of Statistical Planning and Inference*. **136**: 193–198. doi:10.1016/j.jspi.2004.06.004 (https://doi.org/10.1016%2Fj.jspi.2004.06.004).

Further reading

- Kim, Tae-Hwan; White, Halbert (2003). "On More Robust Estimation of Skewness and Kurtosis: Simulation and Application to the S&P500 Index" (http://escholarship.org/uc/item/7b52v07p). *Finance Research Letters*. **1**: 56–70. doi:10.1016/S1544-6123(03)00003-5 (https://doi.org/10.1016%2FS1544-6123%2803%2900003-5). Alternative source (https://web.archive.org/web/20111118123903/http://weber.ucsd.edu/~hwhite/pub_files/hwcv-092.pdf) (Comparison of kurtosis estimators)
- Seier, E.; Bonett, D.G. (2003). "Two families of kurtosis measures". *Metrika*. **58**: 59–70. doi:10.1007/s001840200223 (https://doi.org/10.1007%2Fs001840200223).

External links

- Hazewinkel, Michiel, ed. (2001) [1994], "Excess coefficient" (https://www.encyclopediaofmath.org/index.php?title=p/e036800), *Encyclopedia of Mathematics*, Springer Science+Business Media B.V. / Kluwer Academic Publishers, ISBN 978-1-55608-010-4
- Kurtosis calculator (http://www.fxsolver.com/solve/share/RMqwaVp85T_5rbacksPD4g==/)
- Free Online Software (Calculator) (https://archive.is/20121208231710/http://www.wessa.net/skewkurt.wasp) computes various types of skewness and kurtosis statistics for any dataset (includes small and large sample tests)..
- Kurtosis (http://jeff560.tripod.com/k.html) on the Earliest known uses of some of the words of mathematics (http://jeff560.tripod.com/mathword.html)
- Celebrating 100 years of Kurtosis (http://faculty.etsu.edu/seier/doc/Kurtosis100years.doc) a history of the topic, with different measures of kurtosis.