ON GENERALISED TCHEBYCHEFF THEOREMS IN THE MATHEMATICAL THEORY OF STATISTICS.

BY KARL PEARSON, F.R.S.

(1) *Single Variate.*

Let $y = \phi(x)$ be any law of frequency and let the limits of the distribution be a and b, then if N be the total frequency,

$$N = \int_a^b \phi(x)\, dx,$$

and if $\bar{x}$ be the mean value of the variate,

$$N\bar{x} = \int_a^b x\phi(x)\, dx.$$

Generally, if μ_s be the sth moment-coefficient about the mean,

$$N\mu_s = \int_a^b (x - \bar{x})^s \phi(x)\, dx.$$

Now consider
$$\mu_{2s} = \frac{1}{N}\int_a^b (x - \bar{x})^{2s} \phi(x)\, dx,$$

and let ϵ be any value of $x - \bar{x}$, then

$$\mu_{2s}/\epsilon^{2s} = \frac{1}{N}\int_a^b \frac{(x - \bar{x})^{2s}}{\epsilon^{2s}} \phi(x)\, dx.$$

Now pick out all the values for which $x - \bar{x}$ is greater than ϵ, and let us suppose $b > a$; then

$$\mu_{2s}/\epsilon^{2s} > \frac{1}{N}\int_{\epsilon+\bar{x}}^b \frac{(x - \bar{x})^{2s}}{\epsilon^{2s}} \phi(x)\, dx,$$

and therefore
$$\mu_{2s}/\epsilon^{2s} > \frac{1}{N}\int_{\epsilon+\bar{x}}^b \phi(x)\, dx,$$

since $(x - \bar{x})/\epsilon$ is always greater than unity.

But $\frac{1}{N}\int_{\epsilon+\bar{x}}^b \phi(x)\, dx$ is the chance of an individual occurring with a deviation greater than ϵ from the mean $= 1 - P$ where P is the chance of an individual occurring with a deviation less than ϵ. Hence

$$P > 1 - \frac{\mu_{2s}}{\epsilon^{2s}}.$$

Now let $\epsilon = \lambda\sigma$, where $\sigma = \sqrt{\mu_2}$ is the standard deviation of the distribution.

Thus the chance of a deviation being of less magnitude than $\lambda\sigma$ is

$$P > 1 - \frac{1}{\lambda^{2s}} \cdot \frac{\mu_{2s}}{\mu_2^s} \quad \text{.................................(i).}$$

If we put $s = 1$, the chance P of a deviation less than $\lambda\sigma$ is limited by

$$P > 1 - \frac{1}{\lambda^2} \quad \text{....................................(ii).}$$

This special case is Tchebycheff's Theorem*.

Inequality (i) gives our first generalisation for a single variate of Tchebycheff's Theorem in (ii)†. We can now compare the accuracy of (i) and (ii) by supposing them applied to a normal distribution of frequency for the cases of deviations 1, 2, 3 and 4 times the standard deviation. In this case

$$\mu_{2s} = (2s-1)(2s-3) \dots 1\mu_2^s.$$

TABLE I.

Values of Lower Limit for P given by $1 - \frac{(2s-1)(2s-3)\dots 1}{\lambda^{2s}}$.

s	$\lambda=1\cdot5$	$\lambda=2$	$\lambda=3$	$\lambda=4$
1	**·5556**	·7500	·8889	·9375
2	·4074	**·8125**	·9630	·9883
3	− ·3169	·7656	·9794	·9963
4	—	—	**·9840**	·9984
5	—	—	**·9840**	·9991
6	—	—	·9804	·9994
7	—	—	—	·99950
8	—	—	—	**·99953**
9	—	—	—	·99950
10	—	—	—	·99940
Actual value of P	·8664	·9545	·9970	·99994

Clearly the maximum for any λ will be found by making $(2s-1)/\lambda^2$ equal to unity, or if λ^2 = an odd number, $s = \frac{1}{2}(\lambda^2+1)$ and $\frac{1}{2}(\lambda^2+1) - 1$ will give equal limits. If λ^2 be an even number then $s = \frac{1}{2}\lambda^2$ will give the highest limit.

* It was first proved in the *Recueil des sciences mathématiques*, T. II, according to Liouville, but I cannot trace this reference at all. It was translated from Russian into French in Liouville's *Journal de mathématiques*, Vol. XII, pp. 177—184, Paris, 1867. The proof there given is somewhat lengthy and at first sight the result might appear more general than (ii); but this is not so. Assume $x = u + v + w + \dots$ and suppose u, v, w uncorrelated, so that $\sigma_x^2 = \sigma_u^2 + \sigma_v^2 + \sigma_w^2 + \dots$ then we have with minor differences of notation and terminology (especially the use of the words "mathematical expectation" for our moments) Tchebycheff's own phrasing of his theorem. The remark of Dr Anderson (*Biometrika*, Vol. X, p. 269) with regard to the neglect of the theory of "mathematical expectation" by the English statistical school seems based on a misunderstanding of the moment method.

† This generalised form of Tchebycheff's Theorem was given by me in a paper for the Honours degree of the University of London in Statistics, October, 1915.

