II. Note on the Surface of Constant Association.

By KARL PEARSON, F.R.S.

In the recent memoir by Dr Heron and myself a reference was made to the future publication of this note on the surface of constant association, *Biometrika*, Vol. IX, p. 315. The integral equation to the general surface was given, and also the particular form it took, still very complicated, for the simple case of total or 'marginal' frequencies being normal distributions. It was considered worth while to fully analyse one case of such a surface, namely that of $Q = 0{\cdot}6$ for Gaussian marginal frequencies. The numerical calculations were carried out by Miss Julia Bell, M.A., and from her ordinates of the sections the sections were plotted by Mr H. E. Soper, M.A. with the aid of a Coradi coordinatograph. By interpolation when needful Mr Soper constructed the isoplethes of this surface of constant association, and also an excellent card model, for comparison with the model of a normal surface. The chief features of this surface were referred to in the paper just cited. Although the marginal frequencies are symmetrical, the cross sections are skew-curves, the skewness increasing from zero for the central section to ·16 for the y-array when $x = 1{\cdot}5\sigma_x$ and to ·20 for the y-array when $x = 3{\cdot}5\sigma_x$.

Diagram I gives the series of sections on one side of the mid-section $x = 0$ up to $x = 3{\cdot}5\sigma_x$ by intervals of $0{\cdot}25\sigma_x$; the same sections are repeated in inverse order on the other side of the central section $x = 0$. It will be seen at once from the indicated means how skew the sections are.

Diagram II gives the isoplethes or contour-lines of equal frequency. They are approximately but not accurately ellipses with common principal axial directions, but they are very far indeed from being *similar* ellipses. In the contour corresponding to $z = \frac{9}{10}$ of its maximum value, the major axis is considerably more than twice the minor axis of the oval; in the contour corresponding to $z = \frac{1}{10}$, the major axis is very much less than twice the minor; the ovals tend indeed to less and less ellipticity. On this diagram are also plotted the regression lines of means and of modes. It will be seen that they tend to become parallel to the axis of x, or the regression tends to become zero. There is little doubt, that quite apart from normality of the marginal frequencies, any symmetrical marginal frequency would lead to like results, i.e. the isoplethes would not be similar curves, and the regression would be skew, and tend to asymptote to the horizontal. Thus the constancy of association would depend for its application on the existence of material in which the variates would be intimately related near their mean values and cease to have any relation towards extreme values. Intervening values would exhibit every variety of relationship, from the maximum in the neighbourhood of the means to zero value towards the extremes. These properties of the surface explain why Q is not even approximately constant, for those numerous surfaces of statistical practice in which the regression is approximately constant, i.e. linear.

Plates XXX, I (a) and (b), XXXI, I (c) give photographs of the actual model surface. In Fig. I (a) we see the surface 'end on' to one set of cross-sections and we grasp readily the skewness of the cross-sections and its increasing value as we pass from the central section. Fig. I (b) indicates by the verticals the regression line of means; these verticals have their feet on the regression line and the eye sees at once by their closer and closer approximation to each other, the deviation from linearity. Fig. I (c) gives a diagonal view of the surface of constant association, and underneath it Fig. II (d) has been placed a model of the Gaussian surface* of constant regression for comparison; the angle of deviation was taken roughly about 60°, to give a surface of correlation 0·5. Except for the regression line and skew-sections in Fig. I, the eye does not distinguish very readily in these photographs of Figs. I and II, the fundamental differences of the two surfaces which appear so markedly in the isoplethes of Diagram II.

* On the Brill system of interlaced sections.

