Multimodal distribution

In statistics, a **Multimodal distribution** is a probability distribution with two different modes, which may also be referred to as a bimodal distribution. These appear as distinct peaks (local maxima) in the probability density function, as shown in Figures 1 and 2. Categorical, continuous, and discrete data can all form bimodal distributions.

More generally, a **multimodal distribution** is a probability distribution with two or more modes, as illustrated in Figure 3.
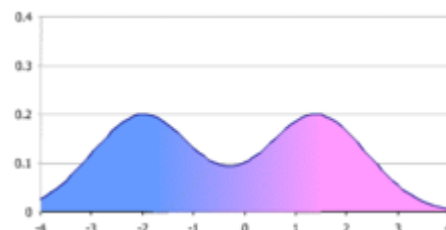

Figure 1. A simple bimodal distribution, in this case a mixture of two normal distributions with the same variance but different means. The figure shows the probability density function (p.d.f.), which is an equally-weighted average of the bell-shaped p.d.f.s of the two normal distributions. If the weights were not equal, the resulting distribution could still be bimodal but with peaks of different heights.
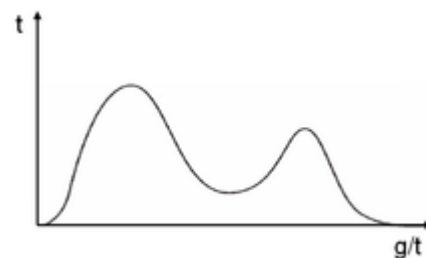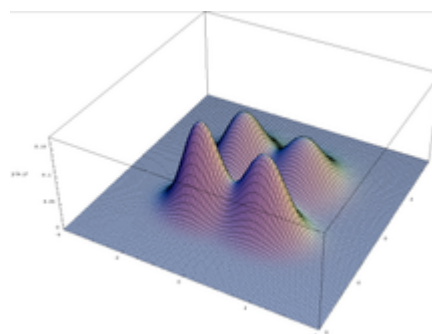

Figure 2. A bimodal distribution.

Figure 3. A bivariate, multimodal distribution

Contents

Terminology

When the two modes are unequal the larger mode is known as the major mode and the other as the minor mode. The least frequent value between the modes is known as the antimode. The difference between the major and minor modes is known as the amplitude. In time series the major mode is called the acrophase and the antimode the batiphase.

Galtung's classification

Galtung introduced a classification system (AJUS) for distributions:[1]

- A: unimodal distribution – peak in the middle
- J: unimodal – peak at either end
- U: bimodal – peaks at both ends
- S: bimodal or multimodal – multiple peaks

This classification has since been modified slightly:

- J: (modified) – peak on right
- L: unimodal – peak on left
- F: no peak (flat)

Under this classification bimodal distributions are classified as type S or U.

Examples

Bimodal distributions occur both in mathematics and in the natural sciences.

Probability distributions

Important bimodal distributions include the arcsine distribution and the beta distribution. Others include the U-quadratic distribution.

The ratio of two normal distributions is also bimodally distributed. Let

$$R = \frac{a + x}{b + y}$$

where a and b are constant and x and y are distributed as normal variables with a mean of 0 and a standard deviation of 1. R has a known density that can be expressed as a confluent hypergeometric function.[2]

The distribution of the reciprocal of a *t* distributed random variable is bimodal when the degrees of freedom are more than one. Similarly the reciprocal of a normally distributed variable is also bimodally distributed.

A *t* statistic generated from data set drawn from a Cauchy distribution is bimodal.[3]

Occurrences in nature

Examples of variables with bimodal distributions include the time between eruptions of certain geysers, the color of galaxies, the size of worker weaver ants, the age of incidence of Hodgkin's lymphoma, the speed of inactivation of the drug isoniazid in US adults, the absolute magnitude of novae, and the circadian activity patterns of those crepuscular animals that are active both in morning and evening twilight. In fishery science multimodal length distributions reflect the different year classes and can thus be used for age distribution- and growth estimates of the fish population.[4] Sediments are usually distributed in a bimodal fashion. Bimodal distributions are also seen in traffic analysis, where traffic peaks in during the AM rush hour and then again in the PM rush hour. This phenomenon is also seen in daily water distribution, as water demand, in the form of showers, cooking, and toilet use, generally peak in the morning and evening periods.

Econometrics

In econometric models, the parameters may be bimodally distributed.[5]

Origins

Mathematical

A bimodal distribution most commonly arises as a mixture of two different unimodal distributions (i.e. distributions having only one mode). In other words, the bimodally distributed random variable X is defined as Y with probability α or Z with probability $(1 - \alpha)$, where Y and Z are unimodal random variables and $0 < \alpha < 1$ is a mixture coefficient.

Mixtures with two distinct components need not be bimodal and two component mixtures of unimodal component densities can have more than two modes. There is no immediate connection between the number of components in a mixture and the number of modes of the resulting density.

Particular distributions

Bimodal distributions, despite their frequent occurrence in data sets, have only rarely been studied. This may be because of the difficulties in estimating their parameters either with frequentist or Bayesian methods. Among those that have been studied are

- Bimodal exponential distribution.[6]
- Alpha-skew-normal distribution.[7]
- Bimodal skew-symmetric normal distribution.[8]
- A mixture of Conway-Maxwell-Poisson distributions has been fitted to bimodal count data.[9]

Bimodality also naturally arises in the cusp catastrophe distribution.

Biology

In biology five factors are known to contribute to bimodal distributions of population sizes:

- the initial distribution of individual sizes
- the distribution of growth rates among the individuals
- the size and time dependence of the growth rate of each individual
- mortality rates that may affect each size class differently
- the DNA methylation in human and mouse genome.

The bimodal distribution of sizes of weaver ant workers arises due to existence of two distinct classes of workers, namely major workers and minor workers.[10]

The distribution of fitness effects of mutations for both whole genomes[11][12] and individual genes[13] is also frequently found to be bimodal with most mutations being either neutral or lethal with relatively few having intermediate effect.

