On the Use of Analytical Geometry to Represent Certain Kinds of Statistics.

By Professor F. Y. Edgeworth, M.A., F.B.A.

(*Continuation.*)

Section II.—The Method of Percentiles.

The Method which is next to be considered has in one respect the advantage over that which we have placed first. It is able to deal with rough imperfect data. The data proper to the method of moments are (a multitude of) single observations. The observations consist each of an integer number in the leading case typified by games of chance; for instance, the number of balls of a certain colour in a sample drawn at random from a bag containing balls of two or more colours (perfectly mixed). Observations relating to continuous quantity, space or time, seem to belong to a different type; but they may be, and for the purposes of Probabilities, I think, should be, reduced to the primary type. Measurements made as accurately as possible may be considered as integer numbers of units or degrees each as small as may be.* In practice, no doubt, we have to put up with much less perfect data. For instance, measurements of human statures are often given in (integer) inches; though, of course, the things measured do not vary thus *per saltum*. But though we have to accept such statistics, we regard their failure to realise the ideal type as an imperfection; we seek to rectify it as far as possible by adjustments and corrections such as Mr. Sheppard has applied. Contrariwise, the data proper to the method of *percentiles* are not single observations, not minimum strips of area bounded by the frequency-curve and two ordinates as close as may be, but integral blocks of area bounded by ordinates at a finite distance from each other. For example, the anthropometric measurements given by Baxter † in units of *two inches*, which are ill suited to the Method of Moments, are well suited to the Method of Percentiles. No doubt it is desirable that the blocks of area demarcated by percentiles should be as small as possible, the number of percentiles being as large as possible consistent with the validity of the method. But it will appear in the sequel that the number of percentiles utilised cannot be large *in relation to the total number of observations* consistently with the validity of the method. Where the

* Cp. "Law of Error," *Cambridge Philosophical Transactions*, 1905, Appendix I.

† Medical Statistics (United States War Department), cited in my "Methods of Statistics," *Jubilee Volume of the Journal of the Royal Statistical Society* (1885), p. 195, as presenting a slightly inaccurate value of the *modulus*.

number of observations is very large, over 300,000 in the last example cited, this restriction on the number of percentiles employed is not serious. But we have often to be content with a much smaller number of (perfectly independent) observations. In general there is a limit below which nothing is gained by increasing the number of percentiles and diminishing the size of the parcels into which they distribute the given total. Or rather, this much only is gained by the possession of further data, that by being able to break up the parcels into their constituent items we may be able to redistribute the items into a new set of parcels more favourable* than the one first coming to hand for the evaluation of the required constants. Still, the number of parcels on which we operate must be small in relation to the total, and may be small absolutely. The number may be as small as *five* even in the general case where there are *four* constants to be determined. If one of the constants is given—for instance, $\kappa = 0$, the group being evidently symmetrical—*four* percentiles only will be necessary. The efficacy of the method to deal with a few comprehensive data might be likened to the power exhibited by the historian when he draws just inferences from a few broad facts, making the most of the little evidence available. But, to complete the parallel, a writer thus gifted may be so wanting in other qualifications that he would not make much use of additional materials supposing that they became accessible. It is natural to presume that where a larger amount of information is intelligently utilised a more accurate result will be attained. The presumption, however, is not universally true. Genius may reconstruct some features of the past more faithfully from imperfect information than mediocrity exhaustively handling more abundant materials. Similarly, in the constructions with which we are concerned, it is possible that the worse materials may yield the better result. For the Method of Moments dealing with the more perfect materials deals with them in a way which, though good, is not the best possible, not being that which is prescribed by Inverse Probability.† Whereas the Method of Moments does employ Inverse Probability, and so makes the best use of its bad materials. However, the Method of Percentiles is not here put forward as superseding, but only as supplementing, the Method of Moments. In general, the method of moments is to be preferred; but the Method of Percentiles should be substituted when the data are too rough for the more refined method.

The new method may be introduced by describing it as a development of the procedure employed by Galton for determining the constants of a normal group, the centre and the "spread" or coefficient of dispersion (probable error, standard deviation or modulus). Galton had been preceded by Quetelet; and with reference to the use of the Median (to determine the centre of the group) by Laplace.

* See below, p. 743.

† Cp. *ante*, p. 304.

Laplace, in his so-called *Method of Situation*, had not only employed the Median to determine the sought central point, but also calculated the error to which this determination is liable.* Laplace's method of determining the error incident to the use of the Median has been extended to other percentiles.† From a determination of the error incident to the use of a percentile it might seem a short step to the error incident to the ascertainment of the constant of dispersion according to the practice of Galton and Quetelet. Thus, suppose that the quartiles having been observed, the *modulus* is ascertained by equating it to half the distance between the quartiles divided by the constant ·4769. . . . The modulus multiplied by twice this constant is equal to the difference between the abscissæ of the two quartiles; and thus the error of the modulus depends entirely on the errors of the *quartiles*. But in estimating the joint effect of those errors there occurs the difficulty that they are not *independent*. This difficulty has been overcome by Dr. Sheppard, who has given an expression for the error incident to the determination of the *Standard Deviation* from any number of percentiles.‡ By minimising this expression, and the corresponding expression for the error of the centre as determined by percentiles, Dr. Sheppard obtains formulæ for the use of percentiles with maximum accuracy. § Well, the problem here attacked may be regarded as an extension of that which Dr. Sheppard has solved. The given group which he contemplates may be considered as *translated* according to our conception from a normal error-curve. His constant " a " is what our constant " a " becomes when our constants κ and λ are zero.

A fuller explanation of the method may begin with the generally accepted hypothesis that a given frequency-group may be regarded as a sample taken from an indefinitely large group obtained by repeating observations under exactly similar conditions. This, as I understand, is the conception which is entertained by Professor Pearson when he speaks of " the fit of an observed to a theoretical frequency-" distribution "‖: a theoretical distribution which may indeed be treated as " known *a priori* " ; but " in a great many cases has " to be judged from the sample itself." This is the conception of a limit applying to any frequency-constant which I have elsewhere adduced¶ referring to the authority of Dr. Venn and Dr. Sheppard. The conception might be illustrated by an area subdivided into compartments in a plane on which raindrops are supposed to fall

* *Théorie Analytique des Probabilités*, supplement ii, and liv. ii, ch. iv, art. 23; *Mecamque Celeste*, liv. iii, art. 39, 40.

† By the present writer, *Philosophical Magazine*, 1886, vol. xxii, p. 375.

‡ *Transactions of the Royal Society*, 1898, vol. cxcii, A., p. 114, *et seq.*

§ " Maximum " rather than " greatest possible " if the views which I express below, p. 743, are correct.

‖ *Philosophical Magazine*, vol. L (1900), pp. 160, 164.

¶ *Encyclopædia Britannica*, eleventh edition. Article on " Probabilities," § 122.

sporadically. After a long continued rain the fall on each constituent compartment would be practically proportioned to its extent. But a slight shower, a mere sprinkling, would result only in an approximation to the distribution which would hold good in the long run. Thus, in Fig. 10, let the area enclosed between the abscissa and the curve (not shown) be unity: and let this area be broken up by five ordinates into six compartments of which the contents are specified. (They are proportioned to the number of observations between contiguous percentiles in a concrete example adduced below.) In the long run each section will receive a number of drops proportioned to its area. To contemplate the effects of a slight, or, more generally, a finite shower, let us substitute items which, unlike drops of water, continue separate and separable after having fallen, say grains of sand. Suppose sand-storms to have raged so chaotically that grains of sand are deposited with sporadic fairness on the areas enclosed between the curve and the abscissa. Of ten thousand grains we may expect that some 668 will be in the compartment furthest to the right, about 919 grains in the next compartment, and so on. The distribution of the ten thousand grains will be very nearly, but not in general, exactly this. If beginning from the extreme right we count 668 grains and draw an ordinate just to the left of the 668th grain, the apparent percentile thus presented will not usually coincide with the true percentile which is at the point 1·5375 measured from 0 or $m + 1{\cdot}5375$ if m is the abscissa of 0 (with reference to an origin on the left). If the apparent percentile lies to the *right* of the true one then there will have occurred an error in the distribution of frequency measured by the little strip of area which is contained between the ordinate erected at the true percentile (the point $m + 1{\cdot}5375$) and the ordinate (not shown) erected at the apparent percentile (the other boundaries of the strip being the frequency-curve—not shown in the figure—which represents the ideal distribution, and the interval in the abscissa between the true and apparent percentile). The compartment on the left of the true ordinate will be depleted to the extent of that strip: the compartment to the right augmented. The strip of area constituting this excess or defect is regarded as an *error*—say E_r pertaining to any the rth compartment—in relation to the ideal or hypothetically true distribution of N particles among the several compartments.

