SIMPLE TESTS FOR BIMODALITY AND BITANGENTIALITY

By J. B. S. HALDANE

It is well known that the graph of a frequency distribution may be bimodal. If the modes are at terminal values of the variate this is a U-shaped distribution, often expressible in Pearsonian terms. If both are maxima at intermediate values, this implies that the cumulative distribution function has two inflexions. Harris & Smith (1949) introduced the important idea of bitangentiality. A frequency distribution which is the sum of two roughly normal distributions may have a double tangent, and thus four inflexions.

In fact we are presented with a set of discrete values of a variate, or with grouped data, and wish to know whether a smoothed curve which might represent the distribution of which they are a sample would probably be bimodal or bitangential. This problem is relatively rare in animal and plant biometry, where further specimens can often be obtained in doubtful cases. It is not rare in human and bacterial biometry, and is likely to arise in palaeontological biometry.

Where the number of individuals is relatively small, a graph may have marked troughs and peaks, but their significance is uncertain. As an example I take the data of Table 1, kindly put at my disposal by Mr T. E. Reed. The variate x is the difference, expressed in arbitrary units, and grouped, between the hair colours of each of 162 pairs of sibs, after a correction for sex difference. n_x is the number of pairs in which the difference x was found. A graph suggests anything up to four modes, and in particular one may ask whether the low values for $x = 3$ and 7, and the high ones for $x = 18$ and 19, are significant.

Table 1. *Frequency distribution of differences between sib pairs*

x	0	1	2	3	4	5	6	7	8	9	10	
n_x	8	12	16	8	13	15	13	5	7	8	10	
x	12	13	14	15	16	17	18	19	20	21	22	23
n_x	7	4	1	4	3	2	5	5	1	2	0	1
x	24	25	26	27	28	29	30	31	32	33	33+	
n_x	1	1	0	1	0	1	0	0	1	1	0	

If a distribution is bimodal it is certainly bitangential. Before we can assert bitangentiality it is a condition that some value of n_x (after further grouping if necessary) should be significantly less than $\frac{1}{2}(n_{x-1}+n_{x+1})$. This will not by itself indicate bitangentiality. It will merely indicate that the frequency curve is locally concave upwards, or if $F(x)$ is the cumulative distribution unction, d^3F/dx^3 is positive. This will be the case near the 'tails' of any bell-shaped distribution.

We are thus led to investigate the distribution of the second central difference

$$d_x = \Delta^2 n_{x-1} = n_{x-1} - 2n_x + n_{x+1}.$$

Let

$$N_x = n_{x-1} + n_x + n_{x+1}.$$

Then

$$d_x = N_x - 3n_x.$$

If d^2F/dx^2 were zero in this region, so that the graph of the frequency distribution was linear, we should expect d_x to be zero. On this null hypothesis we can determine the distribution of d_x for a given N_x. For $\mathscr{E}(n_x) = \frac{1}{3}N_x$, and of the cumulants of the distribution of d_x, $\kappa_1 = 0$, while κ_r

is $(-3)^r$ times the rth cumulant of the distribution of n_x. Thus the cumulants of the distribution of d_x are

$$\kappa_1=0, \quad \kappa_2=2N_x, \quad \kappa_3=-2N_x, \quad \kappa_4=-6N_x, \quad \kappa_5=+30N_x, \quad \kappa_6=+42N_x, \quad \text{etc.},$$

or

$$\sigma=\sqrt{(2N_x)}, \quad \gamma_1=-\sigma^{-1}, \quad \gamma_2=-3\sigma^{-2}, \quad \gamma_3=15\sigma^{-3}, \quad \gamma_4=21\sigma^{-4}.$$

Thus the probability that $|d_x|$ should exceed a given value is roughly twice the probability that a reduced normal variate should exceed $\frac{|d_x|-\frac{3}{2}}{\sqrt{(2N_x)}}$ (cf. David, 1949, p. 53).

This probability can be read off from any table of the Normal Integral. However, if greater accuracy is desired we can use the following device. It follows from Kendall's (1943) equation (6·73) that if a variate y has mean zero and unit standard deviation, but asymmetry and kurtosis measured by γ_1 and γ_2, its distribution is equivalent to that of a normal variate given by

$$z'=y-\tfrac{1}{6}\gamma_1(y^2-1)+\frac{y}{72}[2\gamma_1^2(4y^2-7)-3\gamma_2(y^2-3)]+\ldots.$$

Hence if, for positive d_x,

$$z=\frac{d_x-\frac{3}{2}}{\sqrt{(2N_x)}},$$

then

$$z'=\frac{d_x-\frac{3}{2}-\frac{1}{6}(z^2-1)}{\sqrt{(2N_x)}}$$

is somewhat more normally distributed than z. In practice this refinement is hardly needed.