General properties

A mixture of two unimodal distributions with differing means is not necessarily bimodal. The combined distribution of heights of men and women is sometimes used as an example of a bimodal distribution, but in fact the difference in mean heights of men and women is too small relative to their standard deviations to produce bimodality.[14]

Bimodal distributions have the peculiar property that – unlike the unimodal distributions – the mean may be a more robust sample estimator than the median.[15] This is clearly the case when the distribution is U shaped like the arcsine distribution. It may not be true when the distribution has one or more long tails.

Moments of mixtures

Let

$$f(x) = pg_1(x) + (1-p)g_2(x)$$

where g_i is a probability distribution and p is the mixing parameter.

The moments of $f(x)$ are[16]

$$\mu = p\mu_1 + (1-p)\mu_2$$

$$\nu_2 = p[\sigma_1^2 + \delta_1^2] + (1-p)[\sigma_2^2 + \delta_2^2]$$

$$\nu_3 = p[S_1\sigma_1^3 + 3\delta_1\sigma_1^2 + \delta_1^3] + (1-p)[S_2\sigma_2^3 + 3\delta_2\sigma_2^2 + \delta_2^3]$$

$$\nu_4 = p[K_1\sigma_1^4 + 4S_1\delta_1\sigma_1^3 + 6\delta_1^2\sigma_1^2 + \delta_1^4] + (1-p)[K_2\sigma_2^4 + 4S_2\delta_2\sigma_2^3 + 6\delta_2^2\sigma_2^2 + \delta_2^4]$$

where

$$\mu = \int xf(x)\,dx$$

$$\delta_i = \mu_i - \mu$$

$$\nu_r = \int (x - \mu)^r f(x)\, dx$$

and S_i and K_i are the skewness and kurtosis of the ith distribution.

Mixture of two normal distributions

It is not uncommon to encounter situations where an investigator believes that the data comes from a mixture of two normal distributions. Because of this, this mixture has been studied in some detail.[17]

A mixture of two normal distributions has five parameters to estimate: the two means, the two variances and the mixing parameter. A mixture of two normal distributions with equal standard deviations is bimodal only if their means differ by at least twice the common standard deviation.[14] Estimates of the parameters is simplified if the variances can be assumed to be equal (the homoscedastic case).

If the means of the two normal distributions are equal, then the combined distribution is unimodal. Conditions for unimodality of the combined distribution were derived by Eisenberger.[18] Necessary and sufficient conditions for a mixture of normal distributions to be bimodal have been identified by Ray and Lindsay.[19]

A mixture of two approximately equal mass normal distributions has a negative kurtosis since the two modes on either side of the center of mass effectively reduces the tails of the distribution.

A mixture of two normal distributions with highly unequal mass has a positive kurtosis since the smaller distribution lengthens the tail of the more dominant normal distribution.

Mixtures of other distributions require additional parameters to be estimated.

Tests for unimodality

- The mixture is unimodal if and only if[20]

$$d \leq 1$$

or

$$|\log(1 - p) - \log(p)| \geq 2\log(d - \sqrt{d^2 - 1}) + 2d\sqrt{d^2 - 1},$$

where p is the mixing parameter and

$$d = \frac{|\mu_1 - \mu_2|}{2\sqrt{\sigma_1 \sigma_2}},$$

and where μ_1 and μ_2 are the means of the two normal distributions and σ_1 and σ_2 are their standard deviations.

- The following test for the case $p = 1/2$ was described by Schilling *et al*.[14] Let

$$r = \frac{\sigma_1^2}{\sigma_2^2}.$$

The separation factor (S) is

$$S = \frac{\sqrt{-2 + 3r + 3r^2 - 2r^3 + 2(1 - r + r^2)^{1.5}}}{\sqrt{r}(1 + \sqrt{r})}.$$

If the variances are equal then $S = 1$. The mixture density is unimodal if and only if

$$|\mu_1 - \mu_2| < S|\sigma_1 + \sigma_2|.$$

- A sufficient condition for unimodality is[21]

$$|\mu_1 - \mu_2| \leq 2\min(\sigma_1, \sigma_2).$$

- If the two normal distributions have equal standard deviations σ, a sufficient condition for unimodality is[21]

$$|\mu_1 - \mu_2| \leq 2\sigma\sqrt{1 + \frac{|\log p - \ln(1 - p)|}{2}}.$$

Summary statistics

Bimodal distributions are a commonly used example of how summary statistics such as the mean, median, and standard deviation can be deceptive when used on an arbitrary distribution. For example, in the distribution in Figure 1, the mean and median would be about zero, even though zero is not a typical value. The standard deviation is also larger than deviation of each normal distribution.

Although several have been suggested, there is no presently generally agreed summary statistic (or set of statistics) to quantify the parameters of a general bimodal distribution. For a mixture of two normal distributions the means and standard deviations along with the mixing parameter (the weight for the combination) are usually used – a total of five parameters.

Ashman's D

A statistic that may be useful is Ashman's D:[22]

$$D = (2^{\frac{1}{2}})\frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)}}$$

where μ_1, μ_2 are the means and σ_1 σ_2 are the standard deviations.

For a mixture of two normal distributions $D > 2$ is required for a clean separation of the distributions.

van der Eijk's A

This measure is a weighted average of the degree of agreement the frequency distribution.[23] A ranges from -1 (perfect bimodality) to +1 (perfect unimodality). It is defined as

$$A = U(1 - \frac{S-1}{K-1})$$

where U is the unimodality of the distribution, S the number of categories that have nonzero frequencies and K the total number of categories.

