Examples of incorrect applications of the measures of asymmetry and shape

Emilija Nikolić-Đorić[1], Zagorka Lozanov-Crvenković[2]
Dragan Đorić[3]

[1]Department of Agricultural Economics and Rural Sociology, Faculty of Agriculture, University of Novi Sad, Serbia
[2]Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad, Serbia
[3]Faculty of Organizational Sciences, University of Belgrade, Serbia

Challenges and Innovations in Statistics Education
Szeged, September 7-9, 2017

Motivation

- Incorrect applications of measures of asymmetry and shape for one dimensional distributions.
- The coefficient of asymmetry or skewness is often misinterpreted. The fact that it is zero does not imply that the distribution is symmetric both for discrete and continuous distributions.
- It is also often wrongly thought that the coefficient of kurtosis measures the peakness of a distribution and that leptocurtic curves are more sharply peaked and that platykurtic curves are more flat-topped than the normal curve.

Motivation

- Misconceptions of establishing skewness using relative positions of mean, median and mode frequently occur in introductory textbooks.
- The rule is that if the distribution is skewed to the right, then the median is greater than the mode and vice versa for the skewness to the left
- We give the examples of violations of that rule for discrete and continuous distributions.
- Also, the example which disproves assertion that for skewed distributions the mean lies toward the direction of skew relative to the median is presented.

Motivation

- The normal, Gaussian, distribution is one of the most frequently used distribution in statistics.
- However, we often deal with small samples with sampling distributions other than normal, and hence the central limit theorem cannot be applied. Gery (1947)
- On occasion when we cannot assume that data are normally distributed, we have to measure the magnitude of deviation of the sampling distribution from the normal distribution. This can be achieved by measuring skewness and kurtosis.
- The coefficient of skewness shows asymmetry of one dimensional distributions, and can be calculated on a basis of the first three moments of the distribution.
- In the case of symmetrical distributions, it is important to find out how much the distribution differs from the normal distribution. Kurtosis is the most common coefficient for this purpose.

Measuring asymmetry

- If the probability density function of a continuous random variable X has the line $m = E(X)$ for the axis of symmetry, then the probability density function is said to be symmetrical. If the probability density function is symmetrical and has a unique mode M_0 then the mode, the median M_e and the mean $E(X)$ are equal.
- Some well-known distributions have this property: the normal distribution, the uniform distribution, Student's distribution...
- Let the distribution of the random variable X, continuous or discrete, be symmetrical and $m = E(X)$ its mean value. Then all central moments of odd order μ_{2k+1} (if they exist) are equal to 0:

$$\mu_{2k+1} = E\left((X-m)^{2k+1}\right) = 0, \quad k = 0, 1, 2, \ldots .$$

- If there is at least one central moment of odd order not equal to 0, then the distribution is not symmetric. This makes central moments of odd order a measure of asymmetry.

Measuring asymmetry

- The *coefficient of asymmetry or coefficient of skewness* is given by:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(D(X))^{3/2}}.$$

where σ is the standard deviation of a random variable X.

- The coefficient γ_1 is a real number, i.e. a dimensionless quantity.
- If the distribution is symmetrical, then $\gamma_1 = 0$.

Measuring asymmetry

- Distributions with $\gamma_1 > 0$ are said to be skewed to the right (Figure, left), or to have positive skewness.
- The distributions with $\gamma_1 < 0$ are said to be skewed to the left (Figure, right), that is, to have negative skewness.