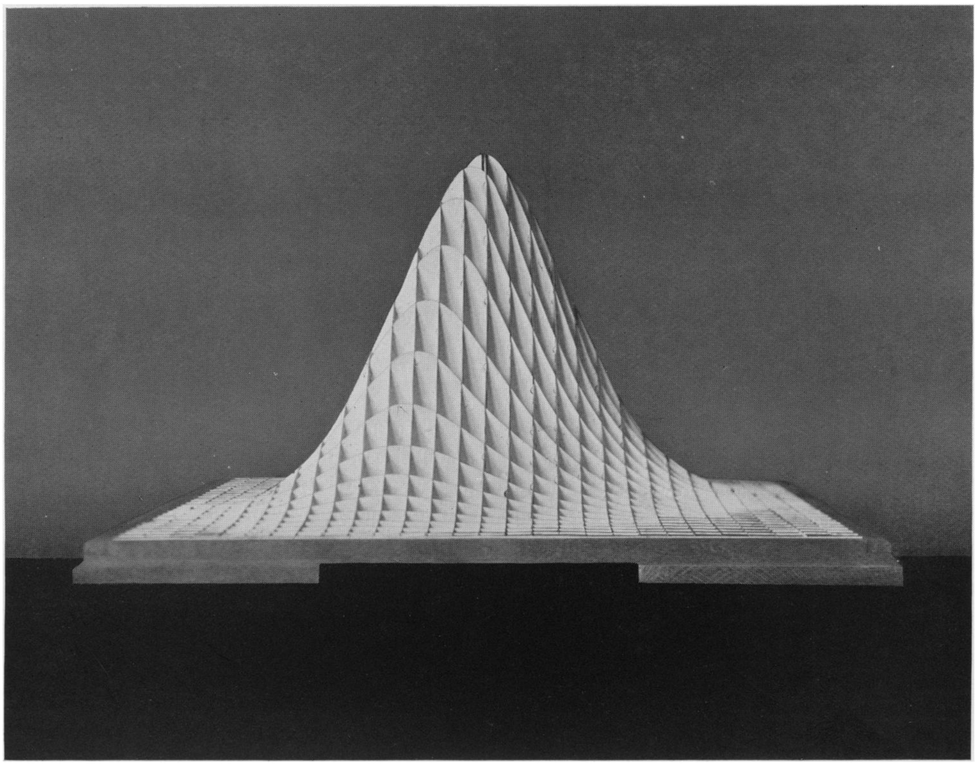


Fig. I (a). Surface of Constant Association.