When N_x is small the deviations from normality are considerable, but we can always use the binomial theorem. It can then be seen that a positive value of d_x cannot be significant at the 5 % level if $N_x<8$; a negative value, which would suggest bitangentiality, cannot be significant if $N_x<4$. Hence we need not tabulate values of d_x corresponding to $N_x<4$.

We can now answer the question: 'Is the evidence sufficient to make it highly probable that the true frequency f_x represented by any of the values of n_x in Table 1 exceeds or falls short of the mean of its neighbours?' Table 2 shows the values of d_x and their standard errors. Thus, in the first column

$$d_x=8-2\times12+16=0, \quad \sigma=\sqrt{\{2(8+12+16)\}}=8\cdot49,$$

and so on. No difference reaches twice its standard error on the null hypothesis. In the most conspicuous case, $\frac{d_3-\frac{3}{2}}{8\cdot60}=1\cdot337$, giving $P=0\cdot181$, the true value from the binomial distribution being 0·175. There is thus no evidence that any value of n_x differs significantly from the mean of its neighbours.

Table 2

x	1	2	3	4	5	6	7	8	9	10
d_x	0	−12	+13	−3	−4	−6	+10	−1	+1	−6
σ	8·49	8·49	8·60	8·49	9·06	8·12	7·07	6·32	7·07	6·93
x	11	12	13	14	15	16	17	18	19	20
d_x	+5	−4	0	+6	−4	0	+4	−3	−4	+5
σ	6·78	5·83	4·90	4·24	4·00	4·24	4·47	4·90	4·69	4·00

To obtain better evidence we must increase the ratio of our d_x values to their standard errors by adopting a coarser grouping. This is particularly likely to be effective where two consecutive values of d_x have the same sign. Table 3 gives the results of such grouping. In Table 3*a* I have grouped n_1 and n_2, n_3 and n_4, and so on; in Table 3*b* I have grouped n_2 and n_3, n_4 and n_5, and so on.

Three values of $| d_x |$ exceed 2σ, but in no case does $| d_x | - \frac{3}{2}$ exceed 2σ. In Table 3*a*, $n_{16\cdot5} = 5$, $n_{18\cdot5} = 10$, $n_{20\cdot5} = 3$. The exact probability of obtaining so large a value of $n_{18\cdot5}$ is 0·0434, giving $P = 0{\cdot}087$. In Table 3*b* the high negative value of $d_{5\cdot5}$ merely indicates a peak in a region where one would be expected, the high positive value of $d_{7\cdot5}$ indicating a trough where a unimodal distribution would probably show little curvature. But $\frac{d - \frac{3}{2}}{\sigma} = 1{\cdot}9034$, so $P = 0{\cdot}057$. Among 19 values of $| d_x | - \frac{3}{2}$ in Table 3 there is nothing surprising in finding one which is almost significant at the 5 % point. Thus, while the distribution of which Table 1 represents a sample may well be multimodal or bitangential, the sample is not large enough to make the contrary hypothesis at all improbable.

Table 3*a*

x	2·5	4·5	6·5	8·5	10·5	12·5	14·5	16·5	18·5	20·5	22·5	24·5
d_x	0	−14	+7	+4	−6	−1	+6	+5	−12	+5	+3	−2
σ	12·00	11·83	11·05	9·90	9·17	8·00	6·48	6·32	6·00	5·29	3·46	2·83

Table 3*b*

x	3·5	5·5	7·5	9·5	11·5	13·5	15·5	17·5	19·5	21·5	23·5
d_x	+14	−23	+22	−11	−3	+10	−2	−1	−3	+4	−1
σ	12·41	11·05	10·77	9·27	8·49	7·07	6·16	6·32	5·48	4·47	3·16

Let us, however, consider what conclusions could have been drawn had one or more values of $| d_x | - \frac{3}{2}$ divided by their standard errors had values giving $P = 0{\cdot}01$ or less. A high negative value near the mean of the distribution merely indicates the presence of a peak there, which is to be expected. A high negative value elsewhere indicates at least bitangentiality. Thus, had $d_{18\cdot5}$ been -18, there would have been good reason to suspect it. A high positive value anywhere indicates a trough, or concavity upward, but if it occurs in a part of the distribution where concavity upward would be expected in any case it does not necessarily indicate bitangentiality.