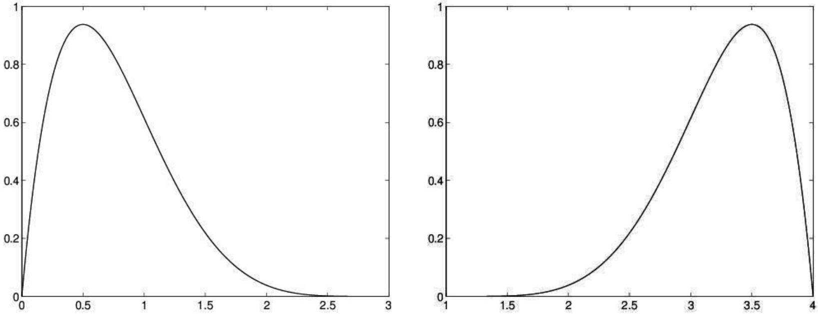


Figure: Skewed distribution

Measuring asymmetry

- Karl Pearson (1895) introduced the coefficient $\beta_1 = \gamma_1^2$ in order to measure asymmetry, but β_1 could not detect positive or negative asymmetry.
- Since 1895 Pearson used $Sk = \dfrac{m - M_o}{\sigma}$ for a measure of asymmetry, where M_o is the mode, and σ is the standard deviation of the random variable X with $E(X) = m$.
- This coefficient satisfies the double inequality: $-1 \leq S_k \leq 1$.

Measuring asymmetry

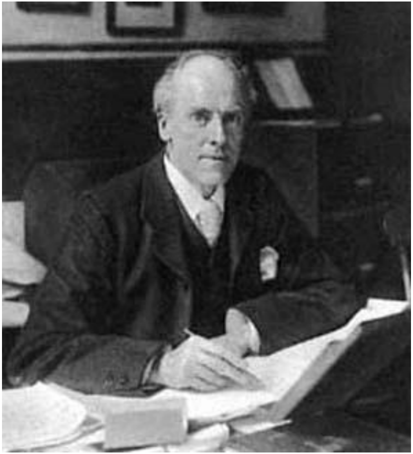


Figure: Karl Pearson, 1857-1936

Measuring asymmetry

- Although positive and negative skewness are usually presented as in previous figure, there are cases when this graphical presentation is misleading.
- For example, in the case of Pareto distribution coefficient γ_1 may be positive or negative depending on parameters, while the shape of density is always the same skewed to the right.

- The following assertion can be found in statistical literature:
- If the distribution is skewed to the right, then the median is greater than the mode and vice versa for the skewness to the left. This is not true for every distribution, as it can be seen from the following example.

Example 1. *Let X be the random variable with probability density function.*

$$g(x) = \begin{cases} \dfrac{1}{\sqrt{\pi}} e^{-x^2}, & x \leq 0 \\ -x/\pi + 1/\sqrt{\pi}, & 0 < x \leq \sqrt{\pi} \\ 0, & x > \sqrt{\pi}. \end{cases}$$

It is easy to find out that the mode and the median are equal, while the distribution is obviously asymmetrical.

Measuring asymmetry

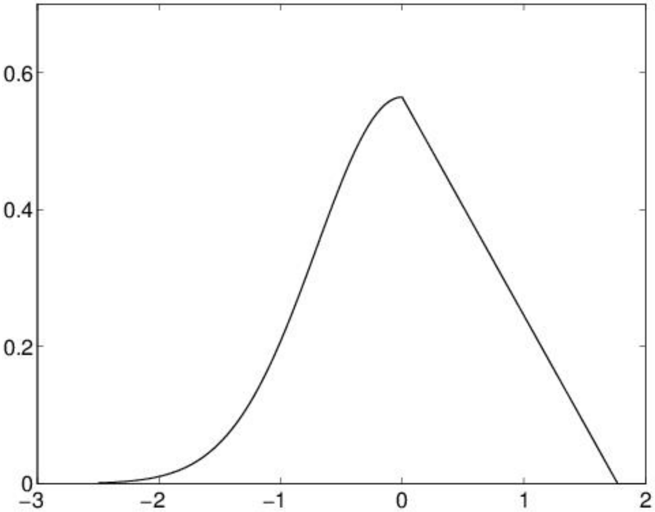


Figure: Density of distribution in Example 1

Measuring asymmetry

- If the discrete random variable X has only two values, then, if $\mu_3 = 0$, the distribution has to be symmetrical.
- If the random variable X is continuous or discrete, with more than two values, then $\mu_3 = 0$ does not imply the symmetry of the distribution.