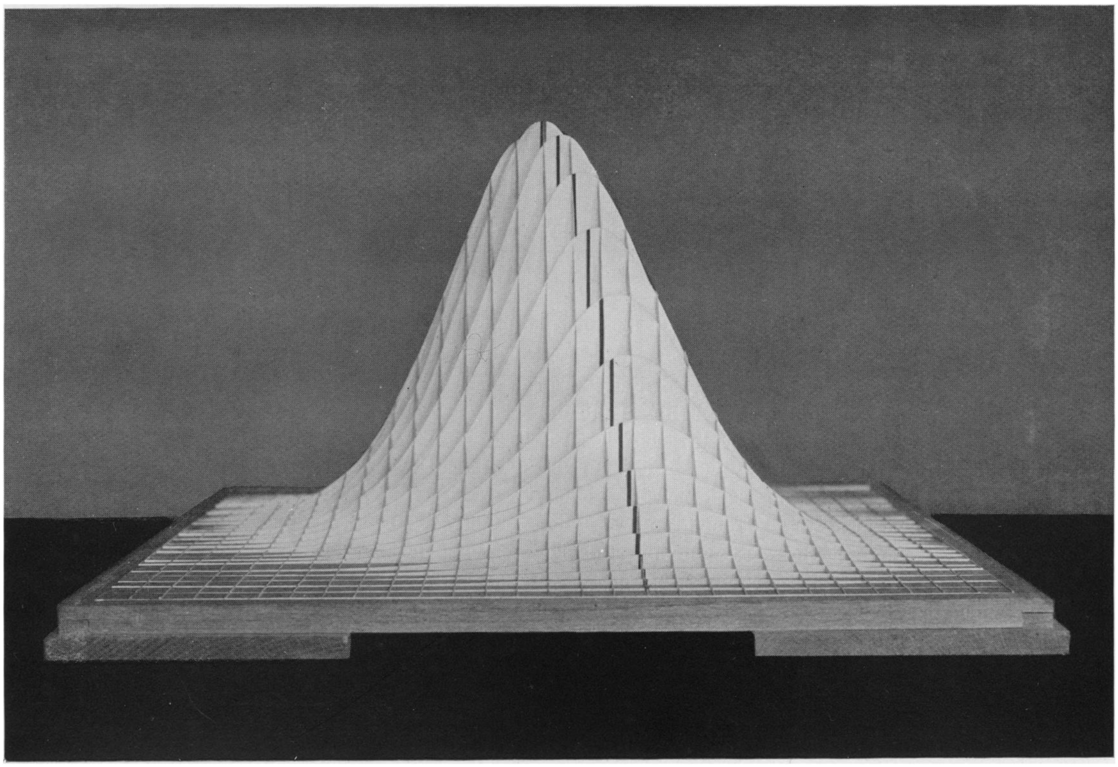


Fig. I (b). Surface of Constant Association.

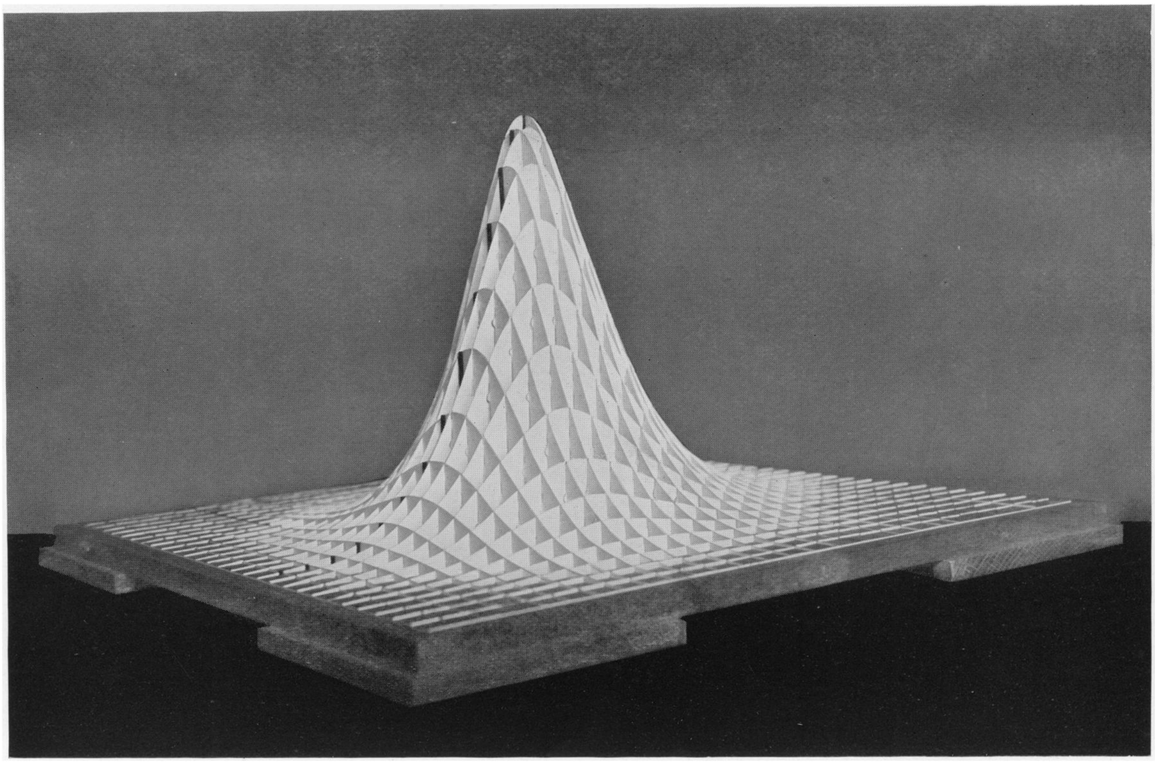


Fig. I (c). Surface of Constant Association.