The value of U is 1 if the distribution has any of the three following characteristics:

- all responses are in a single category
- the responses are evenly distributed among all the categories
- the responses are evenly distributed among two or more contiguous categories, with the other categories with zero responses

With distributions other than these the data must be divided into 'layers'. Within a layer the responses are either equal or zero. The categories do not have to be contiguous. A value for A for each layer (A_i) is calculated and a weighted average for the distribution is determined. The weights (w_i) for each layer are the number of responses in that layer. In symbols

$$A_{overall} = \sum w_i A_i$$

A <u>uniform distribution</u> has $A = 0$: when all the responses fall into one category $A = +1$.

One theoretical problem with this index is that it assumes that the intervals are equally spaced. This may limit its applicability.

Bimodal separation

This index assumes that the distribution is a mixture of two normal distributions with means (μ_1 and μ_2) and standard deviations (σ_1 and σ_2):[24]

$$S = \frac{\mu_1 - \mu_2}{2(\sigma_1 + \sigma_2)}$$

Bimodality coefficient

Sarle's bimodality coefficient b is[25]

$$\beta = \frac{\gamma^2 + 1}{\kappa}$$

where γ is the <u>skewness</u> and κ is the <u>kurtosis</u>. The kurtosis is here defined to be the standardised fourth moment around the mean. The value of b lies between 0 and 1.[26] The logic behind this coefficient is that a bimodal distribution with light tails will have very low kurtosis, an asymmetric character, or both – all of which increase this coefficient.

The formula for a finite sample is[27]

$$b = \frac{g^2 + 1}{k + \frac{3(n-1)^2}{(n-2)(n-3)}}$$

where n is the number of items in the sample, g is the <u>sample skewness</u> and k is the sample <u>excess kurtosis</u>.

The value of b for the <u>uniform distribution</u> is 5/9. This is also its value for the <u>exponential distribution</u>. Values greater than 5/9 may indicate a bimodal or multimodal distribution, though corresponding values can also result for heavily skewed unimodal distributions.[28] The maximum value (1.0) is reached only by a <u>Bernoulli distribution</u> with only two distinct values or the sum of two different <u>Dirac delta functions</u> (a bi-delta distribution).

The distribution of this statistic is unknown. It is related to a statistic proposed earlier by Pearson – the difference between the kurtosis and the square of the skewness (*vide infra*).

Bimodality amplitude

This is defined as[24]

$$A_B = \frac{A_1 - A_{an}}{A_1}$$

where A_1 is the amplitude of the smaller peak and A_{an} is the amplitude of the antimode.

A_{B} is always < 1. Larger values indicate more distinct peaks.

Bimodal ratio

This is the ratio of the left and right peaks.[24] Mathematically

$$R = \frac{A_r}{A_l}$$

where A_{l} and A_{r} are the amplitudes of the left and right peaks respectively.

Bimodality parameter

This parameter (B) is due to Wilcock.[29]

$$B = (\frac{A_r}{A_l})^{0.5} \sum P_i$$

where A_{l} and A_{r} are the amplitudes of the left and right peaks respectively and P_i is the logarithm taken to the base 2 of the proportion of the distribution in the ith interval. The maximal value of the ΣP is 1 but the value of B may be greater than this.

To use this index, the log of the values are taken. The data is then divided into interval of width Φ whose value is log 2. The width of the peaks are taken to be four times 1/4Φ centered on their maximum values.

Bimodality indices

Wang's index

The bimodality index proposed by Wang *et al* assumes that the distribution is a sum of two normal distributions with equal variances but differing means.[30] It is defined as follows:

$$\delta = \frac{|\mu_1 - \mu_2|}{\sigma}$$

where μ_1, μ_2 are the means and σ is the common standard deviation.

$$BI = \delta\sqrt{p(1-p)}$$

where p is the mixing parameter.

Sturrock's index

A different bimodality index has been proposed by Sturrock.[31]

This index (B) is defined as

$$B = \frac{1}{N}\left[\left(\sum_1^N \cos(2\pi m\gamma)\right)^2 + \left(\sum_1^N \sin(2\pi m\gamma)\right)^2\right]$$

When $m = 2$ and γ is uniformly distributed, B is exponentially distributed.[32]

This statistic is a form of underlined periodogram. It suffers from the usual problems of estimation and spectral leakage common to this form of statistic.

de Michele and Accatino's index

Another bimodality index has been proposed by de Michele and Accatino.[33] Their index (B) is

$$B = |\mu - \mu_M|$$

where μ is the arithmetic mean of the sample and

$$\mu_M = \frac{\sum_{i=1}^L m_i x_i}{\sum_{i=1}^L m_i}$$

where m_i is number of data points in the ith bin, x_i is the center of the ith bin and L is the number of bins.

The authors suggested a cut off value of 0.1 for B to distinguish between a bimodal ($B > 0.1$) and unimodal ($B < 0.1$) distribution. No statistical justification was offered for this value.

Sambrook Smith's index

A further index (B) has been proposed by Sambrook Smith *et al*[34]

$$B = |\phi_2 - \phi_1|\frac{p_2}{p_1}$$

where p_1 and p_2 are the proportion contained in the primary (that with the greater amplitude) and secondary (that with the lesser amplitude) mode and φ_1 and φ_2 are the φ-sizes of the primary and secondary mode. The φ-size is defined as minus one times the log of the data size taken to the base 2. This transformation is commonly

used in the study of sediments.

The authors recommended a cut off value of 1.5 with B being greater than 1.5 for a bimodal distribution and less than 1.5 for a unimodal distribution. No statistical justification for this value was given.

Chaudhuri and Agrawal index

Another bimodality parameter has been proposed by Chaudhuri and Agrawal.[35] This parameter requires knowledge of the variances of the two subpopulations that make up the bimodal distribution. It is defined as

$$k = \frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{m \sigma^2}$$

where n_i is the number of data points in the ith subpopulation, σ_i^2 is the variance of the ith subpopulation, m is the total size of the sample and σ^2 is the sample variance.

