if a set of k *sufficient* or of k *quasi-sufficient* statistics can be found, and it is known (from the work of Segal[2], further discussed by Bartlett[3]) that a simultaneous fiducial distribution will then exist, though it may not be *complete* in the sense of Bartlett. The most general parent population admitting a set of sufficient statistics has been determined by Koopman[4], but the parallel problem for quasi-sufficiency (Fisher's "Problem of the Nile") is still unsolved.

Fisher considered quasi-sufficient estimation in relation to two important examples: (1) the estimation of a location parameter; and (2) the estimation of a pair of parameters defining the location and scale of the distribution. In each example he showed that the whole of the relevant information is contained in the estimating statistics provided that these are considered in association with the *configuration* of the sample from which they have been derived. In the first example the configuration is the set of differences between the observations (arranged in descending order); in the second it is the set of ratios of these differences.

I have now generalized Fisher's results, taking the configuration of a sample of n members to consist of $n-k$ functionally independent symmetric functions of the observations, such that the configuration of a sample determines the configurations of all included sub-samples once the observations have been arranged in descending order. The integer k, the *order* of the configuration, corresponds to the number of parameters to be estimated. With this definition (and subject to certain restrictions of detail) it appears that a first-order configuration is always equivalent to the set of differences between the observations on some transformed scale, while a second-order configuration is similarly always equivalent to the set of ratios of the differences. There is a formal analogy with the theory of continuous groups which suggests that a third-order configuration is always equivalent to the set of cross-ratios of tetrads of the transformed observations, and that higher-order configurations do not exist.

Now suppose that samples from a one-parameter distribution $\varphi(x\,;\theta)dx$ admit a first-order configuration in the sense defined above, for which the $n-1$ defining functions have a joint distribution independent of θ. A quasi-sufficient statistic for θ will then exist, and it can be shown that the wider class of distributions described by Fisher[1] and permitting quasi-sufficient estimation of the parameter by virtue of their invariance under certain transformations are included within the present formulation. By means of a characteristic-function argument I have proved that in these circumstances the parent distribution must be of the form $\varphi(X-\lambda)dX$, where $X=X(x)$ and $\lambda=\lambda(\theta)$. (Dr. Olav Reiersøl has pointed out to me that this result bears some resemblance to one mentioned by Segal[2], the precise content of which cannot be determined as the necessary details are not given.)

In the two-parameter case the corresponding situation is less simple, and a number of additional restrictions have to be made; apart from these, however, it appears that the parent distribution must have the form

$$\varphi\left(\frac{X-\lambda}{\mu}\right)\frac{dX}{\mu}, \qquad X = X(x),$$

where (λ, μ) are functions of (θ_1, θ_2).

Thus, within the limitations of the present approach, the two examples for $k \leqslant 2$ discussed by Fisher are not only representative but exhaustive.

Apart from three-parameter distributions of the form

$$d\Phi\left(\frac{\alpha X+\beta}{\gamma X+\delta}\right), \quad \alpha\delta \sim \gamma\beta = 1, \quad X = X(x),$$

the parameters of which can be estimated by a set of quasi-sufficient statistics associated with the 'cross-ratio' type of third-order configuration, my results do not lead to any essentially new solutions of the notoriously difficult "Problem of the Nile", but they cover exhaustively one outstanding line of approach, and suggest that further progress in this field is most likely to be made by attempting to generalize Fisher's solution in a way which does *not* preserve one of its characteristic features—the determination of the configurations of sub-samples by the configuration of their parent.

A detailed account of this work will, it is hoped, be published during the course of the coming year.

DAVID G. KENDALL

Magdalen College,
Oxford.
Aug. 2.

[1] Fisher, R. A., *Proc. Roy. Soc.*, A, **144**, 285 (1934).
[2] Segal, I. E., *Proc. Camb. Phil. Soc.*, **34**, 41 (1938).
[3] Bartlett, M. S., *Ann. Math. Stat.*, **10**, 129 (1939).
[4] Koopman, B. O., *Trans. Amer. Math. Soc.*, **39**, 399 (1936). Darmois, G., *C.R.*, **222**, 164 and 266 (1946), and Féraud, L., *ibid.*, 1272, should also be consulted.