Example 2. *Let random variable X have the distribution*

$$X : \begin{pmatrix} -4 & 1 & 5 \\ 1/3 & 1/2 & 1/6 \end{pmatrix}$$

- *We have*

$$\mu_3 = -4^3 \cdot \frac{1}{3} + \frac{1}{2} + 5^3 \cdot \frac{1}{6} = 0, \quad \mu_5 = -4^5 \cdot \frac{1}{3} + \frac{1}{2} + 5^5 \cdot \frac{1}{6} = 180.$$

- *Values of the next odd moments are $\mu_7 = 7560$ and $\mu_9 = 238140$. It follows that $\gamma_1 = 0$, but the distribution is asymmetrical. The random variable given in this example can be obtained by tossing a die which has one side marked with 5, three with 1, and two sides with -4.*

- It is possible to give an adequate example of a continuous random variable which is asymmetrical, but $\gamma_1 = 0$.

Example 3. *Let probability density function for the random variable X be*

$$g(x) = \begin{cases} \dfrac{\alpha}{1+x^2}, & x < 0 \\ \dfrac{\alpha}{1+x^2} + f(x), & x \geq 0 \end{cases}$$

with

$$f(x) = e^{-x}\left(\frac{1}{10}\sin 2x - \frac{4}{125}\sin x - \frac{48}{3125}\cos x\right), \quad \alpha = \frac{3074}{3125}\cdot\frac{1}{\pi}.$$

We have $m = \mu_1 = \mu_3 = 0$, $\mu_5 = 0.5138$, $\mu_7 = -4.4049$, $\mu_9 = -374.4587$ $\mu_{11} = 9748.4$ *etc.*

Measuring asymmetry

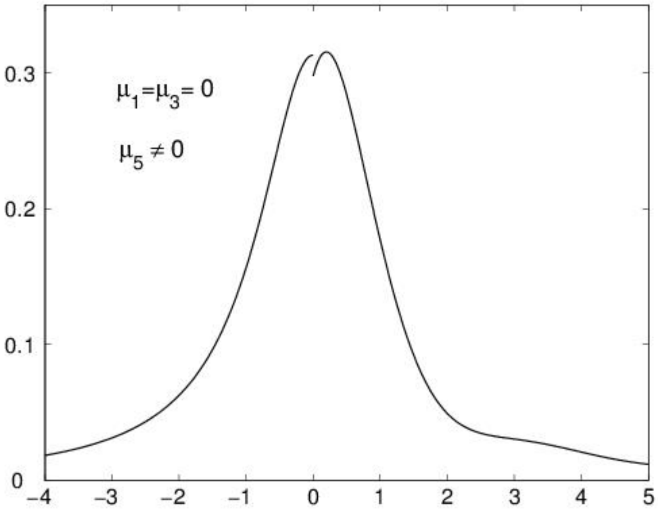


Figure: Density of distribution in Example 3

Example 4. (Stoyanov, 1987) *Let probability density function for the random variable X be*

$$g(x) = \begin{cases} \frac{1}{48}(1+\sin|x|^{1/4})e^{-|x|^{1/4}}, & x < 0 \\ \frac{1}{48}(1-\sin x^{1/4})e^{-x^{1/4}}, & x \geq 0 \end{cases}$$

This distribution is asymmetrical with $\mu_{2n+1}(X) = 0$ *i* $\mu_{2n}(X) = (8n+3)!/6$ *for* $n \in N$.