Upon some such fundamental hypothesis of a true distribution rests a construction which at successive heights involves the following three problems. *Firstly*, supposing that the ideal frequency-distribution being of the general character defined, a sample numbering N is taken at random—N grains fall sporadically on an area ruled in the manner described: what is the probability that there will occur the system of errors E_1, E_2, &c., in the areas of the respective compartment, the error E_r being the difference between the proportionate number of occupants which there would be in the rth stall if the number N was indefinitely large, say U_r, and the proportion which is actually presented, N being finite, viz., $U_r - E_r$? *Secondly*, supposing that the given distribution is the result of taking a sample N from an ideal distribution consisting of a *translated* error-curve with given constants m (for the position of the Median) and a, k, l (for the coefficients of the operator): what is the probability that there should occur an assigned set of percentiles, or, in other words, an assigned set of errors in the observed percentiles—the error of a percentile being the difference between its observed or apparent position, say x_r, and its true position, x_r'? *Thirdly*, supposing that the observed frequency-distribution has resulted from a sample of an ideal distribution consisting of a translated error-curve which is characterised by a certain system of values for the four constants m, a, k, l: what is that system of constants from which the observed distribution has most probably resulted—or which may best be put for the true system?

1. The first problem is an extension of the simplest problem in Probabilities: viz., if N balls are taken at random from an indefinitely large mixture of balls of two different colours, what is the probability of any assigned distribution of colours in the sample? We have merely to substitute for *two* colours *several* colours. Considering the number and complexity of the problems relating to games of chance which have been solved by the classical writers on Probabilities, it is remarkable that, as far as I know, they have not given a solution adapted to practice of this problem. Dr. Sheppard appears to have been the first to give the required solution: in his masterly article on the Normal Law of Error in the Transactions of the Royal Society for 1898. Professor Pearson, in the *Philosophical Magazine* for July, 1900, has given the solution in a very convenient form; which, according to my interpretation and notation, may thus be stated. When there are only two colours occurring with respective probabilities U and V ($U + V = 1$); the probability of a deviation $+E_1$, in the proportion of balls of the first colour from the ideal proportion U, and accordingly a deviation $E_2 = -E_1$ in the proportion of balls of the second colour is proportionate to

$$Exp - \tfrac{1}{2} N \left[\frac{E_1^2}{U} + \frac{E_2^2}{V}\right].$$

This expression is easily identified with the more familiar expression of the probability as an ordinate of a normal error-curve, of which the modulus is $\sqrt{2UV/N}$, the ordinate at the point distant E_1 from the centre. When there are many colours with probabilities (of being drawn) U, V, W . . . the probability of an assigned set of errors in the proportions of a sample numbering N, namely, the errors E_1, E_2, E_3 . . . is proportionate to

$$Exp - \tfrac{1}{2}N\left[\frac{E_1^2}{U} + \frac{E_2^2}{V} + \frac{E_3^2}{W} + \ldots\right].$$

The formula given by Professor Pearson, especially when it is approached as I have attempted to approach it* through the familiar example of differently coloured balls, carries attention to the important incident that, as in the case of two colours or classes, so in the more general case, the reasoning by which the approximate formula is established requires that none of the probabilities U, V, W... should be very small; that the product of the said probability and N should be large. In order that terms neglected should not be relatively large, it is requisite that $\frac{1}{\sqrt{NU}}, \frac{1}{\sqrt{NV}}$, &c., should be small fractions whose third and higher powers may be neglected. Accordingly, in employing our method, it boots not to use percentiles which demarcate portions of area so small that $1/\sqrt{NU}$ is a large fraction. Now if n is the (ideal) number of observations in any compartment $U = n/N$; and $NU = n$. Accordingly $1/\sqrt{n}$ must be a small fraction, say at most one-fifth or one-fourth; n must be at least about 20.† Thus, if we are given only 400 observations, we could hardly employ with advantage, I should think, more than *twenty* percentiles. Perhaps *ten* would give as good a result.

2. In general, according to a theorem of Laplace to which allusion has been made, we pass from an error in *area* of the type E_r to an error in (the position of) a percentile of the type $(x_r' - x)$ by equating E_r to the *algebraic* sum of two errors of the type $(x_r - x_r')Y_r$, $(x_{r+1} - x'_{r+2})Y_{r+1}$, where Y_r is the ordinate of the ideal frequency-curve at the point x_r (the errors being supposed small). In the particular case when the ideal frequency-group consists of a translated error-curve with given constants m, a, k, l, we have to express $(x_r' - x_r)$ and Y in terms of those constants. Now $x_r - x_r' = x_r - (m + X_r)$, where X is the abscissa of a point on the constructed curve relatively to the (true) median, and accordingly $X = a\xi_r + k\xi_r^2 + l\xi_r^3$; ξ_r being an abscissa of an error-curve with unit constant of dispersion. Also Y_r, the ordinate of the constructed curve, is equal to η_r the corresponding ordinate for the generating error-curve, divided by X_r', the differential coefficient of X_r, with respect to ξ_r. Thus (if, as usual, we put Δ prefixed to any function, *e.g.*, ($f(x)$, to designate $f(x+1) - f(x)$), it is proper to replace E_r in our first problem by the expression

$$\left[\Delta\,\frac{(x_r - x_r')\eta_r}{X_r'}\right].$$

* Congress Paper, sect. III.

† A similar reservation is required in the use of the Pearsonian criterion as I have noticed with reference to Dr. Greenwood's paper in the *Journal of the Royal Statistical Society*, 1914, vol. lxxvii, p. 198. But I have some hesitation about imposing a restriction which has not been observed by Professor Pearson and his followers. The rule that compartments with much less than twenty observations should not be utilised would materially alter the received verdicts of the Pearsonian criterion in certain cases; for instance, in the second and fifth examples given by Professor Pearson in his pathbreaking paper in the *Philosophical Magazine* for July, 1900.

Minus the sum of squares of expressions of this type, each multiplied by $\frac{1}{2}$N and divided by U_r, say $-$NT, forms the logarithm of the expression which multiplied by a proper constant represents the probability that the assigned system of percentiles should be presented by a sample numbering N taken from the given true frequency-distribution. For example, let the constants of the given distribution be $m = 0$ $a = 1$, $k = 0$, $l = 0$. In short, let the ideal distribution be a normal error-curve with the origin at the centre and with constant of dispersion, for which we shall here take the *standard-deviation*, equal to unity. Then the probability that a sample of 100 observations should present a distribution such as that represented in Fig. 13 would be, as calculated in a later page, proportional to $Exp - 100 \times \cdot 003$; or rather twice the integral of the error-curve.

$$\frac{1}{\sqrt{\pi}} e^{-\frac{1}{2}x^2} \text{ from } x = \infty \text{ to } x = \sqrt{\cdot 3}$$

gives the measure of the improbability attaching to the assigned occurrence—that is not a very great improbability. If the number of independent observations were 1,000 the improbability would be considerable, the odds against the event being less than 3 to 1,000.

3. It is a short step by way of Inversion from the probability that (an ideal frequency-distribution with) a given system of frequency-constants may produce an assigned set of percentiles to the probability that a given set of percentiles may have been produced by a certain assigned system of frequency-constants. Having obtained the expression for that probability for any assigned system of frequency-constants, we are in a position to determine according to principles set forth in a former Paper* that system of frequency-constants from which the observed set of percentiles most probably resulted, or which may best be taken for the true system of frequency-constants†. This system of values is given by equating to zero the differential coefficients of T with respect to each of the four coefficients m, a, k, and t, provided that the second term of development fulfils the condition of a maximum (for $-$T). The student of our former Paper will remember that the powerful instrument of Inverse Probability at one stroke not only elicits the most probable values of the *quæsita*, but also the probable error to which that determination is liable.