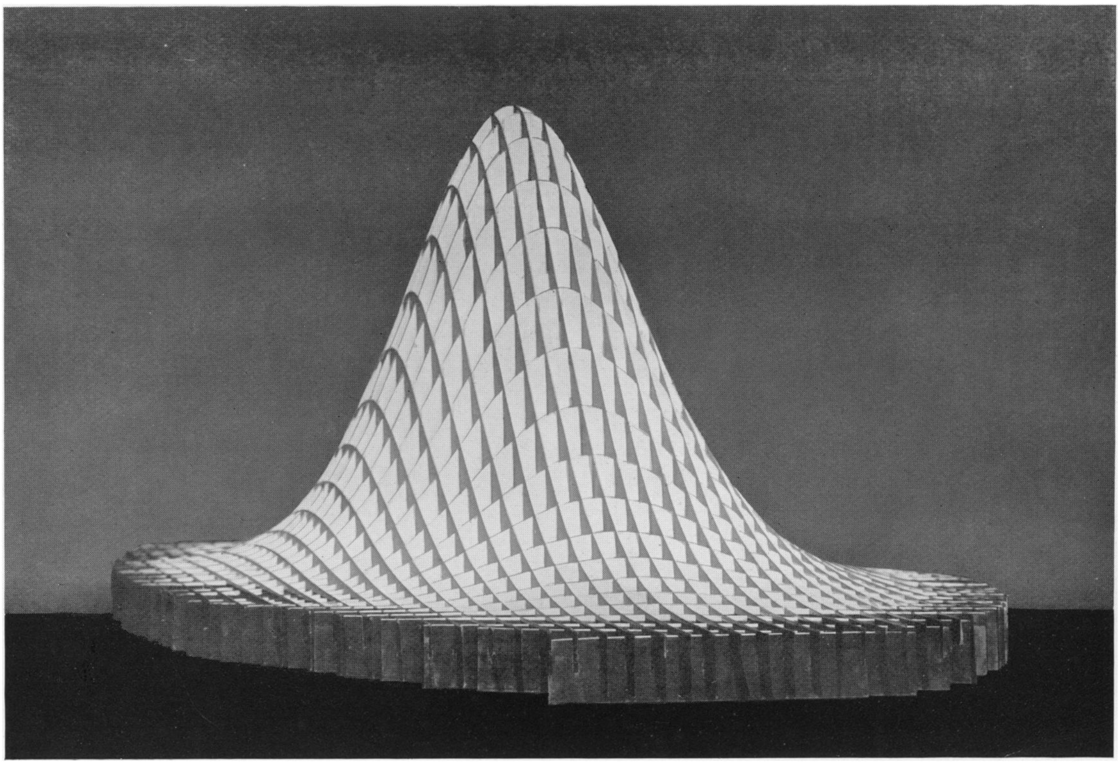