It is a weighted average of the variance. The authors suggest that this parameter can be used as the optimisation target to divide a sample into two subpopulations. No statistical justification for this suggestion was given.

Statistical tests

A number of tests are available to determine if a data set is distributed in a bimodal (or multimodal) fashion.

Graphical methods

In the study of sediments, particle size is frequently bimodal. Empirically, it has been found useful to plot the frequency against the log(size) of the particles.[36][37] This usually gives a clear separation of the particles into a bimodal distribution. In geological applications the logarithm is normally taken to the base 2. The log transformed values are referred to as phi (Φ) units. This system is known as the Krumbein (or phi) scale.

An alternative method is to plot the log of the particle size against the cumulative frequency. This graph will usually consist two reasonably straight lines with a connecting line corresponding to the antimode.

Statistics

Approximate values for several statistics can be derived from the graphic plots.[36]

$$Mean = \frac{\phi_{16} + \phi_{50} + \phi_{84}}{3}$$

$$StdDev = \frac{\phi_{84} - \phi_{16}}{4} + \frac{\phi_{95} - \phi_{5}}{6.6}$$

$$Skew = \frac{\phi_{84} + \phi_{16} - 2\phi_{50}}{2(\phi_{84} - \phi_{16})} + \frac{\phi_{95} + \phi_{5} - 2\phi_{50}}{2(\phi_{95} - \phi_{5})}$$

$$Kurt = \frac{\phi_{95} - \phi_{5}}{2.44(\phi_{75} - \phi_{25})}$$

where *Mean* is the mean, *StdDev* is the standard deviation, *Skew* is the skewness, *Kurt* is the kurtosis and φ_x is the value of the variate φ at the xth percentage of the distribution.

Unimodal vs. bimodal distribution

Pearson in 1894 was the first to devise a procedure to test whether a distribution could be resolved into two normal distributions.[38] This method required the solution of a ninth order polynomial. In a subsequent paper Pearson reported that for any distribution skewness^2 + 1 < kurtosis.[26] Later Pearson showed that[39]

$$b_2 - b_1 \geq 1$$

where b_2 is the kurtosis and b_1 is the square of the skewness. Equality holds only for the two point Bernoulli distribution or the sum of two different Dirac delta functions. These are the most extreme cases of bimodality possible. The kurtosis in both these cases is 1. Since they are both symmetrical their skewness is 0 and the difference is 1.

Baker proposed a transformation to convert a bimodal to a unimodal distribution.[40]

Several tests of unimodality versus bimodality have been proposed: Haldane suggested one based on second central differences.[41] Larkin later introduced a test based on the F test;[42] Benett created one based on Fisher's G test.[43] Tokeshi has proposed a fourth test.[44][45] A test based on a likelihood ratio has been proposed by Holzmann and Vollmer.[20]

A method based on the score and Wald tests has been proposed.[46] This method can distinguish between unimodal and bimodal distributions when the underlying distributions are known.

Antimode tests

Statistical tests for the antimode are known.[47]

Otsu's method

Otsu's method is commonly employed in computer graphics to determine the optimal separation between two distributions.

General tests

To test if a distribution is other than unimodal, several additional tests have been devised: the bandwidth test,[48] the dip test,[49] the excess mass test,[50] the MAP test,[51] the mode existence test,[52] the runt test,[53][54] the span test,[55] and the saddle test.

An implementation of the dip test is available for the R programming language.[56] The p-values for the dip statistic values range between 0 and 1. P-values less than 0.05 indicate significant multimodality and p-values greater than 0.05 but less than 0.10 suggest multimodality with marginal significance [57].

Silverman's test

Silverman introduced a bootstrap method for the number of modes.[48] The test uses a fixed bandwidth which reduces the power of the test and its interpretability. Under smoothed densities may have an excessive number of modes whose count during bootstrapping is unstable.

Bajgier-Aggarwal test

Bajgier and Aggarwal have proposed a test based on the kurtosis of the distribution.[58]

Special cases

Additional tests are available for a number of special cases:

Mixture of two normal distributions

A study of a mixture density of two normal distributions data found that separation into the two normal distributions was difficult unless the means were separated by 4–6 standard deviations.[59]

In astronomy the Kernel Mean Matching algorithm is used to decide if a data set belongs to a single normal distribution or to a mixture of two normal distributions.

Beta-normal distribution

This distribution is bimodal for certain values of is parameters. A test for these values has been described.[60]

Parameter estimation and fitting curves

Assuming that the distribution is known to be bimodal or has been shown to be bimodal by one or more of the tests above, it is frequently desirable to fit a curve to the data. This may be difficult.

Bayesian methods may be useful in difficult cases.

Software

Two normal distributions

A package for R is available for testing for bimodality.[61] This package assumes that the data are distributed as a sum of two normal distributions. If this assumption is not correct the results may not be reliable. It also includes functions for fitting a sum of two normal distributions to the data.

Assuming that the distribution is a mixture of two normal distributions then the expectation-maximization algorithm may be used to determine the parameters. Several programmes are available for this including Cluster,[62] and the R package nor1mix.[63]

Other distributions

The mixtools package available for R can test for and estimate the parameters of a number of different distributions.[64] A package for a mixture of two right-tailed gamma distributions is available.[65]

Several other packages for R are available to fit mixture models; these include flexmix,[66] mcclust,[67], agrmt,[68] and mixdist.[69]

The statistical programming language SAS can also fit a variety of mixed distributions with the PROC FREQ procedure.