Built with the instruments, and resting on the foundation which has been described, the edifice which we are constructing presents as it were two wings: a diversity of construction according as the subject-matter consists of slightly abnormal, or considerably abnormal, frequency-groups. I suggest this *dichotomy* as appropriate to the present topic rather than the *trichotomy* employed in the first section. And I am disposed here to draw the line which demarcates the first class nearer to the limit of perfect normality than in the first section: to define the "slightly abnormal"

* *Journal of the Royal Statistical Society*, 1908, vol. lxxi, p. 388, *et seq.*
† Cf. *loc. cit.*, p. 386.

cases by the retention of the *first* powers only—not the first and second powers—of the variables κ and λ* which indicate divergence from normality.

Sub-section I.—Slightly Abnormal Curves.

By the definition of this class two of the variables, κ and λ, which enter into the expression T which is to be minimised, are small. But we do not enjoy the full advantage of dealing with small quantities as long as the two other variables a and m may be large. Now a may have any value. For a is a number, a *numeric* as some now say, which—multiplied by unity—gives the measure of what may be called the *quasi-modulus* of the constructed curve: what would be the modulus if the group were not only nearly, but perfectly, normal (for "modulus" here, as in other passages of this Paper, "standard-deviation" may be read *mutatis mutandis*)? And m—the abscissa of the Median—may well be larger than a. It may with great propriety be twice or three times as large as a, so as to secure that the origin from which the position of the percentiles is measured should lie entirely outside (the sensible portion of) the curve. I propose to remove these two impediments to the use of small quantities, as follows.

First, I take as the origin a point which may be described as the apparent Median adjusted. No adjustment is required to find the position of the apparent mean when the observations are given in detail, in the neighbourhood at least of the Median. We have merely, as explained above, to take on the abscissa a point on one side of which there are as many observations as on the other. But very generally the observations are not, and indeed cannot be, given in the required detail. There is given only the number of observations in a central compartment: for instance, in an example presently to be adduced, it is given that between 299·5-tenths of an inch and 300·5 there occur 329·5 observations of barometric height; the number of observations outside the compartment being such as to make it evident that the Median was between the points 299·5 and 300·5. The general practice which I propose may conveniently be introduced by applying it to this particular example. A first adjustment is the familiar practice of simple proportion. Whereas there are 1,380·5 observations *below*, that is in the arrangement adopted to the *right* of the point 299·5, and the total number of observations is 2,922, it follows that there must be 80·5 observations between 299·5 and the Median. Now there are 329·5 observations in the compartment between 299·5 and 300·5. Accordingly, *if the observations lie evenly* between those limits, if the frequency-curve be treated as in that neighbourhood a horizontal line, the Median may be taken to be at the point which

* Those who have read the first section will not need to be reminded that $\kappa = k/a$, $\lambda = l/a$.

divides the distance between those limits, namely, unity, in the ratio 80·5 : (329·5 — 80·5). The Median thus calculated proves to be 299·744. It is to be remarked that the condition italicised, though in general, with respect to other percentiles, not safely to be assumed, may be assumed with a minimum of inaccuracy when the *Median* is to be determined. For the normal frequency-curve is approximately horizontal in the neighbourhood of its centre; and the curve which is now to be constructed is approximately normal. Simple proportion then appears to be safe enough. But I propose to improve upon it by utilising the condition that the curve under consideration is translated from a normal curve. Suppose that the true Median is at a point O between 299·5 and 300·5; and let the sought distance from O to 299·5 be x_1; the distance from O to 300·5 being x_{-1} $(= 1 - x_1)$. Now $x_1 = a(\xi_1 + \kappa\xi_1^2 + \lambda\xi_1^3)$ where ξ_1 is the abscissa for the generating unit normal curve corresponding to the abscissa x_1, for the generated curve; and x_r is similarly related to ξ_{-1}. Now κ and λ are by definition small; and ξ_1, ξ_{-1}, if the central compartment is fairly small, are small; and therefore their second and third powers are *very* small. Therefore x_1 and x_{-1} and approximately proportioned to ξ_1 and ξ_{-1}. But ξ_1 and ξ_2 are the abscissæ of the error-curve with unit constant corresponding to the integrals (in Dr. Sheppard's notation) α =·0551 α =·1704; integrals obtained by dividing 80·5 (the number of observations between the Median and 2995) and 249 (the number of observations between the Median and 300·5) by 1461 (half the total number of observations) respectively. These abscissæ are easily determined with the aid of well-known tables, with particular facility by the use of the table which Mr. Sheppard has constructed giving the value of ξ (in our notation) corresponding to each assigned value of the (double) integral (his "α"). Thus, corresponding to the area or integral α = ·0551, the table shows ξ_1 to be ·0691. Similarly, ξ_{-1} is found to be ·2152. Dividing the interval between 299·5 and 300·5 in the ratio $\xi_1 : \xi_{-1}$ we obtain for ξ_1 ·244; and accordingly for the position of the Median 299·743, nearly identical with the value obtained by simple proportion, viz., 299·744.*

I propose to take the apparent Median thus adjusted as the true Median, at least for a first approximation. For a second approximation, when we take into account *second* powers of κ and λ, it will be shown how to re-introduce the variable m as a correction of the adjusted Median. Meanwhile, we shall take the adjusted Median as the origin, measuring from it the observed percentiles; the points x_1, x_2 . . . x_r and $_{-}x_1$, x_{-2} . . . x_{-r} given by observation such that the number of observations observed to occur between (ordinates drawn through) x_r and x_{r+1} is NU_r. The adjusted Median will thus form as it were the central column of

* If we have found approximate values of κ and λ it would be possible to utilise these values for dividing the tract which encloses the Median; but I am not sure that it would be worth the trouble.

our construction; it will play the same part as the Arithmetic Mean (the Centre of Gravity) in the Method of Moments. Let it not be hastily assumed that the Method of Percentiles is at a very great disadvantage in having to depend on the observed Median. For even with respect to the normal error-curve and frequency-curves in its neighbourhood the accuracy of the Median is not so very much less than that of the Arithmetic Mean. Theoretically, the probable error of the Median, in the case of a normal group, is only some 25 per cent. above that to which the Arithmetic Mean is liable. Now I hold with him who teaches that a difference of 25 per cent. or so in probable error is not a practically very significant difference.* Besides, we are not exclusively concerned in this Paper with frequency-curves which are in the neighbourhood of the normal error-curve. For the approximation effected in Subsection II with respect to the curves of considerable abnormality the Median as here obtained will be employed. But for frequency-curves considerably divergent from the normal in the direction of *lepto-kurtosis*, to use Professor Pearson's terminology, the probable error of the Median is *less* than that of the Arithmetic Mean. True, it is greater in the case of divergence in the opposite direction *platy-kurtosis*. But the two cases are not equally probable. For, as shown in our first section, the frequency-curve may well diverge to *any* extent in the former direction; but it can not, while preserving appearances, diverge very far in the latter direction. For example, in the case of symmetrical frequency-curves, the utmost permissible degree of *platy-kurtosis* is the limiting form shown in Fig. 9A, the rectangle. But the probable error of the Median in that extreme case exceeds that of the Arithmetic Mean only in the ratio $\sqrt{3}:1$. Whereas in the more frequent case of the divergence in the direction of *lepto-kurtosis*, the probable error of the Arithmetic Mean may exceed that of the Median to *any* extent. Such excess will be indeed rare. But it is quite within the range of ordinary experience that the probable-error of the Median should be *nearly* as small as that of the Arithmetic Mean. In fact, for the value of the coefficient λ, which may be considered on the limit of ordinary experience, namely ·2, the probable-errors of the two determinations are almost equal (in the case of symmetry).

The Median may even have an advantage in another respect. The observed Median admits of correction by means of the other observations. Though it is the point from which those observations are reckoned, the central pillar on which the other parts of the construction mainly rest, yet archwise it may derive some support from those other parts.† Whereas a similar correction of the Arithmetic Mean, though contemplated perhaps by Demorgan,‡ is not usual, nor I think easy.