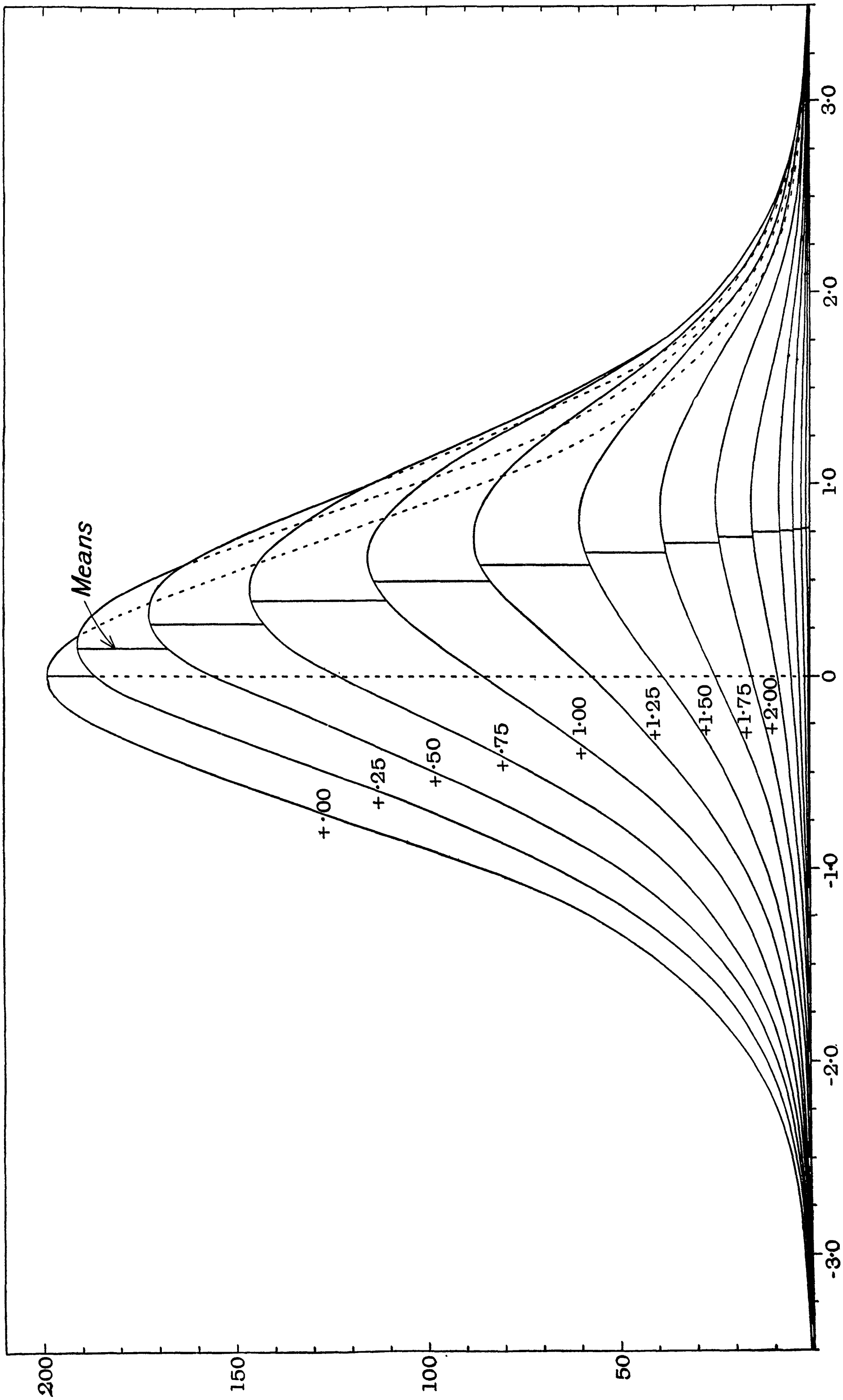


