Brit. J. Psychiat. (1981), **138**, 336–339

Bimodality and the Nature of Depression

B. S. EVERITT

Summary: Arguments concerning the nature of depressive disorders have involved as a central issue the question of the bimodality or otherwise of the distribution of some variable expressing variation in symptomatology. The implications of a particular type of frequency distribution along this dimension, whether uni- or bi-modal, have been misunderstood by a number of workers, and an attempt is made to clarify the situation.

The nature of depressive disorders continues to be an area of controversy and disagreement amongst psychiatrists. A number of studies, for example, Kiloh and Garside (1963), and Pilowsky, Levine and Boulton (1969), indicate that they have evidence for the existence of two subtypes of depression, whilst others, for example, Kendell (1969) and Kendell and Gourlay (1970) argue for the existence of only a single type. The protagonists in the debate have armed themselves with a variety of sophisticated statistical weapons such as discriminant function analysis, cluster analysis and factor analysis, with which to support their particular standpoint and applied them in a variety of circumstances and to various data sets, although in many cases the methods do not appear to have been particularly appropriate. It is not the purpose of this paper to review the debate or the use (and misuse) of statistical techniques in the arguments; instead we shall concentrate on an issue which has been seen as central to both sides, namely that of the bimodality or otherwise of the frequency distribution of some variable expressing variation in symptomatology.

The points to be made in the rest of this paper are not new; indeed they should already be well known since they have been made previously by a number of authors, for example, Murphy (1964) and Fleiss (1972). However, it is thought worthwhile reiterating them here since the debate appears to be gathering momentum once again, as evidenced by a number of recent papers such as Garside and Roth (1978), and Kendell and Brockington (1980), and the bimodality issue still appears to be misunderstood.

Mixture Distributions and Bimodality

In many studies exploring the nature of depression, variation in symptomatology has been expressed by a score on a linear discriminant function found from the analysis of a number of items measured for patients diagnosed originally as neurotic or psychotic depression. The distribution of such scores has then been used to make inferences about the structure of depression, the argument generally being that a bimodal distribution indicates the presence of subtypes, and unimodality the presence of only a single type.

Now for a *population* frequency curve, bimodality is (except in pathological cases) a sufficient (although not a necessary) condition for the presence of subtypes and certainly if, in a fairly large sample, bimodality appeared no matter how the data were arranged, it would be pedantic to insist that it might be an artifact. However with small samples one can often choose class intervals for the histogram which make the distribution bimodal rather than unimodal, and vice-versa. Figs 1 (a) and 1 (b) show histograms of a small data set plotted with different class intervals. One is unimodal, and one has three modes. Further examples are given in Murphy (1964) of samples of size 50 drawn from a single normal distribution, which show signs of bimodality and even, in some cases, trimodality. Consequently the question of the *apparent* bimodality in a sample is a very real one.

If the distribution of discriminant function scores is unimodal, then it is generally inferred that only a single type is present. Unfortunately this again is not necessarily correct, since the presence of subtypes in a population *can* lead to a unimodal frequency distribution when certain conditions are satisfied, for example, when the means of the subtypes are not well separated or when the subtypes are present in widely different proportions. Consequently inferring the presence of subtypes *only* from the bimodality of the distribution is much too stringent a criterion.

The frequency distribution of discriminant function scores may be investigated more formally by fitting *mixture* distributions, in particular mixtures of normal distributions. In essence this involves determining whether a single normal distribution, i.e.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (1)$$

fits the distribution better than a mixture of two normal distributions i.e.

$$h(x) = \frac{p}{\sqrt{2\pi}\,\sigma_1} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} + \frac{(1-p)}{\sqrt{2\pi}\,\sigma_2} e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2} \qquad (2)$$

where p represents the proportion of one of the two subtypes in the population, and μ_1 σ_1, μ_2, σ_2 the subtype means and standard deviations. Estimation of the parameters in the mixture distribution, h (x), is now a fairly routine (although non-trivial) task using maximum likelihood methods. Fleiss (1972) describes an example. Eisenberger (1964) and Behboodian (1970) have investigated the precise conditions under which the distribution, h (x), is unimodal; for interest Figs 2 (a) and 2 (b) illustrate two such cases.