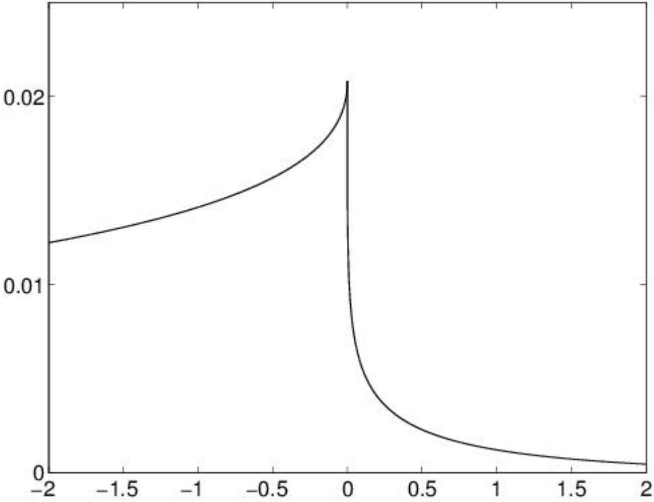


Figure: Density of distribution in Example 4

Measuring asymmetry

- We can also obtain asymmetrical distributions transforming the symmetrical distributions as: the normal distribution, Laplace distribution, Student's distribution etc.
- Asymmetrical distributions that are often used in applications are for example: Poisson, geometric, negative binomial, binomial for $p \neq q$, F, Chi, gamma, lognormal, Weibull, Pareto, Gumbel etc.
- Some of them have constant coefficient of skewness as exponential (2), Gumbel (1.1395)
- Asymmetrical distributions are used in different fields: hydrology, meteorology, biology, economics.

Measuring kurtosis

- If X is normally distributed the shape of the probability density function depends on the variance: the less the variance, the greater is the maximal value of the function, and its tails are thinner.
- Comparison of random variables with the normally distributed random variable can be made by means of fourth central moments.
- Thus, for two random variables X and Y with the same mean and variance, if $\mu_4(X) > \mu_4(Y)$, the probability density of X has larger peak (is less flat) than the probability density function of Y.

Measuring kurtosis

- In order to eliminate units of measure, the forth moment is standardized by square of variance obtaining the coefficient

$$\beta_2 = \frac{\mu_4}{\sigma^4}$$

usually is called *kurtosis*

- Kurtosis ($\kappa\nu\rho\tau\acute{o}\varsigma$) in Greek means convexity, roundness, curvature.
- The value β_2 is also known as the *Second Pearson's coefficient*, or *Pearson's kurtosis*.

Measuring kurtosis

- As for normal distribution $\beta_2 = 3$, the shape of given distribution may be also measured by excess of kurtosis.
- $\gamma_2 = \beta_2 - 3$ known as *Fisher's kurtosis*. Often for both of these coefficients the term kurtosis is used.
- From the inequality $\beta_2 \geq 1 + \beta_1$, it follows that always $\gamma_2 \geq -2$, and there is no upper limit for this coefficient.
- If the probability density function has one mode, then $\gamma_2 \geq -\frac{186}{125}$.
- If it is symmetrical with one mode, then $\gamma_2 \geq -\frac{6}{5}$.

Measuring kurtosis

Pearson:

- Random variables with $\gamma_2 < 0$ - platykurtic. In Greek $\pi\lambda\alpha\tau\acute{\nu}\varsigma$ means wide.
- Random variables with $\gamma_2 > 0$ - leptokurtic. In Greek $\lambda\varepsilon\pi\tau\acute{o}\varsigma$ means narrow.
- Random variables with $\gamma_2 = 0$ - mesokurtic.

Measuring kurtosis

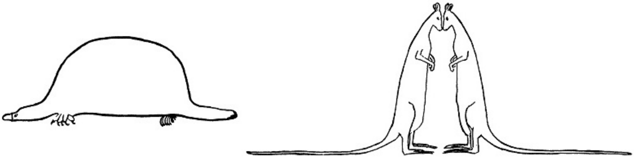
For Gosset (1927) random variables with $\gamma_2 > 0$ have long tails and those with $\gamma_2 < 0$ have short tails.

