for example, where $H=(\mathrm{d}F_1-\mathrm{d}F_2)/\mathrm{d}G$, with the expected value $E(p_1)=\Pi_1$, and variance

$$\sigma^2(p_1) = \frac{\int H^2\mathrm{d}F - (\int H\mathrm{d}F)^2}{n[\int(\mathrm{d}F_1 - \mathrm{d}F_2)H]^2}$$

A considerable choice of H is possible. Thus, considering for definiteness the density case, we define the unweighted least-squares estimator of Π_1 by the choice $\mathrm{d}G=dx$, $H=f_1-f_2$. We should, however, expect weighted estimators to be better. In fact, because the variance of $\mathrm{d}F_s$ is fdx/n, we should try to choose $\mathrm{d}G=fdx$; the term $\int H\mathrm{d}F$ in the expression for $\sigma^2(p_1)$, arising from the covariance of $\mathrm{d}F_s(x)$ and $\mathrm{d}F_s(y)$, then vanishes, and $\sigma^2(p_1)$ becomes identical with the reciprocal of the information function $I(\Pi_1)$. If we take $\mathrm{d}G=\Pi_0\mathrm{d}F_1+(1-\Pi_0)\mathrm{d}F_2$, then this weighting will be most efficient when Π_0 is close to Π_1. Thus if we suspect Π_1 to be near 1, $\frac{1}{2}$ or 0, suitable choices for Π_0 would be 1, $\frac{1}{2}$ or 0, respectively. An alternative to $\frac{1}{2}(\mathrm{d}F_1+\mathrm{d}F_2)$ is the geometric mean of $\mathrm{d}F_1$ and $\mathrm{d}F_2$, and another is $\max(\mathrm{d}F_1, \mathrm{d}F_2)$. The geometric mean has some convenience over the arithmetic mean in theoretical investigations of efficiency (for example, for f_1 and f_2 normal), but the latter (or alternatively $\max(\mathrm{d}F_1, \mathrm{d}F_2)$) has the advantage of approximating to the maximum likelihood estimator, for any value of Π_1, in both the extreme cases of f_1 and f_2 well separated and of $f_1 \rightarrow f_2$. It has been shown by Hill[1] that the information function $I(\Pi_1)$ for Π_1 may be written $n[1-S(\Pi_1)]/(\Pi_1\Pi_2)$, where

$$S(\Pi_1) = \int(f_1f_2/\mathrm{f})\mathrm{d}x$$

If f_1 and f_2 differ by some parameter μ (for example, the mean), and $\Delta\mu=\mu_2-\mu_1$, then also as $f_1 \rightarrow f_2$,

$$I(\Pi_1) \sim n(\Delta\mu)^2I(\mu)$$

where $I(\mu)$ is the information function for μ. Thus for the maximum likelihood estimator $\hat{\Pi}_1$ the variance, which is $\Pi_1\Pi_2/n$ for f_1 and f_2 well separated, is $1/[n(\Delta\mu)^2I(\mu)]$ as $f_1 \rightarrow f_2$.

In the case of further unknown parameters θ_j the problem is obviously more complex, because of the non-linear dependence of f on θ_j; but linearization of the estimation equations may sometimes be convenient. For example, for two populations with unknown parameters θ_1, θ_2, respectively, we might adopt the standard least-squares procedure of taking provisional values v_1, v_2, say, so that in the "sum of squares" to be minimized we write in the numerator

$$f \sim f_2(v_2) + \Pi_1(f_1(v_1) - f_2(v_2)) + \varphi_1\partial f_1/\partial v_1 + \varphi_2\partial f_2/\partial v_2$$

in the unknowns Π_1, φ_1 and φ_2, where

$$\varphi_1 = \Pi_1(\theta_1 - v_1), \varphi_2 = \Pi_2(\theta_2 - v_2)$$

(retaining the provisional values v_1 and v_2 in the weighting-function $\mathrm{d}G$).

M. S. BARTLETT
P. D. M. MACDONALD

Department of Biomathematics,
University of Oxford.

Received December 13, 1967.

[1] Hill, B. M., *J. Amer. Stat. Assoc.*, **58**, 918 (1963).

Speculations on the Use of Orthonormal Functions in the Study of Morphogenesis

THE recent proposal[1] to analyse shapes of biological objects by means of orthonormal functions opens up some interesting possibilities. In the work cited, a particular set of two-dimensional orthonormal functions synthesized from Walsh[2] functions which take only the values +1 and −1 was used. As is well known, however, there is a vast number of other complete sets of orthonormal functions and, in fact, they could be constructed indefinitely. There are also standard linear methods for transforming from one set of orthonormal functions to another; these procedures can be realized as computer programmes.

The Walsh-derived functions are most suitable as an initial set because of their simplicity. Their characteristic chequerboard patterns, however, are not the most natural to use for the analysis of familiar kinds of biological objects; this was indicated by the fact[1] that, for good accuracy, as many as 256 terms might be needed in the expansion for a leaf. It might be possible to construct a different orthonormal set of basic patterns which would be more "leaf-like" in shape; ideally, a development series of leaves would all be adequately described by merely two or three terms, say. What is required then is to devise computer programmes which, for a given series of input shapes, would find the best orthonormal set—using the transformation procedures referred to—best in the sense that fewest terms would be required for adequate description. When this had been achieved, one might hope to discover laws of a general type from the experimental data, for example, by studying the changes of the weights (that is, coefficients in the orthonormal expansion) that occur in development or because of experimentally applied external agents. If this were achieved, the basic patterns would presumably have genuine biological significance.

What follows is rather more speculative. It is well known that most of the standard orthonormal functions of mathematical physics (sinusoidal or Bessel, for example) arise naturally in the solution of differential equations with boundary constraints. Thus a given set of orthonormal functions characterizes a given differential equation with its associated boundary conditions. If then an orthonormal set which was particularly adequate in the sense of the previous paragraph was found, this set might well have a special significance and be indicative of some differential equation or, maybe, difference equation or other type of equation, which characterizes in a general way the development of the patterns concerned. It might thus be possible, merely from examining the optimum orthonormal sets generated by the computer programmes, to draw conclusions about or even to derive the fundamental laws—in the form of differential equations or something similar—which govern the development of form.

Another issue, which it may be fruitful to investigate, is the biological meaning, if any, of the maximal weights previously discussed[1]. If one attempted to reconstruct a form from the maximal weights directly, one would fail, because in general the constructed function would have other values than 0 and 1. Nevertheless, this function, because it contains all the information about the degrees of resemblance of the form concerned to the basic pattern forms, should have a significance which might be usable—especially in those cases where the number of basic patterns effectively used is very small, as discussed previously.

It would also be useful to generalize the orthonormal expansion method by allowing the form function to have more than two values. For example, a representation of the veins of a leaf might be included with a value of 2 being assigned at locations on a vein. This would provide a three-valued function, but many-valued functions could be used also for other purposes such as giving the distribution of colour as well as shape of an object, when a suitable numerical coding of the colours had been chosen.

BERNARD MELTZER

Metamathematics Unit,
University of Edinburgh.

Received December 5, 1967.

[1] Meltzer, B., Searle, N. H., and Brown, R., *Nature*, **216**, 32 (1967).

[2] Walsh, J. L., *Amer. J. Math.*, **45**, 5 (1923).