* Rhind, *Biometrika*, vol. vii, p. 127.

† Thus, as here suggested (below, p. 744), the unknown small quantity *m* may be taken account of at the second approximation.

‡ *Calculus of Probabilities*, p. 136.

So far theoretically, on the supposition that the data have the degree of perfection suited to the Method of Moments. But as pointed out on an earlier page, that degree of perfection is not unfrequently wanting. Then the Method of Moment begins to be suspect. We may have to be content with the Method of Percentiles. It may add to our contentment to know that even if we had a choice there would not be so much to choose between the two species of average.

There remains one large quantity, namely, a. This remaining obstacle to the use of large numbers is easily removed by considering that a differs by a small quantity of the order of κ and λ from the constant above described as the "quasi-modulus," what the constant of dispersion would be if κ and λ were known to be zero. Or rather, it is preferable, as in the investigation of the true Modulus for a normal group,* to take as the first approximation the *inverse* quasi-modus. Thus, let $\frac{1}{a} = h + \alpha$ where a is a small magnitude of the order of κ and λ; all three may easily be evaluated after that the constant h has been ascertained.

To ascertain h we have to find that value thereof which makes $-$T a maximum when κ, λ (and m) are each zero. Now, in general

$$\mathrm{T} = \tfrac{1}{2}\mathrm{N}\Sigma\frac{1}{\mathrm{U}_r}\left[\frac{\Delta(x_r - \mathrm{X}_r)\eta_r}{\mathrm{X}_r'}\right]^2;$$

and in the particular case where κ and $\lambda = 0$, this becomes (the numerator and denominator of E_r, the fraction within square brackets, being each divided by a)

$$\tfrac{1}{2}\mathrm{N}\Sigma\frac{1}{\mathrm{U}_r}\left[h\Delta x_r\eta_r - \Delta\xi_r\eta_r\right]^2.$$

Put $\Delta\xi_r\eta_r = \mathrm{P}_r$ and $\Delta x_r\eta_r = \mathrm{Q}_r$,

and we have for the expression to be minimised (omitting the factor N/2) $\Sigma\,\frac{1}{\mathrm{U}_r}\left[\mathrm{Q}_r h - \mathrm{P}_r\right]^2$. This expression primarily obtained with reference to positive values of r, percentiles on the right of the adjusted Median, which is taken as origin, proves equally true of terms involving negative abscissæ x_{-r} and ξ_{-r} measured from the origin to the left.† From the abbreviated expression for T, differentiating it with respect to h, we at once find

$$h = \Sigma\frac{\mathrm{P}_r\mathrm{Q}_r}{\mathrm{U}_r}\Big/\frac{\Sigma\mathrm{Q}_r}{\mathrm{U}_r}$$

* Compare *J.R.S.S.*, vol. 71 (1908), p. 388 *et seq.*

† As we are concerned only with *squares* of the E's it is indifferent whether for P_{-r} we write $\xi_{-r}\,\eta_{-r} - \xi_{-(r+1)}\,\eta_{-(r+1)}$, or the *negative* thereof; and the sign of Q_{-r} is likewise conventional. I have used the former formulæ mostly in the tables, but in Table VII, having inadvertently used the latter, I have not thought it worth while to make an alteration for the sake of consistency. The *quaternary* quantities, the final coefficients, are not affected by our choice of signs for the P's and Q's pertaining to the negative compartments.

(all the given values positive and negative of r being taken); seeing that the second differential with respect to h is, as it ought to be for a minimum, positive.

Substituting this value for h in T, we have now a function of three small variables to be minimised. This is easily effected by expanding T in ascending powers of κ, λ, and α as far as terms of the second order inclusive, and then differentiating with respect to each of those variables separately. We thus obtain a system of three linear equations for κ, λ, and α.

Let us lighten the task of solving these simultaneous equations by introducing the variables one by one. And first let κ only be variable, λ and α being each zero. We have then for the expression to be minimised

$$\Sigma\frac{1}{U_r}\left[\Delta\eta_r\frac{x_rh - \xi_r - \kappa\xi_r^2}{1 + 2\kappa\xi_1}\right]^2.$$

Expanded in ascending powers of κ the expression within the square brackets, the error of area which we have called E_r, becomes $(Q_rh - P_r) - 1\kappa(\Delta 2x_rh\xi_r\eta_r - \Delta\xi_r^2\eta_r) + \kappa^2(\Delta 4x_r\xi_r^2\eta_r - 2\Delta\xi_r^2\eta_r) + \ . \ . \ .$; where Q_r and P_r have the meaning just now defined. Analogously to that definition I propose to call $x_r\xi_r\eta_r$, Q', and $\xi_r^2\eta_r$, P'; $x_r\xi_r^2\eta_r$, Q'', and $\xi_r^3\eta_r$, P''. It will be observed that $\xi_r\eta_r$ and $\xi_r^3\eta_r$, the makings of P and P'', become negative for negative values of r: whereas the makings of P'—denoted by an *odd* number of dots—remain positive, while r changes its sign. Similar propositions are true of the Q's. I propose to describe P, P', P'' . . . Q, Q', Q'' . . . as *secondary* quantities, in contradistinction to the *data* from which they are derived, the observed percentiles and corresponding areas, and the quantities immediately connected therewith, namely, ξ_r the abscissa of the unit error-curve corresponding to the percentile x_r, η_r the ordinate of the error-curve, and U_r the (proportionate) area of the rth compartment. Again, I propose to name the coefficients of E_r (the expression within square brackets pertaining to any the rth constituent of T) respectively,

$$A(= Q_rh - P_r),\ B(= 2Q_r'h - P_r'),\ F(= 4Q_r''h - 2P_r'');$$

and to describe them as the *tertiary* quantities. We have, then, for any the rth compartment $E_r^2 = A_r - \kappa B_r + \kappa^2 F_r$. Whence we obtain for the expression to be minimised

$$\Sigma\frac{1}{U_r}\left[A_r^2 - 2\kappa A_rB_r + \kappa^2(B_r^2 + 2A_rF_r)\right].$$

Whence $$\kappa = \Sigma\frac{1}{U_r}A_rB_r/\Sigma\frac{B_r^2 + 2A_rF_r}{U_r}.$$

The constituents of the numerator and denominator in the expression for κ may be described as *quaternary* quantities. They may be written respectively $\left(\frac{dT}{d\kappa}\right)_0$, $\left(\frac{d^2T}{d\kappa^2}\right)_0$; meaning that after differentiation zero is put for κ.

Now let us introduce one of the variables which have been abstracted, namely, α. Then E_r will be augmented by terms including α; becoming now

$$E_r = A_r - \kappa B_r + F_r\kappa^2, + Q_r\alpha - 2\kappa\alpha Q'_r.$$

Whence for the constituent of the expression to be minimised, we have

$$\Sigma\frac{1}{U_r}\Big[A_r^2 - 2\kappa A_rB_r + \kappa^2(B_r^2 + 2A_rF_r), \\ + 2\alpha Q_sA_r - 2\kappa\alpha Q_rB_r - 4\kappa\alpha Q_r'A_r + Q_r^2\alpha^2\Big]$$

Of the newly-introduced terms (marked off by a comma) the first vanishes, since $\Sigma\frac{1}{U_r}Q_rA_r = U$ is the equation for h. The coefficient of $-2\kappa\alpha$, namely, $\Sigma\frac{1}{U_r}Q_rB$, together with $2\Sigma\frac{1}{U_r}Q_r'A_r$, is small, of the same order. I think we may say as κ, or rather as κ/h,* since, like (the numerator of) κ, it is made up of positive and negative terms. Accordingly, the new equation introduced by taking α into account, namely,

$$\left(\frac{dT}{d\alpha}\right) = \alpha\Sigma\frac{Q_r^2}{U_r} - \kappa\Sigma\frac{Q_r'B_r + 2Q_r'A_r}{U_r} = 0,$$

while on the one hand it is easily utilised, is on the other hand not very useful. For, combined with the equation $\left(\frac{dT}{d\kappa}\right) = 0$, the new equation furnishes a correction of κ which is of the second order. But we are now seeking only a first approximation to κ: and it is a canon of mathematical elegance not to introduce some while omitting other corrections of a certain order. We may therefore, I think, postpone the introduction of α.