See also

- Overdispersion

References

1. Galtung, J. (1969). *Theory and methods of social research*. Oslo: Universitetsforlaget. ISBN 0-04-300017-7.
2. Fieller E (1932). "The distribution of the index in a normal bivariate population". *Biometrika*. **24** (3–4): 428–440. doi:10.1093/biomet/24.3-4.428 (https://doi.org/10.1093%2Fbiomet%2F24.3-4.428).
3. Fiorio, CV; HajivassILiou, VA; Phillips, PCB (2010). "Bimodal t-ratios: the impact of thick tails on inference" (https://ink.library.smu.edu.sg/soe_research/1817). *The Econometrics Journal*. **13** (2): 271–289. doi:10.1111/j.1368-423X.2010.00315.x (https://doi.org/10.1111%2Fj.1368-423X.2010.00315.x). S2CID 363740 (https://api.semanticscholar.org/CorpusID:363740).
4. Introduction to tropical fish stock assessment (http://www.fao.org/docrep/W5449E/w5449e05.htm.%7CFAO:)
5. Phillips, P. C. B. (2006). "A remark on bimodality and weak instrumentation in structural equation estimation" (http://cowles.yale.edu/sites/default/files/files/pub/d15/d1540.pdf) (PDF). *Econometric Theory*. **22** (5): 947–960. doi:10.1017/S0266466606060439 (https://doi.org/10.1017%2FS0266466606060439). S2CID 16775883 (https://api.semanticscholar.org/CorpusID:16775883).
6. Hassan, MY; Hijazi, RH (2010). "A bimodal exponential power distribution". *Pakistan Journal of Statistics*. **26** (2): 379–396.
7. Elal-Olivero, D (2010). "Alpha-skew-normal distribution" (https://doi.org/10.4067%2Fs0716-09172010000300006). *Proyecciones Journal of Mathematics*. **29** (3): 224–240. doi:10.4067/s0716-09172010000300006 (https://doi.org/10.4067%2Fs0716-09172010000300006).
8. Hassan, M. Y.; El-Bassiouni, M. Y. (2016). "Bimodal skew-symmetric normal distribution". *Communications in Statistics - Theory and Methods*. **45** (5): 1527–1541. doi:10.1080/03610926.2014.882950 (https://doi.org/10.1080%2F03610926.2014.882950). S2CID 124087015 (https://api.semanticscholar.org/CorpusID:124087015).
9. Bosea, S.; Shmuelib, G.; Sura, P.; Dubey, P. (2013). "Fitting Com-Poisson mixtures to bimodal count data" (https://www.galitshmueli.com/system/files/ICIOMS%202013%20Malaysia%20Paper%20ID%2028.pdf) (PDF). *Proceedings of the 2013 International Conference on Information, Operations Management and Statistics (ICIOMS2013), Kuala Lumpur, Malaysia*. pp. 1–8.
10. Weber, NA (1946). "Dimorphism in the African *Oecophylla* worker and an anomaly (Hym.: Formicidae)" (http://antbase.org/ants/publications/10434/10434.pdf) (PDF). *Annals of the Entomological Society of America*. **39**: 7–10. doi:10.1093/aesa/39.1.7 (https://doi.org/10.1093%2Faesa%2F39.1.7).
11. Sanjuán, R (Jun 27, 2010). "Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2880115). *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. **365** (1548): 1975–82. doi:10.1098/rstb.2010.0063 (https://doi.org/10.1098%2Frstb.2010.0063). PMC 2880115 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2880115). PMID 20478892 (https://pubmed.ncbi.nlm.nih.gov/20478892).
12. Eyre-Walker, A; Keightley, PD (Aug 2007). "The distribution of fitness effects of new mutations". *Nature Reviews Genetics*. **8** (8): 610–8. doi:10.1038/nrg2146 (https://doi.org/10.1038%2Fnrg2146). PMID 17637733 (https://pubmed.ncbi.nlm.nih.gov/17637733). S2CID 10868777 (https://api.semanticscholar.org/CorpusID:10868777).
13. Hietpas, RT; Jensen, JD; Bolon, DN (May 10, 2011). "Experimental illumination of a fitness landscape" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3093508). *Proceedings of the National Academy of Sciences of the United States of America*. **108** (19): 7896–901. Bibcode:2011PNAS..108.7896H (https://ui.adsabs.harvard.edu/abs/2011PNAS..108.7896H). doi:10.1073/pnas.1016024108 (https://doi.org/10.1073%2Fpnas.1016024108). PMC 3093508 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3093508). PMID 21464309 (https://pubmed.ncbi.nlm.nih.gov/21464309).

14. Schilling, Mark F.; Watkins, Ann E.; Watkins, William (2002). "Is Human Height Bimodal?". *The American Statistician*. **56** (3): 223–229. doi:10.1198/00031300265 (https://doi.org/10.1198%2F00031300265). S2CID 53495657 (https://api.semanticscholar.org/CorpusID:53495657).

15. Mosteller, F.; Tukey, J. W. (1977). *Data Analysis and Regression: A Second Course in Statistics*. Reading, Mass: Addison-Wesley. ISBN 0-201-04854-X.

16. Kim, T.-H.; White, H. (2003). "On more robust estimation of skewness and kurtosis: Simulation and application to the S & P 500 index" (https://www.cirano.qc.ca/realisations/grandes_conferences/methodes_econometriques/white.pdf) (PDF).

17. Robertson, CA; Fryer, JG (1969). "Some descriptive properties of normal mixtures". *Skandinavisk Aktuarietidskrift*. **69** (3–4): 137–146. doi:10.1080/03461238.1969.10404590 (https://doi.org/10.1080%2F03461238.1969.10404590).

18. Eisenberger, I (1964). "Genesis of bimodal distributions". *Technometrics*. **6** (4): 357–363. doi:10.1080/00401706.1964.10490199 (https://doi.org/10.1080%2F00401706.1964.10490199).

19. Ray, S; Lindsay, BG (2005). "The topography of multivariate normal mixtures". *Annals of Statistics*. **33** (5): 2042–2065. arXiv:math/0602238 (https://arxiv.org/abs/math/0602238). doi:10.1214/009053605000000417 (https://doi.org/10.1214%2F009053605000000417). S2CID 36234163 (https://api.semanticscholar.org/CorpusID:36234163).