Let us see how we could refine our test if we assumed as our null hypothesis that the true frequencies f_x were in a geometrical progression with common ratio p. Suppose then that

$$\mathscr{E}(n_{x-1}) = \frac{pN_x}{p+1+p^{-1}}, \quad \mathscr{E}(n_x) = \frac{N_x}{p+1+p^{-1}}, \quad \mathscr{E}(n_{x+1}) = \frac{p^{-1}N_x}{p+1+p^{-1}}.$$

Let $p + p^{-1} - 2 = g$. Then $$\mathscr{E}(d_x) = \frac{(p-2+p^{-1})\,N_x}{p+1+p^{-1}} = \frac{gN_x}{3+g}.$$

Similarly $$\text{var}\,(d_x) = \frac{9(2+g)}{(3+g)^2}\,N_x = N_x[2 - \tfrac{1}{3}g - \tfrac{2}{27}g^3 + \ldots].$$

Now g is generally small. Thus even if p were 1·5, which would imply very coarse grouping, g would only be $\frac{1}{6}$. In Table 1, $n_2 = 16$, $n_{21} = 2$, so a reasonable value of p is $8^{-\frac{1}{16}}$, or 0·896, giving $g = 0{\cdot}012$, or $\mathscr{E}(d_x) = 0{\cdot}0040\,N_x$, which is clearly a negligible correction.

Thus, in fact, high positive values of d_x anywhere indicate bitangentiality unless the grouping is exceedingly coarse. A single high negative value near the probable mode of the distribution does not do so. But two non-adjacent high negative values do so, provided that they differ significantly from some of the intervening values, even though these latter are not significantly positive.

If good evidence of bitangentiality is found, we can look for evidence of bimodality. For this purpose we consider the significance of first differences $\Delta n_x = x_{n+1} - x_n$. If x_n is significantly

smaller than x_{n-1} *and* x_{n+1}, we have evidence of an antimode. If it is significantly larger than both, we have evidence of a mode.

Penrose (1951) has produced human data showing very strong evidence for bitangentiality. Table 4 (extracted from his Table II) gives the maternal ages in 1288 cases of mongoloid imbecility, grouped in 5-year intervals. The first row in this table gives the central maternal age, the second is the number of births among 1038 cases with no known affected relative, the third row includes another 250 related cases. For the second row $d_{32}=104$, $N_{32}=542$, $\frac{d-\frac{3}{2}}{\sqrt{(2N)}}=3{\cdot}148$ giving $P=0{\cdot}0014$. For the third row $d_{32}=135$, $N_{32}=705$, $\frac{d-\frac{3}{2}}{\sqrt{(2N)}}=3{\cdot}555$, $P=0{\cdot}00038$. The other high positive values of d are where they would be expected, at the 'tails' of the distribution. The 129 cases each related to another case through a father show further evidence of bitangentiality, though not quite significant. The 121 cases related through a mother show no such evidence, though their distribution has a striking plateau suggesting a compound distribution.

Table 4

Central maternal age	x	17	22	27	32	37	42	47	52
Unrelated mongols	n_x	10	64	112	146	284	331	90	1
All mongols	n_x	13	85	167	190	348	396	97	2
(For all mongols)	d_x	+49	+10	−59	+135	−110	−347	+204	+93
(For all mongols)	σ	14·0	22·58	29·73	37·55	43·22	41·09	31·46	9·95

None of the distributions, as grouped, shows any serious evidence of bimodality, but with such extensive data it is conceivable that a finer grouping might give even stronger evidence for bitangentiality, and perhaps even for bimodality.

This bitangential distribution could be due, as Penrose suggests, to the superposition of two distributions, one with a lower mean age for homozygous mothers, and one with a higher mean age for heterozygous mothers. It could be due to the multiplication of the normal maternal age distribution by a probability distribution for the production of a mongol which rose very steeply with maternal age. The former hypothesis is, perhaps, the more likely.

To illustrate a successful search for plurimodality I take the data of Table 5, which were obtained by Mitchison (1951, his Table 2*a*) on 156 cultures of *Mycobacterium tuberculosis* growing, after previous selection, in media containing different concentrations of streptomycin. x is the logarithm, to base 2, of 4 times the minimal inhibitory concentration of streptomycin in milligrams per litre. Thus the value of n_6 means that two cultures were inhibited by 16 mg./l., but not by 8 mg./l. The last entry refers to cultures not inhibited by 2048 mg./l. There are four peaks, at $x=1$, 3, 8 and >13. The peak at $x=8$ could clearly be magnified by suitable grouping. The most doubtful feature is the apparent antimode at $x=2$. The two differences, tested for significance, are $\frac{39-17-1}{\sqrt{(39+17)}}=2{\cdot}806$, and $\frac{33-17-1}{\sqrt{(33+17)}}=2{\cdot}121$.

Table 5

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	>13
n_x	7	39	17	33	7	3	2	0	15	14	5	0	0	0	14

The probabilities of such high values, if the frequencies had been equal, are 0·0025 and 0·017. The probability that at least one of these values is due to chance is approximately their sum, or 0·0195. Now this would not be significant had there been, say, fifty cases where n_x could be

tested for its differences from n_{x-1} and n_{x+1}. Such a test can only be made at the 5 % level if the sum of two consecutive values exceeds 5. Now there are only six cases in this table where $n_{x-1}+n_x$ and n_x+n_{x-1} both exceed 5, and three of these give significant evidence of modes or peaks, without further grouping. The trough at $x=2$ is, therefore, at least significant at the 8 % level, and I think that this figure somewhat underrates its significance.