Before going further it may be well to enforce the portion of doctrine which has been stated by means of a concrete illustration. For this purpose there is required an example with a slight, but sensible, degree of asymmetry, and no, or very little, *kurtosis*. The first condition is well fulfilled by (most of the) statistics of barometric heights which Professor Pearson has compiled.† Let us select one of the groups which fulfils also the second condition. This will easily be effected if we find a case for which the coefficient β is nearly equal to ϵ, both small, but not very small. For then, as shown on an earlier page,‡ $\lambda = \epsilon - \frac{8}{9}\beta, = \frac{1}{9}\beta$ nearly, and so may be considered very small. These conditions

* For the purpose of the *first* Sub-section it might have been more elegant to put for $1/a$ not $h + a$ but $h(1 + \theta)$; so that θ a small *numeric* would be comparable with κ.

† See *Transactions of the Royal Society*, 1897, vol. 190 A.

‡ *Ante*, p. 316.

appear to be fulfilled by the statistics of barometric heights at Churchstoke.* For this group the Pearsonian constants β_1 and η are respectively ·12578 and ·18891. Accordingly our constant $\beta\left(=\frac{1}{8}\beta_1\right) = {}$·01572, and our constant $\epsilon\left(=\frac{1}{12}\eta\right) = {}$·01574, $\lambda = \frac{1}{9}$·0157 = ·0017; while $\chi = \frac{4}{9}\beta =$ ·0069. Whence $\kappa = \sqrt{\chi} =$ ·083. That is the value of κ when the operand, the generating error-curve, has unity for its *modulus*. But, if the *standard deviation* is taken for the unit, it is proper to divide the value of κ, which has been found by $\sqrt{2}$.† In the calculations which follow I shall employ the *standard deviation*, not as essentially preferable, but in order to avail myself of Dr. Sheppard's Tables.

Now, assuming that $\lambda = 0$, let us apply the Method of Percentiles to the data, or rather to a selection of them. For this purpose, having collected and rearranged the frequencies for each degree (of barometric height) as given by Professor Pearson, I take for the positions of the percentiles which are to be utilised the points on the abscissa corresponding to 275·5 (tenths of an inch), 297·5 302·5, together with the Median as determined on an earlier page,‡ viz., 299·744. The distance of the Median from the other percentiles gives us the observations designated x. The corresponding values of the area between each percentile and the Median is obtained by counting the number of observations occurring between those limits and dividing that number by half the total number. We thus obtain that integral which Dr. Sheppard designates "α" and takes as the argument of his convenient Table II.§ By means of that table we obtain for each of the utilised percentiles our ξ (his x) and our η‖ (his z). We thus obtain the *primary* quantities which are exhibited in the first part of Table VII.

TABLE VII.

Showing the evaluation of small constants in a case where there is no *kurtosis*.

PART I. PRIMARY QUANTITIES.

Abscissa or compartment.	x.	$\frac{1}{2}\alpha$.	ξ.	η.	U.
− 1	2·757	·3015	·84699	·278696	·1985
− 0	0	0	0	·39894	·3015
+ 0	0	0	0	·39894	·2264
+ 1	2·243	·2264	·60196	·33283	·1487
+ 2	4·243	·3751	1·15084	·205737	·1249

* *Loc. cit.*, p. 433.

† *Ante*, p. 306.

‡ Above, p. 729.

§ *Biometrika*, vol. II. I have utilised the first and second but not the third differences.

‖ This "η" has nothing to do with the Pearsonian "η" proper to the Method of Moments.

TABLE VII—*contd.*

PART II. SECONDARY QUANTITIES.

	Abscissa or compartment.	$\xi\eta$.	P.	$\xi^2\eta$.	P′.	$\xi^3\eta$.	P″.
A	− 1	− ·23605	+ ·23605	·1999	− ·1999	− ·16932	+ ·16932
	− 0	0	− ·23605	0	+ ·1999	0	− ·16932
	+ 0	0	·20035	0	·12060	0	·0726
	+ 1	·20035	·03642	·12060	·15188	·0726	·2410
	+ 2	·23677	− ·23677	·27248	− ·27248	·31358	− ·31358

	Abscissa or compartment.	$x\eta$.	Q.	$x\xi\eta$.	Q′.	$x\xi^2\eta$.	Q″.
B	− 1	−·76836	+·76836	·65079	− ·65079	−·55121	+ ·55121
	− 0	0	−·76836	0	+ ·65079	0	− ·55121
	+ 0	0	·74654	0	·44936	0	·2705
	+ 1	·74654	·12640	·44936	·55525	·2705	·8856
	+ 2	·89294	−·87294	1·0045	−1·0046	1·15613	−1·15613

	Compartment.	PQ.	PQ/U.	Q^2.	Q^2/U.
C	− 1	·18137	·9137	·590377	2·974
	− 0	·18137	·6015	·590377	1·958
	+ 0	·14957	·661	·557328	2·462
	+ 1	·00460	·031	·15977	·107
	+ 2	·20669	1·655	·762024	6·101
	Sums	....	3·862	....	13·602

PART III. TERTIARY QUANTITIES.

Compartment.	A.	B.	F.
− 1	− ·01784	− ·1697	+ ·2875
− 0	+ ·01784	+ ·1697	− ·2875
+ 0	·01167	·13464	·1621
+ 1	− ·0005	− ·1635	·5241
+ 2	− ·01114	− ·29812	− ·68620

See note to p. 734 as to the signs of P, P′ . . . Q, Q′ . . .

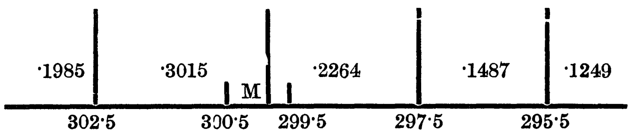
TABLE VII—*contd.*

PART IV. QUATERNARY QUANTITIES.

Compartment.	$-\left(\frac{dT}{d\kappa}\right)_0$	$\left(\frac{d^2T}{d\kappa^2}\right)_0$	$-\left(\frac{d^2T}{d\kappa d\alpha}\right)_0$
− 1	·0152	·093	− ·355
− 0	·0104	·061	− ·355
+ 0	·0069	·097	·49
+ 1	−·0006	·180	·135
+ 2	·0266	·834	1·904
Sums	·059	1·265	1·82

Two of the primary quantities, the x's (by implication) and the U's, are exhibited in Figure 10.

FIG. 10.



To derive the *secondary* quantities we first form $\xi_r\eta_r$ and $x_r\eta_r$ for each percentile, and then multiply each of these products by successive powers of ξ_r; attending carefully to the *signs* of the x's and the ξ's. The *italicised* figures in Part II of the table show the results of these multiplications; from which by a simple subtraction we derive the *secondary* quantities, the P's and Q's shown in Part II, A and B, of the table. In Part II C is shown the calculation by which h is determined. From h, together with the P's and Q's, we derive the *tertiary* quantities, the A, B, and F, shown in Part III of the table.

Part IV shows the *quaternary* quantities from which the value of κ is immediately derived. It proves to be ·059/1·265 = ·05 nearly; a result which agrees fairly well with that above obtained by the Method of Moments (with reference to *standard deviation*), namely, ·06.

By taking account of α we may obtain a correction for κ, while verifying the proposition that the correction is not very important. For the coefficient of $-2\kappa\alpha$ in the expression for T, viz., $\Sigma(Q_rB_r + 2Q_r'A_r)/U_r$, I find from the first three parts of the table 1·82. And the coefficient of α^2 (in T), namely, $\Sigma\frac{Q_r^2}{U_r}$, is as shown in Part II C of the table, 13·6. Whence $\alpha = ·134\kappa$. Also from the equation $\frac{dT}{d\kappa} = 0$ we have $-1·82\alpha + 1·265\kappa = ·059$. Eliminating α we have for the corrected value of κ, ·058, almost identical with the value obtained by Moments, viz., ·059.