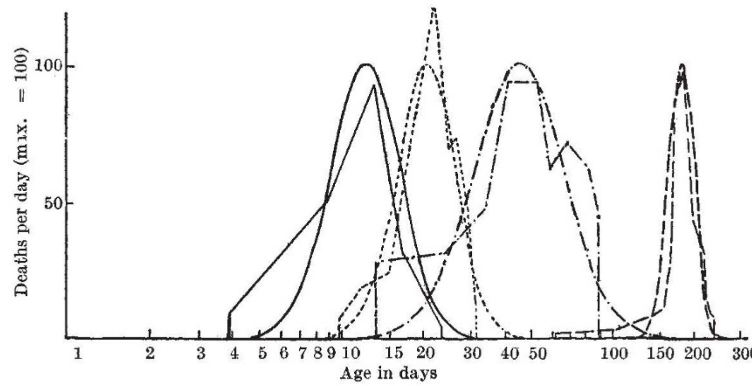
I AGREE entirely with the opinion of Dr. Kendall that future progress in this field should not be restricted to the consideration of cases in which the configuration of sub-samples is determined by the configuration appropriate to the whole sample. Such a restriction is new to me, and indeed appears artificial; it is not surprising that within its limitation the method I had illustrated should be applicable to only a few special cases. Dr. Kendall makes no mention of other illustrations I have given of the use of ancillary information.

R. A. FISHER

Representation of Relative Variability on a Semi-logarithmic Grid

THERE are two types of semi-logarithmic grids: one, called arith-log grid (a system of rectangular co-ordinates, in which the axis of abscissæ is divided arithmetically and that of ordinates logarithmically), is frequently used for graphical representations because, on such a grid, the relative increase or decrease of the intensity of a phenomenon is shown immediately by the slope of the curve obtained by plotting its value at different times; the other, which may be termed 'log-arith' grid (a system of rectangular co-ordinates in which the axis of abscissæ is divided logarithmically and that of the ordinates arithmetically), is used rarely, although this type of grid has some very useful peculiarities. It has been recognized that many moderately skew frequency distributions arising from empirical data[1] or fulfilling certain theoretical conditions[2] are reduced to normal curves of errors when plotted on a log-arith grid. Whereas the log-arith grid shares this peculiarity with other grids where the abscissa is scaled according to different functions[3], another feature, peculiar to the log-arith grid and corresponding to the well-known peculiarity of the arith-log grid, seems to have never been appreciated.

Consider two equal segments, x_1', x_1'' and x_2', x_2'', situated on the x-axis of a log-arith grid, and let their mid-points be x_1 and x_2 respectively. Then it is obvious that $\frac{x_1''-x_1'}{x_1} = \frac{x_2''-x_2'}{x_2}$. If now x_1 and x_2 represent the averages of two frequency distributions and x_1' and x_2' the respective lower (and x_1'' and x_2'' the higher) quartiles, or $x''-x'$ any other measure of dispersion, other than the inter-quartile range, the two distributions will have equal relative dispersions. In general, distribution curves which exhibit equal broadness on a log-arith grid have equal relative dispersions, higher relative dispersion is shown by a broader curve and lower by a narrower one. The log-arith grid is thus eminently suitable for the study of relative dispersion, and, in particular, in the case of distributions approaching a type which has been termed 'log-normal'[1].



DEATH CURVES PLOTTED ON A 'LOG-ARITH' GRID AS FREQUENCY POLYGONS, AND NORMAL CURVES OF ERRORS FITTED TO THEM, SHOWING THE DISTRIBUTION OF STARVED TICKS BY AGE AT DEATH AT A CONSTANT TEMPERATURE AND DIFFERENT HUMIDITIES (—— 20 %; ---- 50 %; —·—·— 80 %; — — — 95 %)

Other methods for the graphical representation of relative dispersion have been devised[4], but their respective advantages and drawbacks will not be discussed here. Yet, the presentation of a practical example is well suited for illustrating the utility of the present method. In the figure, death curves resulting from starvation experiments on ticks (the data were communicated to me by Dr. Feldman-Muhsam, Department of Parasitology of the Hebrew University, and full details will soon be published elsewhere) have been plotted on a log-arith grid, the abscissæ being the logarithm of the age at death in days and the ordinates the number of individuals dying per day at any particular age. It can be seen from the diagram that three of the death curves present approximately the same relative dispersion, the fourth presenting a somewhat higher one. Relative dispersion, as measured by the ratio of the standard deviation to the mean, assumes corresponding values shown in column (4) of the following table.

Relative humidity (1)	Mean length of life (days) (2)	Standard deviation (3)	Coefficient of variation (4)=(3):(2)
20%	13·7	3·7	0·27
50%	23·1	4·5	0·20
80%	48·0	22·4	0·47
95%	162·7	37·0	0·23

At the same time the table shows that both the mean length of life and the standard deviation vary so widely that it would have been impossible to plot these data on an axis of abscissæ divided arithmetically.

H. V. MUHSAM

Department of Statistics,
Government of Palestine,
Jerusalem.
July 14.

[1] Gaddum, J. H., *Nature*, **156**, 624 (1945).
[2] Curtiss, J. H., *Ann. Math. Statist.*, **14**, 107 (1943).
[3] Edgeworth, F. Y., *J. Roy. Statist. Soc.*, **61** (1898).
[4] Pearl, R., "Introduction to Medical Biometry and Statistics" (Philadelphia, 1930).