(2) *Two Variates; Limit to the Frequency within an Elliptic Area round the Mean as Centre.*

Let the law of frequency be $z = \phi(x, y)$ and let the standard deviations of x and y be σ_1, σ_2, and r be the coefficient of correlation between x and y. Let us take as our ellipse,

$$\frac{1}{1-r^2}\left(\theta_{11}\frac{x^2}{\sigma_1^2} - 2\theta_{12}\frac{rxy}{\sigma_1\sigma_2} + \theta_{22}\frac{y^2}{\sigma_2^2}\right) = \chi^2;$$

x and y being measured as deviations from the mean.

Then by giving special values to θ_{11}, θ_{12}, θ_{22} and χ^2 we can get any ellipse we please. Further since the curve is to be an ellipse $r^2\theta_{12}^2 < \theta_{11}\theta_{22}$* and we shall take θ_{11} and θ_{22} always positive. Thus χ^2 and all its powers will invariably be positive. Now consider; if $N = \iint \phi(xy)\, dx dy$,

$$I_s = \frac{1}{N}\iint \phi(xy)\left\{\frac{1}{1-r^2}\left(\theta_{11}\frac{x^2}{\sigma_1^2} - 2\theta_{12}\frac{rxy}{\sigma_1\sigma_2} + \theta_{22}\frac{y^2}{\sigma_2^2}\right)\right\}^s dx dy,$$

the integration extending all over space covered by the frequency surface. Divide both sides by χ_0^{2s},

$$\frac{I_s}{\chi_0^{2s}} = \frac{1}{N}\iint \phi(xy)\left(\frac{\chi}{\chi_0}\right)^{2s} dx dy.$$

Take out all the values for which χ is greater than χ_0, then

$$\frac{I_s}{\chi_0^{2s}} > \frac{1}{N}\iint \phi(xy)\left(\frac{\chi}{\chi_0}\right)^{2s} dx dy,$$

when the integral extends over the area for which χ is $> \chi_0$. Hence

$$\frac{I_s}{\chi_0^{2s}} > \frac{1}{N}\iint \phi(xy)\, dx dy$$

> chance of an observation falling outside the ellipse χ_0.

Let P be the chance of an observation falling inside this ellipse, then we have at once

$$P > 1 - \frac{I_s}{\chi_0^{2s}} \quad \text{.................................(iii)}$$

Now we define

$$p_{ss'} = \frac{1}{N}\iint \phi(xy)(x-\bar{x})^s (y-\bar{y})^{s'} dx dy$$
$$= \frac{1}{N}\iint \phi(xy)\, x^s y^{s'} dx dy$$

in our case, as the s, s'th product moment-coefficient about the mean. And it is very convenient to write

$$q_{ss'} = p_{ss'}/(\sigma_1^s \sigma_2^{s'}) \quad \text{.................................(iv)}$$

and term $q_{ss'}$ a reduced product moment-coefficient.

* We shall generally wish to have symmetry of expression between x and y, and in this case we take $\theta_{22} = \theta_{11} = \theta$ say and write $\theta_{12} r/\theta = \rho$ and we shall have as necessary condition for the ellipse $\rho < 1$.

It is clear that by simple expansion of the trinomial expression, we can always find I_s in terms of $q_{ss'}$.

We have accordingly to study the expansion of

$$\frac{1}{(1-r^2)^s}(\theta_{11}x^2 - 2r\theta_{12}xy + \theta_{22}y^2)^s$$

$$= \frac{1}{(1-r^2)^s} \mathop{S}_{u=0}^{u=s} \mathop{S}_{m=0}^{m=s-u} \left\{(-1)^u 2^u r^u \theta_{12}{}^u \theta_{11}{}^{s-m-u} \theta_{22}{}^m \frac{s!}{(s-u-m)!\,m!\,u!} x^{2s-u-2m} y^{2m+u}\right\},$$

and if this value be substituted in the integral expression for I_s we find

$$I_s = \frac{1}{(1-r^2)^s} \mathop{S}_{u=0}^{u=s} \mathop{S}_{m=0}^{m=s-u} \left\{(-1)^u 2^u r^u \theta_{12}{}^u \theta_{11}{}^{s-m-u} \theta_{22}{}^m \frac{s!}{(s-u-m)!\,m!\,u!} q_{2s-u-2m,\,2m+u}\right\} \quad \text{......(v).}$$

The lower values can be equally readily found by the expansion of

$$\left(\theta_{11}\frac{x^2}{\sigma_1{}^2} - 2\theta_{12}r\frac{xy}{\sigma_1\sigma_2} + \theta_{22}\frac{y^2}{\sigma_2{}^2}\right)^s$$

in powers of r by aid of the binomial.