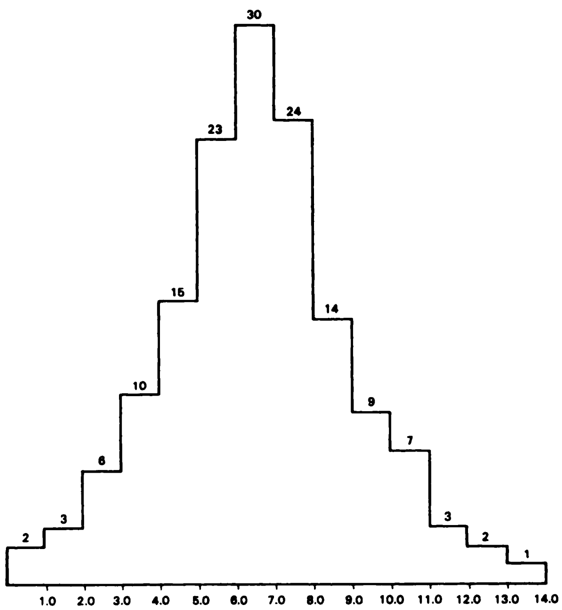
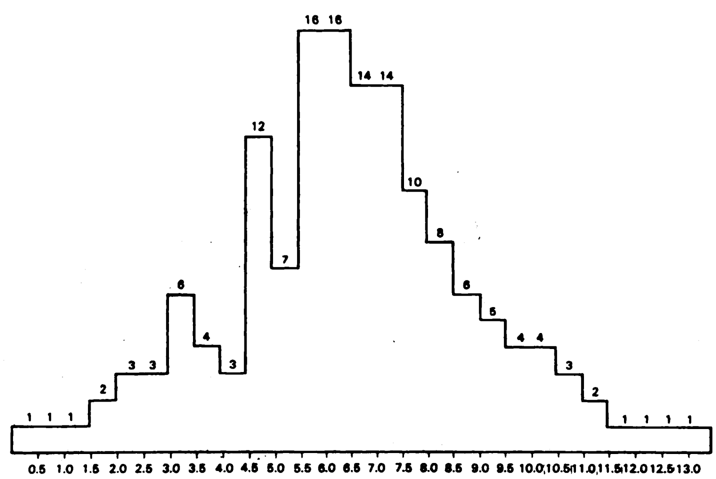


FIG 1(a)

Discussion

Understanding the nature of depression is likely to be a difficult and protracted undertaking. Focussing too much on one particular feature such as the bimodality or otherwise of scores on some dimension expressing variation in symptomatology is too simplistic and, in many situations, a waste of time. The mixing of two unimodal frequency curves produces a bimodal distribution only if the components are fairly widely separated, and it is the mixing *not* the bimodality which is fundamental (see Cox, 1966). On the other hand, bimodality can arise for a variety of reasons other than the presence of distinct subtypes, including non-representativeness of sampling, observer differences, or inadequate definition of variables. Consequently it may be more appropriate for workers



FIGS 1(a) and 1(b).—Histograms of same data set plotted with different class intervals.

https://doi.org/10.1192/bjp.138.4.336 Published online by Cambridge University Press