Let us now introduce the variable λ, at first by itself. We may now write the content of the square brackets in any constituent of the expression for T

$$E_r = A_r - \lambda C_r + \lambda^2 H_r ;$$

where A has the signification already defined; $C = 3Q''h - 2P''$; $H = 9 Q'^r h - 6 P'^r$; $Q_r'^r$ and $P_r'^r$ being devised from Q_r and P_r by continued multiplication of ξ_r after the analogy of Q', Q'', and P', P''. We may now proceed at once to evaluate λ by squaring the last-written expression for each element, dividing by U_r, summing up all the constituents of T, and differentiating the sum with respect to λ. But the value so obtained for λ would not be as good as the value similarly obtained for κ. For in the case of λ it is found that α cannot safely be ignored. The coefficient of $\alpha\lambda$ in the expression for T, unlike that of α, proves not to be small for small values of λ, but substantial. The expression to be squared for each constituent is now—

$$A - \lambda C + \lambda^2 H, + Q\alpha - \alpha\lambda 3Q''A,$$

whence squaring, dividing by U_r, summing, and differentiating by λ and by α, we obtain the simultaneous linear equations (the absolute term in the first equation vanishing as before)—

$$\Sigma \frac{Q_r^2}{U_r}\alpha - \lambda\Sigma\frac{Q_rC_r + 3Q_r''A_r}{U_r} = 0$$

$$-\alpha\Sigma\frac{Q_rC_r + 3Q_r''A_r}{U_r} + \lambda\frac{C_r^2 + 2A_rH_r}{U_r} = \Sigma\frac{A_rC_r}{U_r}.$$

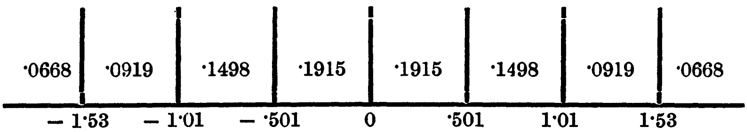
The values of α and λ are easily found from these equations.

What further explanation of the method is required may be conveyed by working an example. Unfortunately none of the concrete symmetrical groups with which I am acquainted are quite suited for this purpose. One is too nearly normal, so that the λ coefficient elicited is almost invisible; another is so far from normal as to raise a doubt whether it does not belong to Sub-section II. Accordingly, I have thought it best to construct a fictitious example. I have "translated" a normal error-curve with standard deviation unity by putting for the abscissa of the constructed curve $X = \xi + \cdot 01\, \xi^3 \ldots$. We have thus the coefficient $a = 1$, $\lambda = \cdot 01$. Supposing that these coefficients have to be discovered from observed percentiles of the constructed group; let us take for the percentiles to be utilised the values of x (the abscissa of the observed frequency-group or in Professor Pearson's terminology "histogram"), the value, which correspond to the round values of ξ (the abscissa of the generating normal curve) $\pm \cdot 5 \pm 1 \pm 1\cdot 5$. The observed values of x on this supposition are :—

$$x_1 = x_{-1} = \cdot 5 + \cdot 01 \times \cdot 5^3 = 0\cdot 50125$$
$$x_2 = x_{-2} = 1 + \cdot 01 \times 1^3 = 1\cdot 01$$
$$x_3 = x_{-3} = 1\cdot 5 + \cdot 01 \times 1\cdot 5^3 = 1\cdot 53375.$$

From Dr. Sheppard's Table I we easily find for each value of our ξ (his x) the corresponding value of our η (his z), and of twice the area intercepted between x and the centre (his α). Our U's are immediately obtained from (the halves of) his α's. Figure 11 shows the x's and U's. The values supposed to be the

FIG. 11.



primary data, are shown in Part I of Table VIII. Part II of the Table shows the *secondary* quantities, the P's and Q's; which are obtained with less trouble than in the case of an asymmetrical group, since the P's and Q's with which we have now to deal have all

TABLE VIII.

Showing the evaluation of small constants in a case where there is no asymmetry.

PART I. PRIMARY QUANTITIES.

Abscissa or compartment.	ξ.	x.	η.	$\frac{1}{2}\alpha$.	U.
± 0	0	0	·39894	0	·1915
± 1	·5	·50125	·35206	·19146	·1498
± 2	1	1·01	·24197	·34134	·0919
± 3	1·5	1·53375	·12952	·43319	·0668

PART II. SECONDARY QUANTITIES.

	Abscissa or compartment.	$\xi\eta$.	P.	$\xi^3\eta$.	P″.	$\xi^5\eta$.	P^{IV}.
A	± 0	0	·176	0	·044	0	·011
	± 1	±·176	·066	±·044	·19797	±·011	·23097
	± 2	±·24197	·0477	±·24197	·19509	±·24197	·74141
	± 3	±·19425	−·19425	±·43706	−·43706	±·98388	−·98338

	Abscissa or compartment.	$x\eta$.	Q.	$x\xi^2\eta$.	Q″.	$x\xi^4\eta$.	Q^{IV}.
B	± 0	0	·17644	0	·04411	0	·01103
	± 1	±·17644	·06795	±·04411	·20028	± ·01103	·23336
	± 2	±·24439	−·04577	±·24439	·20250	± ·24439	·76112
	± 3	±·19862	−·19862	±·44689	−·44689	±1·0055	−1·0055

TABLE VIII—*contd.*

PART II. SECONDARY QUANTITIES—*contd.*

Compartment.		PQ.	PQ/U.	Q^2.	Q^2/U.
C	± 0	·03105	·1622	·031131	·1625
	± 1	·00448	·0299	·004617	·0308
	± 2	·002184	·0238	·002095	·0228
	± 3	·038581	·5776	·039446	·5905
		....	·7935	....	·8066

PART III. TERTIARY QUANTITIES.

Compartment.	A.	C.	H.
± 0	− ·00243	·04218	·03166
± 1	+ ·00085	·19513	·6803
± 2	+ ·00267	·20745	2·2903
± 3	− ·00114	− ·44477	− 3·00217

PART IV. QUATERNARY QUANTITIES.

Compartment.	$-\left(\frac{dT}{d\lambda}\right)_0$	$\left(\frac{d^2T}{d\lambda^2}\right)_0$	$-\left(\frac{d^2T}{d\lambda d\alpha}\right)_0$
± 0	− ·0005	·0085	·037
± 1	+ ·0011	·262	·092
± 2	+ ·0059	·601	− ·086
± 3	+ ·0076	3·06	1·35
Sums	·014	3·93	1·4

an *even* number of dots. Inserting the P's and Q's into the proper formula, we elicit the value of h, viz., ·98375. Combining this value with those of the P's and Q's we calculate the *tertiary* quantities A, C, and H, shown in Part III of the table. From these we derive the *quaternary* quantities which are the values of

$$\left(\frac{dT}{d\lambda}\right)_0, \left(\frac{d^2T}{d\lambda^2}\right)_0, \left(\frac{d^2T}{d\lambda d\alpha}\right)_0.$$

Differentiating the expression for T with respect to λ and α separately we obtain for these variables the simultaneous equations—

$$\cdot 8066\alpha - 1\cdot 4\lambda = 0$$
$$-1\cdot 4\alpha + 3\cdot 93\lambda = \cdot 014.$$

Solving these equations we find $\lambda = \cdot 009$ $\alpha = \cdot 0155$; good approximations to the true values (known in this case), namely, $\lambda = \cdot 01$, $\alpha\ (= 1 - \cdot 98375) = \cdot 01625$.

It will occur to the expert in Probabilities who may cast his eye over this calculation that the result depends in an undue degree upon one item of the data, one strand of the coil, that which pertains to the outmost double compartment, designated by the subscript ± 2. The character of a *republic* proper to the constituents in the calculation of a frequency-constant seems wanting. What has happened is, I think, that the set of percentiles utilised is one not very favourable for the accurate determination of the sought coefficients. *Any* set of percentiles, if not too few or too numerous, supply data adequate for the determination of the coefficients by Inverse Probability. But the determination is not equally good for all sets of percentiles; the "probable errors" to which the result is liable are greater or less according as one or other set of percentiles is utilised. The matter is particularly simple where we have to deal with only one coefficient, our constant a, which we may regard as the *standard deviation*. Taking the *reciprocal* of a, say as before h, for the *quæsitum*, and determining it by Inverse Probability as applied in an earlier page and explained in an earlier Paper,* we find that the errors to which the determination is liable fluctuates according to a normal error-curve of which twice the mean square of error (= the square of the Modulus) is the reciprocal of $N\Sigma Q_r^2/U_r$.† This measure of inaccuracy becomes smaller the larger N is, and the larger the sum of squares which is multiplied by N. Given the number of observations, and using a number of percentiles consistent with the number of observations, it is desirable to select the percentiles so that the sum of terms of the type Q_r^2/U_r should be as large as possible. Starting with the set of percentiles which comes first to hand, we may improve upon it by a series of somewhat laborious tentatives. When there are two or three *quæsita* (a, κ, λ) the arithmetical labour is aggravated by a philosophical difficulty. But it will be time enough to prescribe for the selection of the best set of percentiles when the proposal to use percentiles has been accepted by experts.