The first few cases are

$$I_1 = \frac{1}{(1-r^2)}\{\theta_{11}q_{20} - 2\theta_{12}rq_{11} + \theta_{22}q_{02}\},$$

$$I_2 = \frac{1}{(1-r^2)^2}\{\theta_{11}{}^2q_{40} + \theta_{22}{}^2q_{04} + 2\theta_{11}\theta_{22}q_{22} - 4\theta_{12}r(\theta_{11}q_{31} + \theta_{22}q_{13}) + 4\theta_{12}{}^2r^2q_{22}\},$$

$$I_3 = \frac{1}{(1-r^2)^3}\{\theta_{11}{}^3q_{60} + \theta_{22}{}^3q_{06} + 3\theta_{11}\theta_{22}(\theta_{11}q_{42} + \theta_{22}q_{24})$$
$$- 6\theta_{12}r(\theta_{11}{}^2q_{51} + \theta_{22}{}^2q_{15} + 2\theta_{11}\theta_{22}q_{33})$$
$$+ 12\theta_{12}{}^2r^2(\theta_{11}q_{42} + \theta_{22}q_{24}) - 8\theta_{11}{}^3r^3q_{33}\},$$

$$I_4 = \frac{1}{(1-r^2)^4}\{\theta_{11}{}^4q_{80} + \theta_{22}{}^4q_{08} + 4\theta_{11}\theta_{22}(\theta_{11}{}^2q_{62} + \theta_{22}{}^2q_{26}) + 6\theta_{11}{}^2\theta_{22}{}^2q_{44}$$
$$- 8\theta_{12}r(\theta_{11}{}^3q_{71} + 3\theta_{11}\theta_{22}(\theta_{11}q_{53} + \theta_{22}q_{35}) + \theta_{22}{}^3q_{17})$$
$$+ 24\theta_{12}{}^2r^2(\theta_{11}{}^2q_{62} + \theta_{22}{}^2q_{26} + 2\theta_{11}\theta_{22}q_{44})$$
$$- 32\theta_{12}{}^3r^3(\theta_{11}q_{53} + \theta_{22}q_{35}) + 16r^4\theta_{12}{}^4q_{44}\} \quad \text{......(vi).}$$

These expressions simplify for various cases, but it is clear that for the general case of unknown type of distribution we shall have to find very high product moments from the observations in order to use our generalised Tchebycheff's Theorem. Otherwise we shall have to make assumptions as to the relations between high order and low order q's.

Since generally $q_{20} = q_{02} = 1$ and $q_{11} = r$, we have

$$I_1 = \frac{1}{1-r^2}(\theta_{11} + \theta_{22} - 2\theta_{12}r^2).$$

This suggests that for all cases we are likely to get simplified results, if we take $\theta_{11} = \theta_{22} = \theta_{12} = 1$ when we find $I_1 = 2$. In other words, simplification arises if we make our ellipse that of the normal contours, although of course for the general case this will not be a contour of equal probability, although it may roughly approximate to it.

Thus we find for this case,

$$I_1 = 2,$$
$$I_2 = \frac{1}{(1-r^2)^2}\{q_{40} + q_{04} + 2q_{22} - 4r(q_{31} + q_{13}) + 4r^2 q_{22}\},$$
$$I_3 = \frac{1}{(1-r^2)^3}\{q_{60} + q_{06} + 3(q_{42} + q_{24}) - 6r(q_{51} + q_{15} + 2q_{33}) + 12r^2(q_{42} + q_{24}) - 8r^3 q_{33}\},$$
$$I_4 = \frac{1}{(1-r^2)^4}\{q_{80} + q_{08} + 4(q_{62} + q_{26}) + 6q_{44} - 8r(q_{71} + q_{17} + 3q_{53} + 3q_{35}) + 24r^2(q_{62} + q_{26} + 2q_{44}) - 32r^3(q_{53} + q_{35}) + 16r^4 q_{44}\} \quad \text{.........(vii)},$$

and the general value of I_s will be

$$I_s = \frac{1}{(1-r^2)^s} \mathop{S}_{u=0}^{u=s} \mathop{S}_{m=0}^{m=s-u} \left\{(-1)^u 2^u r^u \frac{s!}{(s-u-m)!\, m!\, u!} q_{2s-u-2m,\ 2m+u}\right\}.$$

For the case of a normal distribution the q's are all given in terms of r (*Biometrika*, Vol. XII, p. 87) and on substitution we find

$$I_1 = 2, \qquad I_2 = 8, \qquad I_3 = 48, \qquad I_4 = 384;$$

generally $I_s = 2s(2s-2)(2s-4)\ldots 2$, which can be shown directly, thus:

$$I_s = \iint \phi(x, y) \frac{1}{1-r^2}\left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)^s dxdy$$
$$= \int_0^\infty e^{-\frac{1}{2}\chi^2} \chi^{2s} \chi d\chi = 2sI_{s-1},$$

if we integrate by parts,

$$= 2s(2s-2)(2s-4)\ldots 2 \times \int_0^\infty e^{-\frac{1}{2}\chi^2} \chi d\chi$$
$$= 2s(2s-2)(2s-4)\ldots 2.$$

Accordingly our generalised Tchebycheff's limit becomes

$$P > 1 - \frac{2s(2s-2)(2s-4)\ldots 2}{\chi_0^{2s}} \quad \text{.................(viii)*},$$

and our best value of s will be determinable from $2s < \chi_0^2$, or s must be the greatest integer less than or the integer equal to $\frac{1}{2}\chi_0^2$.