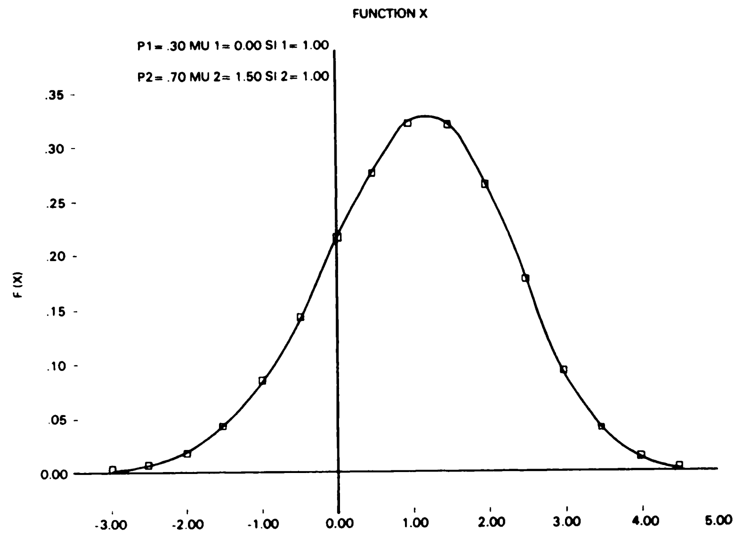
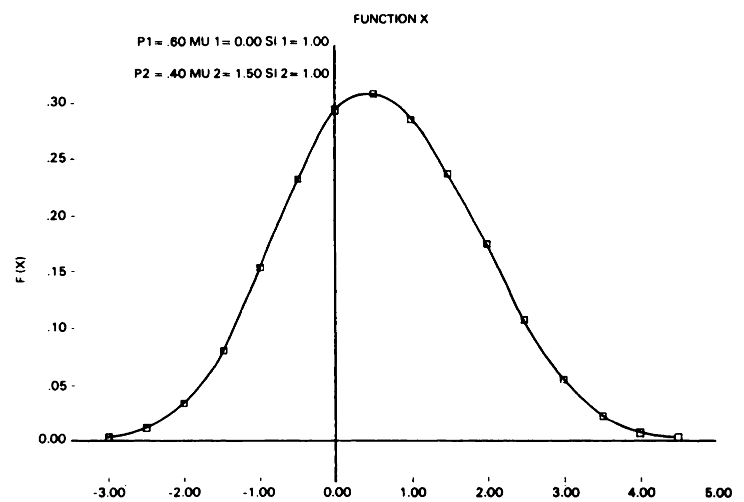


FIG 2(a)



FIGS 2(a) and 2(b).—Frequency curves of mixtures of two normal distributions. (a) $p = 0.3$, $\mu_1 = 0.0$, $\sigma_1 = 1.0$, $\mu_2 = 1.5$, $\sigma_2 = 1.0$: (b) $p = 0.6$, $\mu_1 = 0.0$, $\sigma_1 = 1.0$, $\mu_2 = 1.5$, $\sigma_2 = 1.0$.

in this area to consider fitting mixture distributions to their data in their attempts to gain evidence for or against the existence of two types of depression. The computations involved are heavy, but no more so than those involved in the variety of other statistical techniques routinely employed. Details are available in Everitt and Hand (1981).

References

BEHBOODIAN, J. (1970) On a mixture of normal distributions. *Biometrika*, **57**, 215–17.

COX, D. R. (1966) Notes on the analysis of mixed frequency distributions. *British Journal of Mathematical and Statistical Psychology*, **19**, 39–47.

EISENBERGER, I. (1964) Genesis of bimodal distributions. *Technometrics*, **6**, 357–64.

https://doi.org/10.1192/bjp.138.4.336 Published online by Cambridge University Press

EVERITT, B. S. & HAND, D. J. (1981) *Finite Mixture Distributions*. London: Chapman and Hall.

FLEISS, J. L. (1972) Classification of the depressive disorders by numerical typology. *Journal of Psychiatric Research*, **9**, 141–53.

GARSIDE, R. F. & ROTH, M. (1978) Multivariate statistical methods and problems of classification in psychiatry. *British Journal of Psychiatry*, **133**, 53–67.

KENDELL, R. E. (1969) The continuum model of depressive illness. *Proceedings of the Royal Society of Medicine*, **62**, 335–9.

—— & BROCKINGTON, I. F. (1980) The identification of disease entities and the relationship between schizophrenic and affective psychoses. *British Journal of Psychiatry*, **137**, 324–31.

—— & GOURLAY, J. (1970) The clinical distinction between psychotic and neurotic depressions. *British Journal of Psychiatry*, **117**, 257–60.

KILOH, L. G. & GARSIDE, R. F. (1963) The independence of neurotic depression and endogenous depression. *British Journal of Psychiatry*, **109**, 451–63.

MURPHY, E. D. (1964) One cause? Many causes? The argument from the bimodal distribution. *Journal of Chronic Disorders*, **17**, 301–24.

PILOWSKY, I., LEVINE, S. & BOULTON, D. M. (1969) The classification of depression by numerical taxonomy. *British Journal of Psychiatry*, **115**, 937–45.

B. S. Everitt, B.Sc., M.Sc., *Head of Biometrics Unit, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF*

(*Received 10 December 1980*)