I likewise postpone the full exposition and exemplification of the general case which subsumes the particular cases which have

* *J.R.S.S.*, 1908.

† It may excite suspicion that this expression for the mean square of the error incurred should differ from that which has been given by Dr. Sheppard in his classical treatment of the subject (*Transactions of the Royal Society*, 1898A, vol. 192, pp. 131, 132). The explanation of the difference is, I think, that he has employed for the determination of the standard-deviation from observed percentiles a rule which, though a good rule, and, it may be, practically the best, as being the simplest, rule, is not the theoretically best rule—that which is given by Inverse Probability. Between Dr. Sheppard's and our formula for the Standard Deviation, is there not the same relation as between Galton's simple formula for the determination of the *correlation-coefficient* (in the case of two variables) and the less simple, but more accurate formula which Professor Pearson obtained by a stroke of Inverse Probability (*cp.* *J.R.S.S.*, 1908, vol. lxxi, p. 395)?

been treated, the case in which both κ and λ as well as α are sought variables. Suffice it to say that in the general case the error in area E_r receives an additional term $\kappa\lambda G_r$, where $G_r = 12\,Q_r''' - 7\,P_r'''$; becoming $A_r - \kappa B_r - \lambda C_r + F\kappa^2 + G\kappa\lambda + H\lambda^2$ Whence the quantity which is to be maximised,

$$-T = \tfrac{1}{2}N\Sigma\frac{1}{U_r}\Big[A_r^2 - 2\kappa A_rB_r - 2\lambda A_rC_r + \kappa^2(B_r^2 + 2A_rF_r) + \lambda^2(C_r^2 + 2A_rH_r) + \alpha^2Q_r^2 + 2\kappa\lambda(B_rC_r + A_rG_r) - 2\lambda\alpha(QC + 3Q''A)\Big]$$

—the term $\kappa\alpha$ with its coefficient being for a reason before given. Differentiating with respect to κ, λ and α separately, we obtain the three simultaneous linear equations for those variables.

Good examples of the general case may be found among the statistics of barometric heights to which reference has been already made. I have begun the work for one of them, the one which I long ago used to exemplify simpler species of translation,* the group of barometric heights at Babbacombe. By method of interpolation before practised,† as well as by the more refined adjustment now proposed, I find for the *Median* 300 (tenths of an inch). For the other percentiles to be utilised I take 295·5, 297·5, 302·5, 034·5. Counting the numbers of observations intercepted by these points, as stated in my earlier Paper,‡ I have obtained with the aid of Dr. Sheppard's tables the *primary* quantities. From a rough approximation to the *secondary* and subsequent classes of coefficients, I am led to anticipate for the calculation if accurately performed a plausible result, one in accordance with previous determinations. I trust that the work will be completed by some more industrious and accurate arithmetician.

It is a nice question whether there should be added to this subsection a praxis for the determination of *second* approximations to the values of κ, λ, and α. With their corrections of the second order, say, $\delta\kappa$, $\delta\lambda$, might be ranged m, the correction of the adjusted Median presumed to be of the second order. We should expand each constituent of the type E_r so as to include (before differentiation) quantities of the *third* order, such as κ^3, $\kappa^2\lambda$, $\kappa\delta\kappa$, $\kappa\delta\lambda$, κm . . . The new small quantity m will enter by substituting for each positive abscissa x $(x_r + m)$ (supposing that the apparent Median which is to be corrected is on the *right* of the true Median) and for each negative abscissa—x_r, $-(x_r - m)$. But I doubt whether it is worth the trouble to institute so laborious a calculation, appropriate only to a particular degree of abnormality. I proceed to the more general case of considerable abnormality.

* See *J.R.S.S.* "Mathematical Representation of Statistics," 1898, p. 680 *et seq.*

† *Loc. cit.*, p. 698.

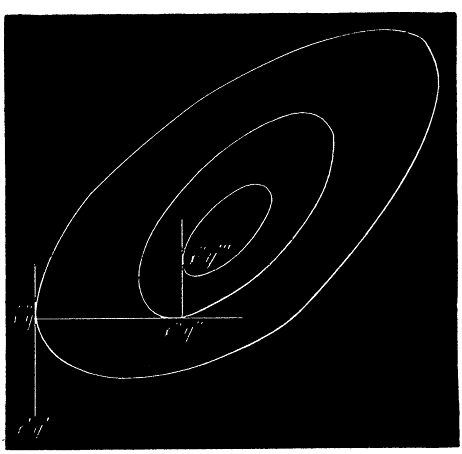
‡ The observations given by Pearson, *Phil. Trans.*, 1897, vol. 190, A, p. 433.

SUB-SECTION II.—CONSIDERABLY ABNORMAL CURVES.

The class of frequency-groups which is now to be considered is characterised by constants of such magnitude that the expansion in ascending powers of the function which is to be maximised is no longer available. The subject may be approached by way of the following lemmas.

Lemma 1. It is required to find an approximation to the values of x and y which make $f(xy)$ a maximum; when f is such a function that if $\left(\frac{df}{dx}\right)$, say f_1, and $\left(\frac{df}{dy}\right)$, say f_2, are simultaneously equated to zero the solution of the system would be impracticably laborious, but the solution of one of the derivative functions, *e.g.*, $f_1(xy)$ equated to zero for the corresponding variable x, the other variable being treated as constant, is not impracticable. (For example, the conditions would be fulfilled if $f(xy)$ were a rational algebraic function of the fourth degree.) To obtain a rough approximation I propose the following process. Considering $z = f(xy)$, as the equations of a surface, find a point $x'y'$ in the neighbourhood of a maximum value of z, or at least at a finite distance therefrom. Substituting one of the co-ordinates, say x', in the function derived by the differentiation with respect to the other variable, *i.e.*, in the case supposed $f_2(xy)$, solve the equation $f_2(x'y) = 0$ for y, and obtain thereby a maximum value of z (relative to x' constant), which is greater than $f(x'y')$, say $f(x'y'')$. Now put y'' constant in the equation $f_1(xy'') = 0$; and, solving for x, obtain a value of z, $f(x''y'')$, which is greater than $f(x'y'')$. And so on. Fig. 12 is designed to illustrate the process of climbing up hill, so to speak. The successive curves are members of the

FIG. 12.



family z = constant; one of which is *touched* at each stopping-point by one of the co-ordinates of that point. The method, enounced with reference to two variables, is capable of extension to functions of three or more variables.

Lemma 2. The other conditions being the same, let the function f be such that the equations $f_1(xy') = 0$, $f_2(x_2'y) = 0$, are both —or at least one of them—impracticable. Starting from a suitable point $x'y'$, observe the sign of one of the derivate functions, say $f_1(x'y')$; and take a step Δx (the direction corresponding to the function selected, say f_1), such that Δx (with its sign) multiplied by $f(x'\,y')$ is positive. Putting x'' for $x' + \Delta x$ observe whether $f_1(x''y')$ is greater than $f_1(x'y')$. If not, take a shorter step (from the point $x'y'$). If $f_1(x''y') > f_1(x'y')$, either take another (and another) step parallel to the axis of x; or take a step parallel to the axis of y, observing the sign of f_2, and taking Δy of such a sign that $\Delta y f_2$ may be positive. And so on. That is supposing *both* the equations $f_1(xy') = 0$, $f_2(x'y) = 0$ to be impracticable. But, if one of them, *e.g.*, $f_2(x'y) = 0$, is soluble, it may be used to determine the *extent*—as well as the direction—of a step parallel to the axis of y. Like the first lemma, the second lemma may be extended to three or more dimensions.