Figure: William Sealy Gosset, 1876-1937

Measuring kurtosis

Errors of Routine Analysis
Author(s): Student
Source: *Biometrika*, Vol. 19, No. 1/2 (Jul., 1927), pp. 151-164

* In case any of my readers may be unfamiliar with the term "kurtosis" we may define mesokurtic as "having β_2 equal to 3," while platykurtic curves have $\beta_2 < 3$ and leptokurtic > 3. The important property which follows from this is that platykurtic curves have shorter "tails" than the



normal curve of error and leptokurtic longer "tails." I myself bear in mind the meaning of the words by the above *memoria technica*, where the first figure represents platypus, and the second kangaroos, noted for "lepping," though, perhaps, with equal reason they should be hares!

Figure: Gosset memoria technica for kurtosis

Measuring kurtosis

It is often thought that leptocurtic curves were more sharply peaked and platykurtic curves more flat-topped than the normal curve. The next examples show that it may be, but it is not necessarily true.

Example 6. *Let X be a random variable with the probability density function*

$$g_X(x) = \begin{cases} 0, & |x| \geq 1 \\ x+1, & -1 < x < 0 \\ 1-x, & 0 < x < 1 \end{cases}$$

(the triangle or Simpson's distribution). Then

$$E(X) = 0,\ E(X^2) = 2\int_0^1 x^2(1-x)dx = \frac{1}{6},\ E(X^4) = 2\int_0^1 x^4(1-x)dx,$$

and

$$\sigma^2(X) = \frac{1}{6}, \quad \mu_4(X) = \frac{1}{15}, \quad \gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{36}{15} - 3 = -0.6.$$

Measuring kurtosis

- The probability density functions g_X and the corresponding density function of normal distribution with the same mean and variance are given in the figure. Since it is equal to 0 for $|x| > 1$, it has thinner tails than normal distribution but it is more flat than the normal distribution.

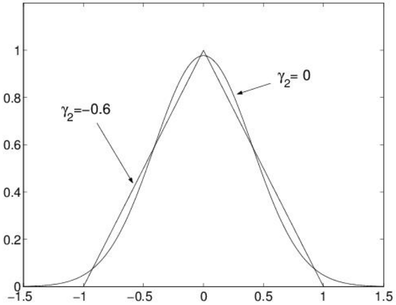


Figure: Density of distribution in Example 6

Example 7. (Kaplansky, 1943) *Let X and Y be random variables with probability density functions*

$$g_X(x) = \frac{1}{3\sqrt{\pi}}\left(x^4 + \frac{9}{4}\right)e^{-x^2}, \quad g_Y(x) = \frac{3\sqrt{3}}{16\sqrt{\pi}}(x^2+2)e^{-3x^2/4}.$$

Both distributions are symmetric with mean 0 and variance 1, while $EX = EY = 0$, $\mu_2(X) = \mu_2(Y) = 1$, $\mu_4(X) = 2.75$ i $\mu_4(Y) = 2.6667$, so their tails are below the tail of a normal distribution, but their maximal values are 0.423 and 0.366 respectively and are different from the maximal value (0.399) of a corresponding normal distribution $N(0,1)$.

Measuring kurtosis

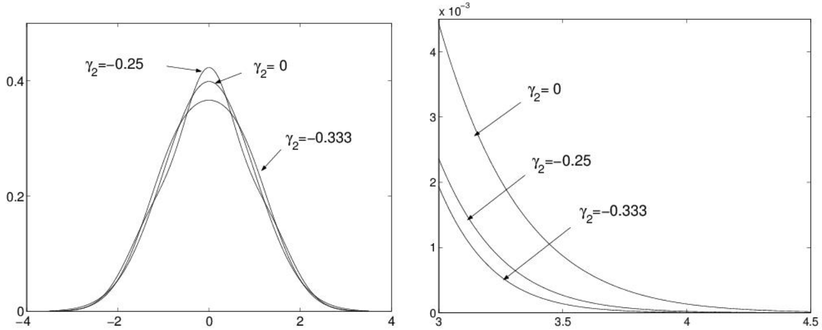


Figure: Standardized densities with $\gamma_2 = -0.25$ and $\gamma_2 = -0.333$ and standard normal distribution (left) and tails of this distributions (right)