Now the actual volume of the frequency surface inside the contour

$$\chi_0^2 = \frac{1}{1-r^2}\left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)$$

is known to be $1 - e^{-\frac{1}{2}\chi_0^2}$, and it is thus easy to test the present generalised Tchebycheff limit as applied to this case.

* This result is almost at once extensible to any number of variates following the normal distribution, but as the actual value of the probability is known there is no value in writing down this limiting value.

TABLE II.

Generalised Tchebycheff Limit applied to the Probability that an association of two variables lies inside a given contour χ_0^2 of a normal frequency surface.

χ_0^2	Actual Probability	Minimum value of P
4	·8647	·5000 (I_1)
5	·9179	·6800 (I_2)
6	·9502	·7778 (I_2)
7	·9698	·8600 (I_3)
8	·9817	·9062 (I_3)
9	·9889	·9415 (I_4)
10	·9933	·9616 (I_4)
12	·9975	·9846 (I_5)
14	·9991	·9939 (I_6)
16	·9997	·9976 (I_7)
18	·9999	·9991 (I_8)
20	·99995	·99964 (I_9)

Here as in the case of a single variate the generalised Tchebycheff limit is not very useful for low values of χ_0^2. But if in any particular type of observation we consider it desirable to look with suspicion on an observation which has occurred and yet the odds against which are greater than 50 to 1, the Tchebycheff limit may be of value. As illustration, suppose two variates are correlated with intensity ·7, what suspicion should we cast on an observation which gave the deviation of one variate 3·8 times its standard deviation and of the other 3·2 times? Here

$$\chi_0^2 = \frac{1}{1-r^2}\left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)$$
$$= \frac{1}{\cdot 51}\{(3\cdot 8)^2 - 1\cdot 4\,(3\cdot 8)(3\cdot 2) + (3\cdot 2)^2\}$$
$$= 15\cdot 01, \text{ or say } 15.$$

Then
$$P > 1 - \frac{2^7(7\,!)}{15^7} > \cdot 9962,$$

or the odds are greater than 250 to 1 against it. Actually the probability of the occurrence of anything as unusual as or more unusual than this is ·9994, or the actual odds 1700 to 1 about. For many purposes the odds of 250 to 1 would amply suffice to mark suspicion, although of course in the case of *normal* frequency it would be as easy or even easier to calculate the real probability as the generalised Tchebycheff limit.

The chief interest of the investigation thus far is to show that unless we use an I_s of a high order the Tchebycheff limit is unlikely to be of very much service. We can obtain it in the case of material following a normal distribution, but then we know the exact result and do not need it!

I have considered very carefully the possibilities of deducing higher q's from lower q's for non-normal systems on various hypotheses as to the nature of the regression and the scedasticity. The simplest hypothesis is to suppose linearity of regression, homoscedasticity and homocliticity of both sets of arrays.

Let $\beta_{2s-2} = S(x^{2s})/\sigma^{2s}$, and $\beta_{2s-1}/\sqrt{\beta_1} = S(x^{2s+1})/\sigma^{2s+1}$

as usual; let a single dash mark the β's for the y variate, and double dashes the β's for the y arrays of x's and triple dashes the β's for the x arrays of y's. Then if $\bar{y}_x$ be the mean of the x-array of y's,

$$\sqrt{\beta_1'} = \frac{1}{N} S(y^3)/\sigma_2^3 = \frac{1}{N} \Sigma S(\bar{y}_x + y')^3/\sigma_2^3,$$

where y' is measured from the mean of the array, S is the sum for all members of the array and Σ the sum for all arrays. Thus if $\bar{y}_x = \frac{r\sigma_2}{\sigma_1} x$ be the regression line,

$$\sqrt{\beta_1'} = \frac{1}{N}\left\{r^3 \frac{\Sigma(n_x x^3)}{\sigma_1^3} + 3r^2 \frac{\Sigma(n_x x^2)}{\sigma_1^2}\frac{S(y')}{n_x\sigma_2} + 3r\frac{\Sigma(n_x x)}{\sigma_1}\frac{S(y'^2)}{n_x\sigma_2^2} + \frac{S(y'^3)}{n_x\sigma_2^3}\Sigma(n_x)\right\},$$

since $S\frac{(y'^2)}{n_x}$, $S\frac{(y'^3)}{n_x}$, is to be the same for every array. Thus

$$\sqrt{\beta_1'} = r^3\sqrt{\beta_1} + \sqrt{\beta_1'''}(1-r^2)^{\frac{3}{2}},$$

or
$$\sqrt{\beta_1'''} = \frac{\sqrt{\beta_1'} - r^3\sqrt{\beta_1}}{(1-r^2)^{\frac{3}{2}}} \quad \text{.............................(ix).}$$

Similarly
$$\sqrt{\beta_1''} = \frac{\sqrt{\beta_1} - r^3\sqrt{\beta_1'}}{(1-r^2)^{\frac{3}{2}}} \quad \text{.............................(x).}$$

Thus it is impossible in homoclitic systems for the skewness of the arrays to be equal to the skewness of the marginal totals if there be correlation*.