Fig. II (d). Brill type model of Gaussian surface.

DIAGRAM I. Surface of Constant Association, $Q=0{\cdot}6$.



Showing skew-sections at intervals of $x=0{\cdot}25\sigma_x$.

The equation to the surface of $Q=0{\cdot}6$ with normal marginal frequencies, for a total population of 1000, is:

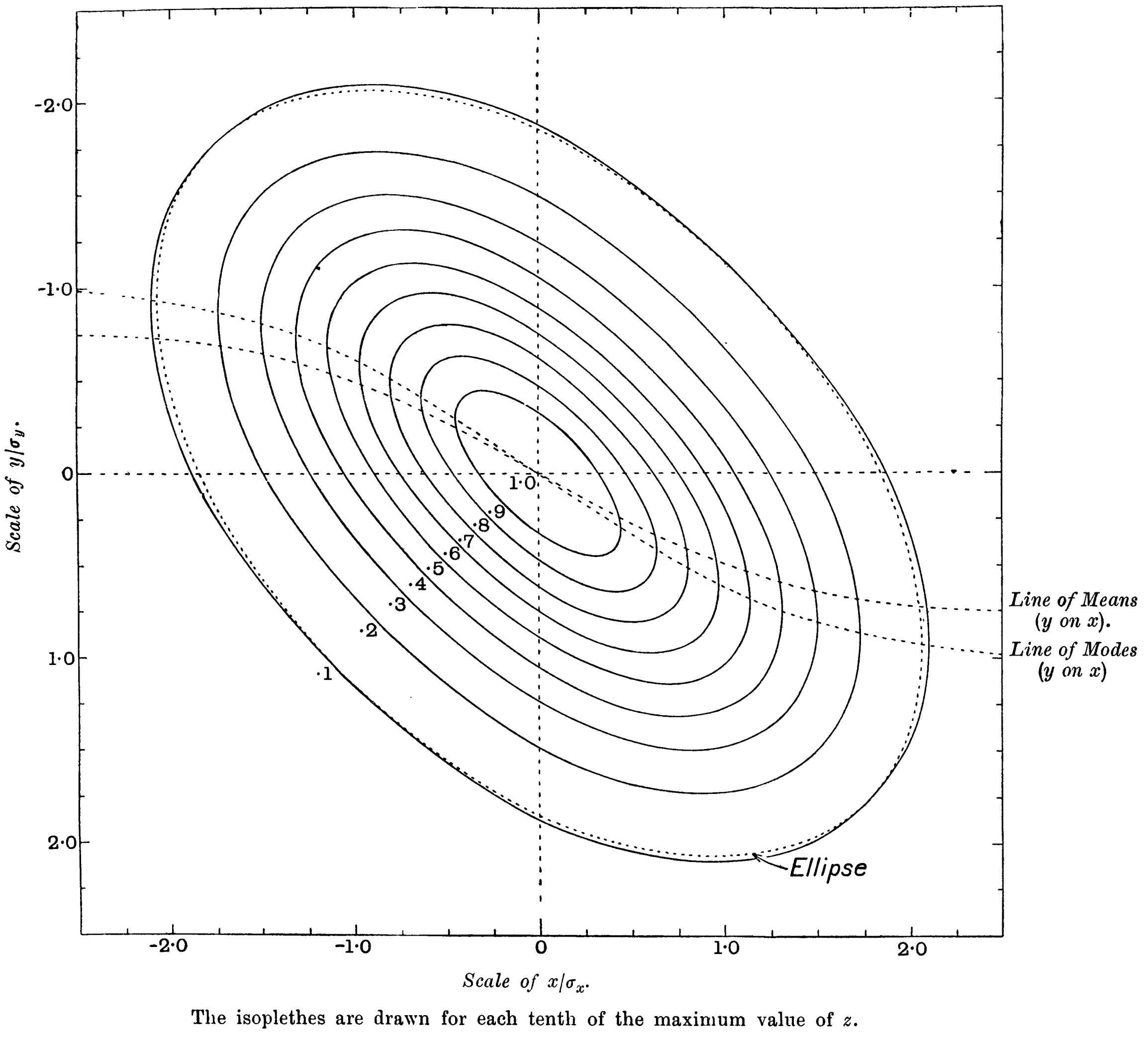
$$z=\frac{4000\frac{dp}{dx}\frac{dq}{dy}(1+3(p+q)-6pq)}{\{(1+3(p+q))^2-48pq\}^{\frac{3}{2}}},$$

where $p=\frac{1}{2}(1+a)$ for x/σ_x,

$q=\frac{1}{2}(1+a')$ for y/σ_y,

a having the usual significance attributed to it in Tables of the Probability Integral, e.g. Sheppard's.