20. Holzmann, Hajo; Vollmer, Sebastian (2008). "A likelihood ratio test for bimodality in two-component mixtures with application to regional income distribution in the EU". *AStA Advances in Statistical Analysis*. **2** (1): 57–69. doi:10.1007/s10182-008-0057-2 (https://doi.org/10.1007%2Fs10182-008-0057-2). S2CID 14470055 (https://api.semanticscholar.org/CorpusID:14470055).

21. Behboodian, J (1970). "On the modes of a mixture of two normal distributions". *Technometrics*. **12** (1): 131–139. doi:10.2307/1267357 (https://doi.org/10.2307%2F1267357). JSTOR 1267357 (https://www.jstor.org/stable/1267357).

22. Ashman KM; Bird CM; Zepf SE (1994). "Detecting bimodality in astronomical datasets". *The Astronomical Journal*. **108**: 2348–2361. arXiv:astro-ph/9408030 (https://arxiv.org/abs/astro-ph/9408030). Bibcode:1994AJ....108.2348A (https://ui.adsabs.harvard.edu/abs/1994AJ....108.2348A). doi:10.1086/117248 (https://doi.org/10.1086%2F117248). S2CID 13464256 (https://api.semanticscholar.org/CorpusID:13464256).

23. Van der Eijk, C (2001). "Measuring agreement in ordered rating scales". *Quality & Quantity*. **35** (3): 325–341. doi:10.1023/a:1010374114305 (https://doi.org/10.1023%2Fa%3A1010374114305).

24. Zhang, C; Mapes, BE; Soden, BJ (2003). "Bimodality in tropical water vapour". *Quarterly Journal of the Royal Meteorological Society*. **129** (594): 2847–2866. Bibcode:2003QJRMS.129.2847Z (https://ui.adsabs.harvard.edu/abs/2003QJRMS.129.2847Z). doi:10.1256/qj.02.166 (https://doi.org/10.1256%2Fqj.02.166).

25. Ellison, AM (1987). "Effect of seed dimorphism on the density-dependent dynamics of experimental populations of *Atriplex triangularis* (Chenopodiaceae)". *American Journal of Botany*. **74** (8): 1280–1288. doi:10.2307/2444163 (https://doi.org/10.2307%2F2444163). JSTOR 2444163 (https://www.jstor.org/stable/2444163).

26. Pearson, K (1916). "Mathematical contributions to the theory of evolution, XIX: Second supplement to a memoir on skew variation" (https://doi.org/10.1098%2Frsta.1916.0009). *Philosophical Transactions of the Royal Society A*. **216** (538–548): 429–457. Bibcode:1916RSPTA.216..429P (https://ui.adsabs.harvard.edu/abs/1916RSPTA.216..429P). doi:10.1098/rsta.1916.0009 (https://doi.org/10.1098%2Frsta.1916.0009). JSTOR 91092 (https://www.jstor.org/stable/91092).

27. SAS Institute Inc. (2012). SAS/STAT 12.1 user's guide. Cary, NC: Author.

28. Pfister, R; Schwarz, KA; Janczyk, M.; Dale, R; Freeman, JB (2013). "Good things peak in pairs: A note on the bimodality coefficient" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3791391). *Frontiers in Psychology*. **4**: 700. doi:10.3389/fpsyg.2013.00700 (https://doi.org/10.3389%2Ffpsyg.2013.00700). PMC 3791391 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3791391). PMID 24109465 (https://pubmed.ncbi.nlm.nih.gov/24109465).

29. Wilcock, PR (1993). "The critical shear stress of natural sediments". *Journal of Hydraulic Engineering*. **119** (4): 491–505. doi:10.1061/(asce)0733-9429(1993)119:4(491) (https://doi.org/10.1061%2F%28asce%290733-9429%281993%29119%3A4%28491%29).

30. Wang, J; Wen, S; Symmans, WF; Pusztai, L; Coombes, KR (2009). "The bimodality index: a criterion for discovering and ranking bimodal signatures from cancer gene expression profiling data" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2730180). *Cancer Informatics*. **7**: 199–216. doi:10.4137/CIN.S2846 (https://doi.org/10.4137%2FCIN.S2846). PMC 2730180 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2730180). PMID 19718451 (https://pubmed.ncbi.nlm.nih.gov/19718451).

31. Sturrock, P (2008). "Analysis of bimodality in histograms formed from GALLEX and GNO solar neutrino data". *Solar Physics*. **249** (1): 1–10. arXiv:0711.0216 (https://arxiv.org/abs/0711.0216). Bibcode:2008SoPh..249....1S (https://ui.adsabs.harvard.edu/abs/2008SoPh..249....1S). doi:10.1007/s11207-008-9170-3 (https://doi.org/10.1007%2Fs11207-008-9170-3). S2CID 118389173 (https://api.semanticscholar.org/CorpusID:118389173).

32. Scargle, JD (1982). "Studies in astronomical time series analysis. II – Statistical aspects of spectral analysis of unevenly spaced data". *The Astrophysical Journal*. **263** (1): 835–853. Bibcode:1982ApJ...263..835S (https://ui.adsabs.harvard.edu/abs/1982ApJ...263..835S). doi:10.1086/160554 (https://doi.org/10.1086%2F160554).

33. De Michele, C; Accatino, F (2014). "Tree cover bimodality in savannas and forests emerging from the switching between two fire dynamics" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3963849). *PLOS ONE*. **9** (3): e91195. Bibcode:2014PLoSO...991195D (https://ui.adsabs.harvard.edu/abs/2014PLoSO...991195D). doi:10.1371/journal.pone.0091195 (https://doi.org/10.1371%2Fjournal.pone.0091195). PMC 3963849 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3963849). PMID 24663432 (https://pubmed.ncbi.nlm.nih.gov/24663432).