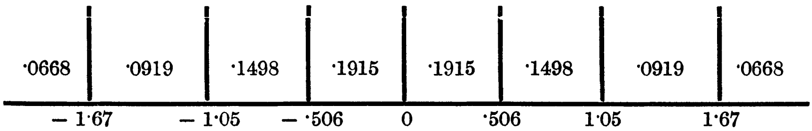
That variety of the second lemma, in which there are *three* independent variables, the derived function relative to *one* of them being manageable, is in general appropriate to our problem. But for the purpose of illustration, an example with only two variables will suffice. I have constructed such an example by *translating* a normal error-curve with unit standard deviation to a curve in which the abscissa χ (measured from the Median) $= \xi + {\cdot}05\, \xi^3$. That it is not safe to neglect the higher powers of λ is suggested by the circumstance that if we were to do so when using the Method of Moments, we should put $\lambda = {\cdot}077$, that being the value of the constant ϵ, the first term in the expression for λ. A practitioner who should be so incautious as to apply here the method proper to our first sub-section would run the risk of tumbling into a pitfall; for instance, obtaining values for the *quaternary* coefficients which are the second differentials of T inconsistent with the condition that T should be a minimum ($-$ T a maximum).* As in a former example, let us take for the percentiles to be utilised those which have (for the generating curve) $\xi_1 = -\xi_{-1} = {\cdot}05$, $\xi_2 = -\xi_{-2} = 1$, $\xi_3 = -\xi_{-3} = 1{\cdot}5$. The corresponding values of x are given in Part I of Table IX. Figure 13 exhibits the U's and x's.

* Thus, in the case of two variables λ and α we should have

$$\frac{d^2T}{d\lambda^2}\,\frac{d^2T}{d\alpha^2} > \left(\frac{d^2T}{d\lambda d\alpha}\right)^2.$$

It may be observed that this condition is fulfilled by the values of λ and α found in our second example above, p. 742 (*a fortiori*, by our first example).

FIG. 13.



Our first step is to find a suitable initial point, two values of a and λ corresponding to the x' and y' of the lemma. As appropriate are suggested the values of those variables which would obtain if the group were normal; that is, the quasi-standard-deviation for a, and for λ zero. The reciprocal of the required value of a may be determined by the process appropriate to Sub-section I. From the *primary* data presented in Part I of Table IX we advance by a now familiar procedure to the *secondary* quantities tabulated in Part II. Whence h, the required constant, is found to be ·92 ($a = 1·087$).

TABLE IX.

Showing first steps in approximation to large constants.

PART I. PRIMARY QUANTITIES.

Abscissa or compartment.	ξ.	η.	x.	U.
± 0	± 0	....	± 0	·1915
± 1	± ·5	·352	± ·50625	·1498
± 2	± 1	·24197	± 1·05	·0919
± 3	± 1·5	·1295	± 1·66875	·0668

PART II. SECONDARY QUANTITIES.

Compartment.	P.	P″.	Q.	Q″.	PQ/U.	Q^2/U.
± 0	·176	·044	·17819	·04705	·1637	·1658
± 1	·066	·198	·0159	·207	·0334	·0384
± 2	− ·0477	− ·195	− ·03797	·2322	·0197	·0157
± 3	− ·1942	− ·4371	− ·2161	− ·4862	·6284	·6990
Sums	....	....	....	....	·8452	·9189

PART III. TERTIARY QUANTITIES.

Compartment.	A.	C.	A^2/U.
± 0	− ·012	·0419	·00074
± 1	·0038	·1753	·0001
± 2	·0128	·2504	·0018
± 3	− ·00455	− ·4671	·0003
	...	....	·0029

TABLE IX—*contd.*

PART IV. QUATERNARY QUANTITIES.

Compartment.	AC.	AC/U.
± 0	−·0005	−·0026
± 1	·0007	·0045
± 2	·0032	·035
± 3	·00213	·032
Sum	....	·07

Our next step is to vary λ from its initial value, zero. In order to determine in which direction we ought to move (which for the sake of the example we must be supposed not to know), we have to observe the sign of $\left(\frac{dT}{d\lambda}\right)$. This is one of the *quaternary* quantities specified in the first sub-section, and it is to be ascertained according to the rule there given, by summing the elements of the type A_rC_r/U_r. From an examination of (the sign of) the quantity so obtained, we learn that it is desirable to *increase* λ. Let us take a round number, small yet sensible, viz., ·02, for the increment of λ. We have now to form the expression ΣE_r^2 with h = ·92 and λ = ·02. For this purpose there may be suggested a variant of the former procedure. Substitute for the η_r of the first Sub-section H_r* $= \eta_r/(1 + 3\lambda\xi_r^2)$, where λ = ·02; for $\eta_r\xi_r$ substitute $H_r\Xi_r$, where $\Xi_r = \xi_r + \lambda\xi_r^3 (\lambda = \cdot 02)$, and call $\Delta H_r\Xi_r = \Pi_r$. Likewise let $\Delta x_r H_r = \Gamma_r$. We thus obtain a set of *modified secondary* quantities shown in Table X; which may be used in calculating the value of ΣE_r^2 for $a = 1/\cdot 92$, $\lambda = \cdot 02$. We have only to form the *tertiary* quantities A according to the familiar rules, from Π's and Γ's now, as before, from P's and Q's; and then to form the sum $\Sigma A_r^2/U_r$. Whereas the sum so formed was, prior to the variation of λ, about ·0029, it has now become ·0019.

To proceed to a smaller value of T, a higher value of −T, let us now take a step in the α (or $1/a$) direction; a step which need not be in the dark, as it were, like that of λ, but may be determined in extent as well as direction by a simple equation. We have only to operate on the modified secondary quantities as if we were seeking the value of h *de novo*. We thus obtain for h_1 the best value of h where λ = ·02, $h_1 = (\Sigma\, \Pi_r\, \Gamma_r / U_r) / (\Sigma\, \Gamma_r^2 / U_r)$, that is ·955. In fact, substituting ·95 for h and thus re-calculating the A's, I now find for $\Sigma\, E_r^2$, ·0011. The way is open for another step in a direction parallel to the axis of λ.

When we have reached values of λ and $1/a$, say λ_0 and h_0, in the near neighbourhood of the maximum, it will be possible by

* This symbol is to be understood as capital *eta*, quite distinct from the "H" of Sub-section I.

TABLE X.

Showing method of approximation to large constants (*continued*).

PART I. PRIMARY QUANTITIES.

Abscissa or compartment.	ξ.	Ξ.	η.	H.	x.	U.
± 0	0	0	....	....	0	·1915
± 1	± ·5	± ·5025	·352	·3468	± ·50625	·1498
± 2	± 1	± 1·02	·24197	·22828	± 1·05	·0919
± 3	± 1·5	± 1·56	·1295	·1141	± 1·66875	·0668

PART II. SECONDARY QUANTITIES.

Abscissa or compartment.	ΞH.	Π.	xH.	Γ.	ΠΓ/U.	Γ²/U.
± 0	*0*	·17427	*0*	·1756	·160	·161
± 1	± *·17427*	·05858	± *·17557*	·0641	·0251	·0274
± 2	± *·23285*	− ·05400	± *·23969*	− ·0493	·0292	·0266
± 3	± *·17885*	− ·17885	± *·1914*	− ·1904	·51	·543
Sums	....	....	....	....	·724	·758

substituting $\lambda_0 + \delta\lambda$ for λ, and $h_0 + \delta h$ for $1/a$, in the expression which is to be minimised, expanding in ascending powers of $\delta\lambda$ and δh and differentiating, to obtain simultaneous linear equations for $\delta\lambda$ and δh, and by repeating the operation to obtain as near an approximation as possible (as the inaccuracy of the data permits) to the true values of the coefficients. But query whether it would be worth the trouble thus to improve upon the method of the main process of approximation.

The procedure which has been exemplified for two variables may be extended to several variables, to the coefficient κ in addition to λ and α, and to other variables attributes in addition to x—the case of *surfaces* which is to be considered in the following, and concluding, section.

(*To be continued.*)