Example 8. (Kaplansky, 1943) *Let X and Y be random variables with probability density functions*

$$g_X(x) = \frac{1}{6\sqrt{\pi}}(4e^{-x^2} + e^{-x^2/4})$$

$$g_Y(x) = \frac{3}{2\sqrt{2\pi}}e^{-x^2/2} - \frac{1}{6\sqrt{\pi}}\left(x^4 + \frac{9}{4}\right)e^{-x^2}.$$

These distribution are symmetric with $EX = EY = 0$, $\mu_2(X) = \mu_2(Y) = 1$, $\mu_4(X) = 4.5$ i $\mu_4(Y) = 3.125$ and they both have tails above the tail of a corresponding normal distribution $N(0,1)$, while their maximal values are 0.47 and 0.387.

Measuring kurtosis

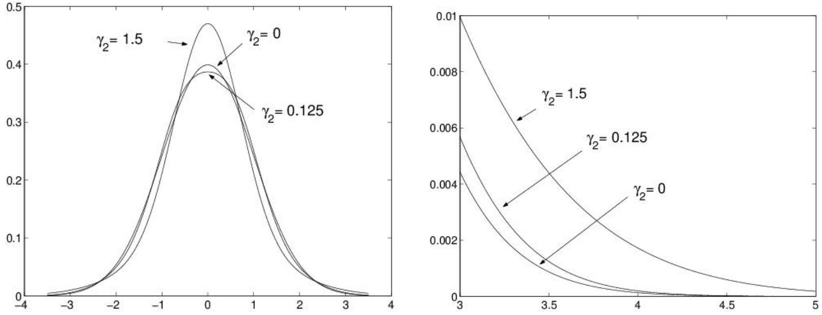


Figure: Standardized densities with $\gamma_2 = 1.5$ and $\gamma_2 = 0.125$ and standard normal distribution (left) and tails of this distributions (right)

Measuring kurtosis

- After these examples we can conclude that the values of the coefficient γ_2 cannot describe the shape of one distribution exactly in comparison with the corresponding normal distribution, having the same mean and variance.
- The kurtosis for asymmetric distributions is always greater than the kurtosis for the normal distribution (Hopkins & Weaks, 1990). There are other measures of kurtosis often used nowadays, some of them eliminate the effects of asymmetry (Blest, 2003).

Measuring kurtosis

- It is wrong to assume any major dependence between the coefficient of kurtosis and the shape of a distribution.
- It is also unacceptable to conclude that small values of the coefficient of kurtosis imply that the variance is big, since there are distributions with the same coefficient of kurtosis and different variances, as well as those with the same variance but different coefficients of kurtosis.

Measuring kurtosis

- In the teaching there is it is often a misunderstanding of the shape of the Student's distribution. It is theoretical fact that with increasing degrees of freedom Student distribution converges to normal distribution with mean value 0 and variance 1. This result is usually illustrated by a graph similar to the next graph:

Measuring kurtosis

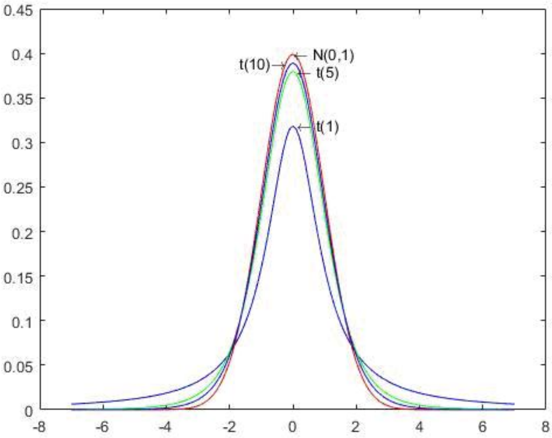


Figure: Densities of standard normal and t-distributions

Measuring kurtosis

- Based on the graph, students conclude that Student's distribution is a platykurtic distribution as a peak will be flatter compared to normal distribution. When comparing the shape of the Student's and the normal distribution, one should present the normal and Student's distribution at the same time with the same variability.
- As standard deviation of Student distribution with degrees of freedom is $\sigma = \sqrt{\nu/(\nu-2)}$, for $\nu = 5$, $\sigma = \sqrt{5/3}$. If we present $t(5)$ and normal distribution with same standard deviation on the same graph it can be seen that Student's distribution is more peaked and has a heavier tails compared to normal distribution. Student distribution $t(5)$ crosses normal distribution twice.