The peak at $x>13$ does not formally satisfy our criteria, but Dr Mitchison tells me that he could not use higher concentrations, since the medium was saturated with streptomycin. A similar difficulty arises with phenylthiourea. However, in other experiments of the same type, a few cultures were inhibited at $x=13$, so there is probably a genuine peak in this region.

Mitchison gives data on four other groups of cultures, in which similar peaks are shown. Not all of them give significant evidence of all the 'peaks' and 'troughs', but they serve to enhance the significance of those shown in Table 4. Thus, if his Tables 2*a* and 2*b*, which refer to the same strain, are added together, the values of $\frac{|\Delta n_x|-1}{\sigma}$ for the trough at $x=2$ are raised to 2·73 and 3·47. There can be no doubt that Mitchison was statistically justified in his conclusion 'It is clear that there are four such peaks'. It is, of course, possible that had he worked on a larger scale he would have discovered a fifth peak. However, his data do not give significant evidence even of bitangentiality other than that due to the four modes, or peaks, which he describes.

It may be remarked that Mitchison's data, as would be expected, furnish stronger evidence for bitangentiality than for bimodality. Thus, from Table 4, $d_2=+38$, $N_2=89$, $\frac{d-\frac{3}{2}}{\sqrt{N}}=3{\cdot}87$. So large a positive value would occur by chance with a frequency of 0·001 if the frequencies in the three classes were equal. However, a bacteriologist might be dubious as to the biological meaning of bitangentiality, whereas he would regard bimodality as good evidence for the existence of a sharp difference, even if he could not always classify a given culture with certainty.

Finally, it is perhaps worth generalizing the problem, and considering the distribution of the rth difference $\Delta^r n_x$, if the expectation of each n_x is $\frac{N}{r+1}$.

If
$$d=n_0-\binom{r}{1}n_1+\binom{r}{2}n_2+\ldots+(-1)^r n_r,$$

then
$$\mathscr{E}(n_i)=\frac{N}{r+1},\quad \mathscr{E}[n_i^{(2)}]=\mathscr{E}[n_i n_j]=\frac{N(N-1)}{(r+1)^2}$$

and so on. Hence
$$\mathscr{E}(d)=0,$$
$$\mathscr{E}(d^2)=\frac{N}{r+1}\sum_{i=0}^{r}\binom{r}{i}^2$$
$$=\frac{(2r)!\,N}{(r!)^2\,(r+1)}.$$

Similarly, it can be shown that the distribution is symmetrical if r is odd, whilst if r is even
$$\mathscr{E}(d^3)=\frac{N}{r+1}\sum_{i=0}^{r}\left[(-1)^i\binom{r}{i}^3\right].$$

Thus successive values of σ_d^2/N are

1, 2, 5, 14, 42, 132, etc.

Successive values of $\gamma_1\sigma$ are

$$0, \quad -1, \quad 0, \quad \frac{+9}{7}, \quad 0, \quad \frac{-20}{11}, \text{ etc.}$$

A significantly high third difference would indicate an inflexion in the frequency distribution graph, and so on.

The method here developed is not elegant. It calls for a certain amount of common sense in its application. It is possible to suggest methods which could be translated into instructions for a calculating machine, as, for example, the methods for estimating the significance of departures from symmetry or normality could be coded. The test for bimodality is as follows: Tabulate all values of Δn_x for different groupings. If a significantly positive value immediately succeeds a significantly negative one, or if in two cases significantly negative values immediately succeed significantly positive ones, we have evidence of bimodality, but the precise degree of significance to be attached to such evidence is not clear, and it is harder to frame a clear-cut test for bitangentiality. While it is to be hoped that more precise tests can be devised, it is believed that those here given may be of use in doubtful cases.

SUMMARY

Simple tests are given for the bitangentiality and bimodality of a distribution represented by a sample.

I have to thank Mr Reed for permission to publish his data.

REFERENCES

DAVID, F. N. (1949). *Probability Theory for Statistical Methods*. Cambridge.

HARRIS, H. & SMITH, C. A. B. (1949). The sib-sib age of onset correlation among individuals suffering from a hereditary syndrome produced by more than one gene. *Ann. Eugen., Lond.*, **14**, 309.

KENDALL, M. G. (1943). *The Advanced Theory of Statistics*, **1**. London.

MITCHISON, D. A. (1951). The segregation of streptomycin-resistant variants of *Mycobacterium tuberculosis* into groups with characteristic levels of resistance. *J. Gen. Microbiol.* **5**, 3.

PENROSE, L. S. (1951). Maternal age in familial mongolism. *J. Ment. Sci.* **97**, 408, 738.