Again we have

$$\beta_2' = \frac{1}{N} S(y^4)/\sigma_2^4 = \frac{1}{N} S\left(\frac{r\sigma_2}{\sigma_1}x + y'\right)^4 \Big/ \sigma_2^4$$

$$= r^4\beta_2 + 6r^2(1-r^2) + \beta_2'''(1-r^2)^2,$$

or
$$\beta_2''' = \frac{\beta_2' - r^4\beta_2 - 6r^2(1-r^2)}{(1-r^2)^2},$$

or, again,
$$\beta_2''' - 3 = \frac{\beta_2' - 3 - r^4(\beta_2 - 3)}{(1-r^2)^2} \quad \text{..........................(xi),}$$

and similarly,
$$\beta_2'' - 3 = \frac{\beta_2 - 3 - r^4(\beta_2' - 3)}{(1-r^2)^2} \quad \text{.......................(xii).}$$

* We note that if the marginal totals be both without skewness, all the arrays will also be symmetrical. Equations (xi) and (xii) show us that if the marginal totals be mesokurtic the arrays will also be mesokurtic.

Now consider q_{22},

$$q_{22} = \frac{1}{N} S(x^2y^2)/(\sigma_1^2\sigma_2^2) = \frac{1}{N}\Sigma Sx^2\left(\frac{r\sigma_2}{\sigma_1}x + y'\right)^2 \Big/ \sigma_1^2\sigma_2^2$$

$$\left.\begin{aligned} &= \beta_2 r^2 + (1 - r^2) \\ &= \beta_2' r^2 + (1 - r^2) \end{aligned}\right\} \quad \text{.........(xiii)}$$

by symmetry. Hence it follows that in linear homoscedastic systems $\beta_2 = \beta_2'$, and accordingly

$$\beta_2'' - 3 = \beta_2''' - 3 = (\beta_2 - 3)\frac{1 + r^2}{1 - r^2} \quad \text{.........(xiv).}$$

This is of interest as indicating that in linear homoscedastic systems the one with mesokurtic margins is the only one in which the kurtosis of the arrays can be the same as that of the margins.

Again $$q_{33} = \frac{1}{N} S(x^3y^3)/(\sigma_1^3\sigma_2^3) = \frac{1}{N}\Sigma Sx^3\left(r\frac{\sigma_2}{\sigma_1}x + y'\right)^3 \Big/ \sigma_1^3\sigma_2^3$$

$$= r^3\beta_4 + \frac{3r}{N}\frac{\Sigma(n_x x^4)}{\sigma_1^4} S\left(\frac{y'^2}{n_x\sigma_2^2}\right) + \frac{1}{N}\Sigma\left(\frac{n_x x^3}{\sigma_1^3}\right) S\left(\frac{y'^3}{n_x\sigma_2^3}\right)$$

$$= r^3\beta_4 + 3r(1 - r^2)\beta_2 + \sqrt{\beta_1}\sqrt{\beta_1'} - r^3\beta_1 \quad \text{.........(xv)}$$

$$= r^3\beta_4' + 3r(1 - r^2)\beta_2' + \sqrt{\beta_1'}\sqrt{\beta_1} - r^3\beta_1' \quad \text{.........(xvi)}$$

by symmetry.

It follows from (xv) and (xvi) that it is needful for

$$\beta_4 - \beta_1 = \beta_4' - \beta_1' \quad \text{.........(xvii).}$$

Finally we have

$$q_{44} = \frac{1}{N} S(x^4y^4)/(\sigma_1^4\sigma_2^4)$$

$$= \frac{1}{N}\Sigma Sx^4\left(r\frac{\sigma_2}{\sigma_1}x + y'\right)^4 \Big/ \sigma_1^4\sigma_2^4$$

$$= r^4\beta_6 + 6r^2(1 - r^2)\beta_4 - 4r^2\beta_3 + 4\beta_3\sqrt{\beta_1'/\beta_1} + \beta_2\beta_2'''(1 - r^2)^2,$$

or

$$q_{44} = r^4\beta_6 + 6r^2(1 - r^2)\beta_4 - 4r^2\beta_3 + 4\beta_3\sqrt{\beta_1'/\beta_1} - r^4\beta_2^2 - 6r^2(1 - r^2)\beta_2 + \beta_2\beta_2'$$

$$= r^4\beta_6' + 6r^2(1 - r^2)\beta_4' - 4r^2\beta_3' + 4\beta_3'\sqrt{\beta_1/\beta_1'} - r^4\beta_2'^2 - 6r^2(1 - r^2)\beta_2' + \beta_2'\beta_2 \quad \text{......(xviii),}$$

which again involves the complicated β-relation

$$r^4(\beta_6 - \beta_6') + 6r^2(1 - r^2)(\beta_4 - \beta_4') - 4r^2(\beta_3 - \beta_3') + 4(\beta_3\beta_1' - \beta_3'\beta_1)\sqrt{\beta_1\beta_1'} = 0 \quad \text{......(xix).}$$