DIAGRAM II. Isoplethes of Surface of Constant Association, $Q=0{\cdot}6$.



The isoplethes are drawn for each tenth of the maximum value of z.

The following short table will be sufficient to indicate the nature of the cross-sections to any-one wishing to pursue the matter further.

z-ordinates of y-arrays for given special values of x/σ_x.

Array at $x/\sigma_x =$

Values of y/σ_y	*0·0*	*0·5*	*1·0*	*1·5*	*2·0*	*2·5*	*3·0*	*3·5*
3·5	0·223	0·332	0·388	0·314	0·165	0·059	0·015	0·003
3·0	1·135	1·688	1·970	1·589	0·834	0·297	0·077	0·015
2·5	4·536	6·736	7·811	6·223	3·230	1·145	0·297	0·059
2·0	14·881	21·381	24·195	18·497	9·269	3·230	0·834	0·165
1·5	38·080	55·053	57·904	39·890	18·497	6·223	1·589	0·314
1·0	85·096	113·670	100·499	57·904	24·195	7·811	1·970	0·388
0·5	155·888	168·289	113·670	55·053	21·381	6·736	1·688	0·332
0·0	198·944	155·888	85·096	38·080	14·881	4·536	1·135	0·223
−0·5	155·888	97·562	49·567	22·036	8·414	2·644	0·662	0·130
−1·0	85·096	49·567	25·217	11·369	4·379	1·381	0·346	0·068
−1·5	38·080	22·036	11·369	5·186	2·010	0·635	0·159	0·031
−2·0	14·881	8·414	4·379	2·010	0·781	0·247	0·062	0·012
−2·5	4·536	2·644	1·381	0·635	0·247	0·078	0·020	0·004
−3·0	1·135	0·662	0·346	0·159	0·062	0·020	0·005	0·001
−3·5	0·223	0·130	0·068	0·031	0·012	0·004	0·001	0·000
Means	0·000	0·268	0·494	0·646	0·723	0·757	0·766	0·769

III.

Studies in the Meaning and Relationships of Birth and Death Rates. I. The Relationship between "corrected" Death-rates and Life Table Death-rates, by JOHN BROWNLEE, M.D., D.Sc. Journal of Hygiene, Vol. XIII. No. 2, pp. 178—190.

There are three "death-rates" used by those who deal with the statistics of public health etc.; the "crude" death-rate found by dividing the total number of deaths in a district in one year by the total number of inhabitants; the "corrected" death-rate found by applying the death-rates for age and sex to a standard population and calculating the rate from the figures so found; and the "life table" death-rate which is the rate that would be found by working out a complete stationary population from the death-rates for each age and calculating the ratio of the total deaths among the assumed stationary population to the total stationary population. This result is simply the reciprocal of the "expectation of life" at birth. Dr Brownlee thinks that this last measure is the most satisfactory death-rate and his paper is an attempt to reach approximate values for it from the corrected death-rates. Although Dr Brownlee does not set them out in that way these approximations appear to be based on the use of an imaginary stationary population or populations and require, we think, to be tested more extensively than has yet been done before they are used for any practical conclusions.

It appears to us that the use of the "expectation of life" at birth or of a single death-rate for a population as a whole cannot, however it is calculated, gauge completely the mortality of that population, and cannot therefore form an entirely satisfactory basis for comparative purposes: two populations may show the same death-rate or expectation of life and yet the mortality of one of them may, for instance, be heavy only in the first five years of life, while the other is light for the first five years but heavy at later ages. There must therefore be limitations in the use of any single figure, and if more than one figure is used we do not see why the death-rates at various ages should not be employed as they stand. They are easier to interpret and no more difficult to deal with than the expectation of life at a series of ages,—an alternative measure implied by Dr Brownlee.

In the paper before us Dr Brownlee mentions that further work is to follow, and possibly it will show how he intends to overcome these difficulties. W. P. E.