34. Sambrook Smith, GH; Nicholas, AP; Ferguson, RI (1997). "Measuring and defining bimodal sediments: Problems and implications". *Water Resources Research*. **33** (5): 1179–1185. Bibcode:1997WRR....33.1179S (https://ui.adsabs.harvard.edu/abs/1997WRR....33.1179S). doi:10.1029/97wr00365 (https://doi.org/10.1029%2F97wr00365).

35. Chaudhuri, D; Agrawal, A (2010). "Split-and-merge procedure for image segmentation using bimodality detection approach" (https://doi.org/10.14429%2Fdsj.60.356). *Defence Science Journal*. **60** (3): 290–301. doi:10.14429/dsj.60.356 (https://doi.org/10.14429%2Fdsj.60.356).

36. Folk, RL; Ward, WC (1957). "Brazos River bar: a study in the significance of grain size parameters" (https://doi.pangaea.de/10.1594/PANGAEA.896129). *Journal of Sedimentary Research*. **27** (1): 3–26. Bibcode:1957JSedR..27....3F (https://ui.adsabs.harvard.edu/abs/1957JSedR..27....3F). doi:10.1306/74d70646-2b21-11d7-8648000102c1865d (https://doi.org/10.1306%2F74d70646-2b21-11d7-8648000102c1865d).

37. Dyer, KR (1970). "Grain-size parameters for sandy gravels". *Journal of Sedimentary Research*. **40** (2): 616–620. doi:10.1306/74D71FE6-2B21-11D7-8648000102C1865D (https://doi.org/10.1306%2F74D71FE6-2B21-11D7-8648000102C1865D).

38. Pearson, K (1894). "Contributions to the mathematical theory of evolution: On the dissection of asymmetrical frequency-curves" (https://doi.org/10.1098%2Frsta.1894.0003). *Philosophical Transactions of the Royal Society A*. **185**: 71–90. Bibcode:1894RSPTA.185...71P (https://ui.adsabs.harvard.edu/abs/1894RSPTA.185...71P). doi:10.1098/rsta.1894.0003 (https://doi.org/10.1098%2Frsta.1894.0003).

39. Pearson, K (1929). "Editorial note". *Biometrika*. **21**: 370–375.

40. Baker, GA (1930). "Transformations of bimodal distributions" (https://doi.org/10.1214%2Faoms%2F1177733063). *Annals of Mathematical Statistics*. **1** (4): 334–344. doi:10.1214/aoms/1177733063 (https://doi.org/10.1214%2Faoms%2F1177733063).

41. Haldane, JBS (1951). "Simple tests for bimodality and bitangentiality". *Annals of Eugenics*. **16** (1): 359–364. doi:10.1111/j.1469-1809.1951.tb02488.x (https://doi.org/10.1111%2Fj.1469-1809.1951.tb02488.x). PMID 14953132 (https://pubmed.ncbi.nlm.nih.gov/14953132).

42. Larkin, RP (1979). "An algorithm for assessing bimodality vs. unimodality in a univariate distribution" (https://doi.org/10.3758%2FBF03205709). *Behavior Research Methods & Instrumentation*. **11** (4): 467–468. doi:10.3758/BF03205709 (https://doi.org/10.3758%2FBF032 05709).

43. Bennett, SC (1992). "Sexual dimorphism of *Pteranodon* and other pterosaurs, with comments on cranial crests". *Journal of Vertebrate Paleontology*. **12** (4): 422–434. doi:10.1080/02724634.1992.10011472 (https://doi.org/10.1080%2F02724634.1992.10011472).

44. Tokeshi, M (1992). "Dynamics and distribution in animal communities; theory and analysis". *Researches on Population Ecology*. **34** (2): 249–273. doi:10.1007/bf02514796 (https://doi.org/1 0.1007%2Fbf02514796). S2CID 22912914 (https://api.semanticscholar.org/CorpusID:2291291 4).

45. Barreto, S; Borges, PAV; Guo, Q (2003). "A typing error in Tokeshi's test of bimodality". *Global Ecology and Biogeography*. **12** (2): 173–174. doi:10.1046/j.1466-822x.2003.00018.x (https://do i.org/10.1046%2Fj.1466-822x.2003.00018.x). hdl:10400.3/1408 (https://hdl.handle.net/10400. 3%2F1408).

46. Carolan, AM; Rayner, JCW (2001). "One sample tests for the location of modes of nonnormal data". *Journal of Applied Mathematics and Decision Sciences*. **5** (1): 1–19. CiteSeerX 10.1.1.504.4999 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.504.499 9). doi:10.1155/s1173912601000013 (https://doi.org/10.1155%2Fs1173912601000013).

47. Hartigan, J. A. (2000). "Testing for Antimodes" (https://www.google.com/books/edition/Data_An alysis/WVDmCAAAQBAJ?hl=en&gbpv=1&pg=PA169). In Gaul W; Opitz O; Schader M (eds.). *Data Analysis*. Studies in Classification, Data Analysis, and Knowledge Organization. Springer. pp. 169–181. ISBN 3-540-67731-3.

48. Silverman, B. W. (1981). "Using kernel density estimates to investigate multimodality". *Journal of the Royal Statistical Society, Series B*. **43** (1): 97–99. Bibcode:1981JRSSB..43...97S (https:// ui.adsabs.harvard.edu/abs/1981JRSSB..43...97S). doi:10.1111/j.2517-6161.1981.tb01155.x (ht tps://doi.org/10.1111%2Fj.2517-6161.1981.tb01155.x). JSTOR 2985156 (https://www.jstor.org/s table/2985156).

49. Hartigan, JA; Hartigan, PM (1985). "The dip test of unimodality" (https://doi.org/10.1214%2Fao s%2F1176346577). *Annals of Statistics*. **13** (1): 70–84. doi:10.1214/aos/1176346577 (https://do i.org/10.1214%2Faos%2F1176346577).