It is difficult to see how the form of variation of one character can be related by the correlation between that and another character to the form of variation of the second character as (xix) would indicate. If it were we should get into great difficulties in dealing with similar conditions to (xix) for a large number of characters with different correlations. If as it appears to me (xix) would need to be satisfied independently of r, then we must have

$$\left.\begin{aligned} \beta_6 - 6\beta_4 &= \beta_6' - 6\beta_4' \\ 3\beta_4 - 2\beta_3 &= 3\beta_4' - 2\beta_3' \\ \beta_3/\beta_1 &= \beta_3'/\beta_1' \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(xx)}.$$

The second of (xx) by aid of (xvii) leads us to

$$3\beta_1\left(1 - \frac{2}{3}\frac{\beta_3}{\beta_1}\right) = 3\beta_1'\left(1 - \frac{2}{3}\frac{\beta_3'}{\beta_1'}\right),$$

whence $\beta_1 = \beta_1'$, and as $\beta_2 = \beta_2'$ it follows that $\beta_3 = \beta_3'$, $\beta_4 = \beta_4'$, $\beta_6 = \beta_6'$, that is to say the total frequencies of the two correlated characters must possess variation practically of the same type.

Now I find this is very far from being the case in distributions which differ widely from the normal correlation surface. Thus it follows that the hypothesis of homoscedasticity, linear regression and homocliticity fails for such cases. I therefore modified the linear regression and adopted skew regression, homoscedasticity and homocliticity. I again got relations between the β's, but of a much higher degree of complexity. These were tested by Mr A. W. Young and myself on the skew correlation surfaces of barometric data, but were found to fail. Direct investigation afterwards showed me that while the regression differed to some extent from linearity, it was the homoscedasticity which was in the first place the erroneous assumption. The arrays were very far from having the same standard deviations.

Until therefore some theoretical advance is made in the investigation of skew regression surfaces, especially for those which have linear or nearly linear regression combined with heteroscedasticity, it is unlikely that we shall have any adequate method of determining high product moment-coefficients from low ones. We are accordingly thrown back on direct determination of the high product moment-coefficients, if we wish to determine a Tchebycheff limit. The work of determining I_4 would involve a whole round of 8th order moment-coefficients and product moment-coefficients. It would then give us a limit of the order ·95 for ·99. Lower order I's would hardly give values of much importance, and it may be questioned whether a rough limit of the kind required could not be better obtained by inserting the desired contour on a "scatter diagram" and simply counting the dots which fall outside it, or indeed by taking the best fitting normal surface to the actual distribution. The reader may question whether something better could not be achieved for skew correlation Tchebycheff limits by some contour other than the ellipse. This would undoubtedly be the case, if we knew the forms of the skew-correlation contours, for then we should undoubtedly choose this equi-probable locus for our boundary. But as we have only a knowledge of these empirically—experience shows them to be frequently pear or lemniscate loop shaped—we get little help for our present problem.

One other aspect of the matter may be briefly considered. We may find a limit to the probability that an event or individual will lie within a circle of radius R

round the origin. This corresponds to Schols' Problem*. It may be useful to have a Tchebycheff limit for this case, although we have yet to meet the particular instance in practical statistics where it would be of marked advantage†.

We can best investigate this problem *de novo*.

Let $$I_s = \iint (x^2 + y^2)^s \phi(xy)\, dx\, dy.$$

Then if R be any radius round the origin,

$$I_s/R^{2s} = \iint \left(\frac{x^2 + y^2}{R^2}\right)^s \phi(xy)\, dx\, dy,$$

the integral being taken to include the whole volume of the probability surface $z = \phi(x, y)$. Now pick out those elements of the integral for which $x^2 + y^2$ is $> R^2$, then

$$I_s/R^{2s} > \iint \left(\frac{x^2 + y^2}{R^2}\right)^s \phi(xy)\, dx\, dy,$$

where the integration extends over the above-mentioned elements only, and is therefore

$$> \iint \phi(xy)\, dx\, dy,$$

but this integral is $1 - P$, where P is the probability that the individual falls within the distance R of the origin. Thus the Tchebycheff limit is given by

$$P > 1 - \frac{I_s}{R^{2s}} \text{ ....}$$

Now clearly we have

$$\begin{aligned} I_s &= \iint (x^2 + y^2)^s \phi(xy)\, dx\, dy \\ &= p_{2s,3} + s p_{2s-2,2} + \frac{s(s-1)}{1 \cdot 2} p_{2s-4,4} + \dots \\ &= \sigma_1^{2s} q_{2s,0} + s \sigma_1^{2s-2} \sigma_2^2 q_{2s-2,2} + \frac{s(s-1)}{1 \cdot 2} \sigma_1^{2s-4} \sigma_2^4 q_{2s-4,4} + \dots. \end{aligned}$$

Now write $R = \lambda \sqrt{\sigma_1^2 + \sigma_2^2}$, and further take $\tan \theta = \sigma_2/\sigma_1$. Then

$$\frac{I_s}{R^{2s}} = \frac{1}{\lambda^{2s}} \left\{ \cos^{2s}\theta\, q_{2s,0} + s \cos^{2s-2}\theta \sin^2\theta\, q_{2s-2,2} + \frac{s(s-1)}{1 \cdot 2} \cos^{2s-4}\theta \sin^4\theta\, q_{2s-4,4} \right.$$
$$\left. + \frac{s(s-1)(s-2)}{1 \cdot 2 \cdot 3} \cos^{2s-6}\theta \sin^6\theta\, q_{2s-6,6} + \dots \right\} \text{ ......}$$