Measuring kurtosis

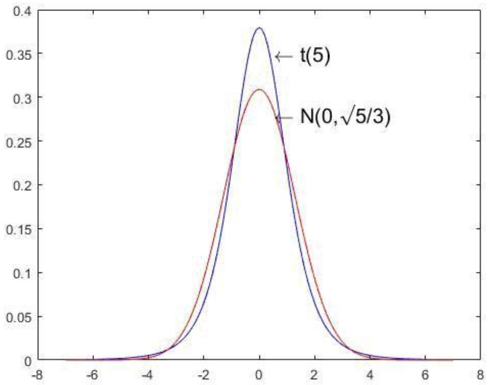


Figure: Densities of standard normal and t5 distribution

Measuring kurtosis

In order to simultaneously present the kurtosis of several distributions correctly, all distributions should have the same variance. The examples of platykurtic, mesokurtic and leptokurtic distributions with zero mean and standard deviation 1 are uniform distribution defined on interval $[-\sqrt{3}, \sqrt{3}]$, normal distribution with mean 0 and standard deviation 1 and two parameters logistic distribution with parameters $a = 0$ and $b = \sqrt{3/\pi^2}$.

Measuring kurtosis

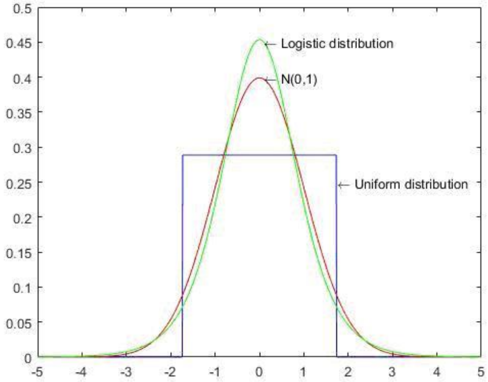


Figure: Densities of uniform, standard normal and logistic distributions

References

- Djorić, D., Nikolić-Djorić, E., Jevremović, V., Mališić, J.(2009) On measuring skewness and kurtosis, *Qual Quant*, 43 pp. 481-493.
- Blest, D. C. (2003) A new measure of kurtosis adjusted for skewness, *Australian & New Zealand Journal of Statistics*, 45, pp. 175-179.
- Cramer, H. (1957) *Mathematical methods of statistics,* Princeton University Press, Seventh Printing, Princeton.
- Darlington, R. B. (1970) Is kurtosis really 'peakedness', *The American Statistician,* 24, pp. 19-22.
- De Carlo, L. T. (1997) On the meaning and use of kurtosis, *Psychological Methods,* 2, pp. 292-307.
- Fisher, R. A. (1930) The moments of the distribution for normal samples of measures of departures from normality, *Proceedings of the Royal Society, London, Series A,* 130, pp. 16-28.

References

Gosset, W. S. "Student" (1927) Errors of Routine Analysis, *Biometrika*, 19, pp. 151-164.

Kaplansky, I. (1945) A common error concerning kurtosis, *Journal of the American Statistical Association*, 40, pp. 259.

Kendall, M. G. and Stuart, A. (1958) *The Advanced Theory of Statistics*, Vol.1, Charles Griffin, London.

Pearson, K. (1895) Contribution to the mathematical theory of evoluition, II: Skew variation in homogenous material, *Philosophical Transactions of the Royal Society of London*, 186, pp. 343-414.

Stoyanov, J. M. (1987) *Counterexamples in Probability*, John Wiley & Sons, New York.

THANK YOU!