50. Mueller, DW; Sawitzki, G (1991). "Excess mass estimates and tests for multimodality". *Journal of the American Statistical Association*. **86** (415): 738–746. doi:10.1080/01621459.1991.10475103 (https://doi.org/10.1080%2F01621459.1991.10475103). JSTOR 2290406 (https://www.jstor.org/stable/2290406).

51. Rozál, GPM Hartigan JA (1994). "The MAP test for multimodality". *Journal of Classification*. **11** (1): 5–36. doi:10.1007/BF01201021 (https://doi.org/10.1007%2FBF01201021). S2CID 118500771 (https://api.semanticscholar.org/CorpusID:118500771).

52. Minnotte, MC (1997). "Nonparametric testing of the existence of modes" (https://doi.org/10.121 4%2Faos%2F1031594735). *Annals of Statistics*. **25** (4): 1646–1660. doi:10.1214/aos/1031594735 (https://doi.org/10.1214%2Faos%2F1031594735).

53. Hartigan, JA; Mohanty, S (1992). "The RUNT test for multimodality". *Journal of Classification*. **9**: 63–70. doi:10.1007/bf02618468 (https://doi.org/10.1007%2Fbf02618468). S2CID 121960832 (https://api.semanticscholar.org/CorpusID:121960832).

54. Andrushkiw RI; Klyushin DD; Petunin YI (2008). "A new test for unimodality". *Theory of Stochastic Processes*. **14** (1): 1–6.

55. Hartigan, J. A. (1988). "The Span Test of Multimodality". In Bock, H. H. (ed.). *Classification and Related Methods of Data Analysis*. Amsterdam: North-Holland. pp. 229–236. ISBN 0-444-70404-3.

56. Ringach, Martin Maechler (originally from Fortran and S.-plus by Dario; NYU.edu) (5 December 2016). "diptest: Hartigan's Dip Test Statistic for Unimodality - Corrected" (https://cran.r-project.or g/web/packages/diptest/index.html) – via R-Packages.

57. Freeman; Dale (2012). "Assessing bimodality to detect the presence of a dual cognitive process" (http://psych.nyu.edu/freemanlab/pubs/2012_BRM.pdf) (PDF). *Behavior Research Methods*. **45** (1): 83–97. doi:10.3758/s13428-012-0225-x (https://doi.org/10.3758%2Fs13428-012-0225-x). PMID 22806703 (https://pubmed.ncbi.nlm.nih.gov/22806703). S2CID 14500508 (https://api.semanticscholar.org/CorpusID:14500508).

58. Bajgier SM; Aggarwal LK (1991). "Powers of goodness-of-fit tests in detecting balanced mixed normal distributions". *Educational and Psychological Measurement*. **51** (2): 253–269. doi:10.1177/0013164491512001 (https://doi.org/10.1177%2F0013164491512001). S2CID 121113601 (https://api.semanticscholar.org/CorpusID:121113601).

59. Jackson, PR; Tucker, GT; Woods, HF (1989). "Testing for bimodality in frequency distributions of data suggesting polymorphisms of drug metabolism--hypothesis testing" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1380036). *British Journal of Clinical Pharmacology*. **28** (6): 655–662. doi:10.1111/j.1365-2125.1989.tb03558.x (https://doi.org/10.1111%2Fj.1365-2125.1989.tb03558.x). PMC 1380036 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1380036). PMID 2611088 (https://pubmed.ncbi.nlm.nih.gov/2611088).

60. Inc., Advanced Solutions International. "Sections & Interest Groups" (http://www.amstat.org/sections/srms/Proceedings/y2002/Files/JSM2002-000150.pdf) (PDF). *www.amstat.org*.

61. "Archived copy" (https://web.archive.org/web/20131103100209/http://www.uni-marburg.de/fb12/stoch/research/rpackage/manualbimodlilitytest.pdf) (PDF). Archived from the original (http://www.uni-marburg.de/fb12/stoch/research/rpackage/manualbimodlilitytest.pdf) (PDF) on 2013-11-03. Retrieved 2013-11-01.

62. "Cluster home page" (https://engineering.purdue.edu/~bouman/software/cluster/). *engineering.purdue.edu*.

63. Mächler, Martin (25 August 2016). "nor1mix: Normal (1-d) Mixture Models (S3 Classes and Methods)" (https://cran.r-project.org/web/packages/nor1mix/index.html) – via R-Packages.

64. Young, Derek; Benaglia, Tatiana; Chauveau, Didier; Hunter, David; Elmore, Ryan; Hettmansperger, Thomas; Thomas, Hoben; Xuan, Fengjuan (10 March 2017). "mixtools: Tools for Analyzing Finite Mixture Models" (https://cran.r-project.org/web/packages/mixtools/index.html) – via R-Packages.

65. "discrimARTs" (https://cran.r-project.org/web/packages/discrimARTs/discrimARTs.pdf) (PDF). *cran.r-project.org*. Retrieved 22 March 2018.

66. Gruen, Bettina; Leisch, Friedrich; Sarkar, Deepayan; Mortier, Frederic; Picard, Nicolas (28 April 2017). "flexmix: Flexible Mixture Modeling" (https://cran.r-project.org/web/packages/flexmix/index.html) – via R-Packages.

67. Fraley, Chris; Raftery, Adrian E.; Scrucca, Luca; Murphy, Thomas Brendan; Fop, Michael (21 May 2017). "mclust: Gaussian Mixture Modelling for Model-Based Clustering, Classification, and Density Estimation" (https://cran.r-project.org/web/packages/mclust/index.html) – via R-Packages.

68. Ruedin, Didier (2 April 2016). "agrmt" (https://cran.r-project.org/web/packages/agrmt/index.html). cran.r-project.org.

69. Macdonald, Peter; Du, with contributions from Juan (29 October 2012). "mixdist: Finite Mixture Distribution Models" (https://cran.r-project.org/web/packages/mixdist/index.html) – via R-Packages.