For the particular case in which $s = 1$,

$$\frac{I_1}{R^2} = \frac{1}{\lambda^2} (\cos^2\theta + \sin^2\theta) = \frac{1}{\lambda^2} \text{ ....}$$

For $s = 2$, $$\frac{I_2}{R^4} = \frac{1}{\lambda^4} (\cos^4\theta\, \beta_2 + 2 \cos^2\theta \sin^2\theta\, q_{22} + \sin^4\theta\, \beta_2').$$

* Over de Theorie der Fouten in Ruimte en in het platte Vlak, *Verhandlingen der K. Akademie van Wetenschapen*, Deel XV, pp. 1—68, Amsterdam, 1875. Translated into French in the *Annales de l'École polytechnique de Delft*, Tome II, pp. 123—178. Leide, 1886.

† It is conceivable that the solutions given might be serviceable in the case of testing machine guns against a target.

Now a good approximation to q_{22} by (xiii) must be $\frac{1}{2}(\beta_2 + \beta_2')\, r^2 + (1 - r^2)$; hence substituting

$$\frac{I_2}{R^4} = \frac{1}{\lambda^4}\{(\beta_2 - 3)(\cos^4\theta + r^2\sin^2\theta) + (\beta_2' - 3)(\sin^4\theta + r^2\cos^2\theta) + 3 - 4(1 - r^2)\cos^2\theta\sin^2\theta\} \ldots\ldots.$$

For the special case of normal distribution, if we write $\kappa^2 = 4(1 - r^2)\cos^2\theta\sin^2\theta$,

$$\frac{I_2}{R^4} = \frac{1}{\lambda^4}(3 - \kappa^2) \ldots\ldots.$$

Again

$$\frac{I_3}{R^6} = \frac{1}{\lambda^6}\{\cos^6\theta\beta_6 + 3\cos^2\theta\sin^2\theta(\cos^2\theta q_{42} + \sin^2\theta q_{24}) + \sin^6\theta\beta_6'\} \ldots\ldots$$

and for a normal distribution,

$$\frac{I_3}{R^6} = \frac{1}{\lambda^6}(15 - 9\kappa^2) \ldots\ldots.$$

Further general cases can be at once written down, but it will suffice to give here the leading values of I_s for a normal distribution:

$$\frac{I_1}{R^2} = \frac{1}{\lambda^2}, \quad \frac{I_2}{R^4} = \frac{1}{\lambda^4}(3 - \kappa^2), \quad \frac{I_3}{R^6} = \frac{1}{\lambda^6}(15 - 9\kappa^2),$$

$$\frac{I_4}{R^8} = \frac{1}{\lambda^8}(105 - 90\kappa^2 + 9\kappa^4), \quad \frac{I_5}{R^{10}} = \frac{1}{\lambda^{10}}(945 - 1050\kappa^2 + 225\kappa^4),$$

$$\frac{I_6}{R^{12}} = \frac{1}{\lambda^{12}}(10395 - 14175\kappa^2 + 4725\kappa^4 - 225\kappa^6),$$

$$\frac{I_7}{R^{14}} = \frac{1}{\lambda^{14}}(135{,}135 - 218{,}295\kappa^2 + 99{,}225\kappa^4 - 11{,}025\kappa^6),$$

$$\frac{I_8}{R^{16}} = \frac{1}{\lambda^{16}}(2{,}027{,}025 - 3{,}783{,}780\kappa^2 + 2{,}182{,}950\kappa^4 - 396{,}900\kappa^6 + 11{,}025\kappa^8) \ldots\ldots.$$

The following table gives the maximum Tchebycheff limit for the probability of an individual falling within the circle $\lambda\sqrt{\sigma_1^2 + \sigma_2^2}$ for various values of

$$\kappa^2 = 4(1 - r^2)\,\sigma_1^2\sigma_2^2/(\sigma_1^2 + \sigma_2^2)^2.$$

(I_s) denotes the particular I from which the maximum limit is found. $(I_8\,?)$ denotes that the corresponding numerical value is a Tchebycheff limit found from I_8, but it is not known whether I_9 would not give a higher value, I_9 not having been tabled. The second part of the table provides the values of I_s from which the first part has been computed. They may be useful in the determination of the Tchebycheff limits for other values of λ.

I. *Generalised Tchebycheff Limit for Schols' Problem with a Normal Distribution.*

Radius of circle $= \lambda \sqrt{\sigma_1^2 + \sigma_2^2}$, $\kappa^2 = 4(1-r^2)\sigma_1^2\sigma_2^2/(\sigma_1^2+\sigma_2^2)^2$.

Values of κ^2. $\lambda=$	1	1·25	1·5	2·0	2·5	3	3·5	4·0
0·0	0 (I_1)	·36 (I_1)	·5556 (I_1)	·8125 (I_2)	·9386 (I_3)	·98400 ($I_4=I_5$)	·996924 (I_6)	·999528 (I_8)
0·1	0 (I_1)	·36 (I_1)	·5556 (I_1)	·81875 (I_2)	·9422 (I_3)	·98574 (I_5)	·997329 (I_6)	·999611 (I_8)
0·2	0 (I_1)	·36 (I_1)	·5556 (I_1)	·8275 (I_2)	·9459 (I_3)	·98740 (I_5)	·997707 (I_6)	·999685 (I_8?)
0·3	0 (I_1)	·36 (I_1)	·5556 (I_1)	·83125 (I_2)	·9496 (I_3)	·98899 (I_5)	·998109 (I_7)	·999749 (I_8?)
0·4	0 (I_1)	·36 (I_1)	·5556 (I_1)	·8375 (I_2)	·9538 (I_4)	·99099 (I_5)	·998478 (I_7)	·999805 (I_8?)
0·5	0 (I_1)	·36 (I_1)	·5556 (I_1)	·84375 (I_2)	·9592 (I_4)	·99193 (I_5)	·998806 (I_7)	·999853 (I_8?)
0·6	0 (I_1)	·36 (I_1)	·5556 (I_1)	·8500 ($I_2=I_3$)	·9645 (I_4)	·99333 (I_6)	·999096 (I_8)	·999893 (I_8?)
0·7	0 (I_1)	·36 (I_1)	·5556 (I_1)	·8641 (I_3)	·9696 (I_4)	·99490 (I_6)	·999380 (I_8)	·999927 (I_8?)
0·8	0 (I_1)	·36 (I_1)	·5654 (I_2)	·8781 (I_3)	·9746 (I_4)	·99631 (I_6)	·999609 (I_8)	·999954 (I_8?)
0·9	0 (I_1)	·36 (I_1)	·5852 (I_2)	·8922 (I_3)	·9809 (I_5)	·99770 (I_7)	·999788 (I_8)	·999975 (I_8?)
1·0	0 (I_1)	·36 (I_1)	·6049 (I_2)	·90625 ($I_3=I_4$)	·9879 (I_6)	·99895 (I_7)	·999920 (I_8?)	·999991 (I_8?)

II. *Values of the functions I_s forming the denominator of the Tchebycheff Limit to the probability that an Individual will fall for the case of Normal Bi-variate Frequency within a given circle of radius $\lambda\sqrt{\sigma_1^2+\sigma_2^2}$.*

κ^2	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8
0·0	1	3·0	15·0	105·00	945·00	10,395·000	135,135·000	2,027,025·0000
0·1	1	2·9	14·1	96·09	842·25	9,024·525	114,286·725	1,670,080·7025
0·2	1	2·8	13·2	87·36	744·00	7,747·200	95,356·800	1,354,429·4400
0·3	1	2·7	12·3	78·81	650·25	6,561·675	78,279·075	1,077,729·5025
0·4	1	2·6	11·4	70·44	561·00	5,466·600	62,987·400	837,665·6400
0·5	1	2·5	10·5	62·25	476·25	4,460·625	49,415·625	631,949·0625
0·6	1	2·4	9·6	54·24	396·00	3,542·400	37,497·600	458,317·4400
0·7	1	2·3	8·7	46·41	320·25	2,710·575	27,167·175	314,534·9025
0·8	1	2·2	7·8	38·76	249·00	1,963·800	18,358·200	198,392·0400
0·9	1	2·1	6·9	31·29	182·25	1,300·725	11,004·525	107,705·9025
1·0	1	2·0	6·0	24·00	120·00	740·000	5,040·000	40,320·0000

The reader may be curious to know whether the Tchebycheff limit gives a better result for Schols' circles than for the elliptic contours. The actual probability of an individual falling within the circle of radius $\lambda\sqrt{\sigma_1^2+\sigma_2^2}$ is given by

$$P = 1 - \frac{\kappa}{\pi}\int_0^\pi \frac{e^{-\frac{\lambda^2}{\kappa^2}(1-\kappa'\cos\theta)}}{1-\kappa'\cos\theta}\,d\theta,$$

where $\kappa' = \sqrt{1-\kappa^2}$ and $\kappa^2 = 4(1-r^2)\sigma_1^2\sigma_2^2/(\sigma_1^2+\sigma_2^2)^2$ as before.

I have not succeeded in finding any rapidly converging expansion for this expression*, and have been reduced to evaluating its argument and using a quadrature formula. Thus for $\lambda = 2$, $\kappa^2 = \cdot 4$, I find

$$P = \cdot 963{,}3694.$$

* Unfortunately Schols has not tabled P, but only gives the values of λ for ten values of κ', which occur when $P=1/2$, i.e. radial values for generalised "probable errors."

The process is not as long as it might seem. Indeed if we only need four decimal places, it is quite adequate to integrate only through the first quadrant, the second contributes nothing of importance. The value given by the last Tchebycheff limit is

$$P > \cdot 8375.$$

This is of the same order of divergence as we found for the elliptic contour, i.e. for $\chi_0^2 = 7$, we had $P = \cdot 9698$, with a Tchebycheff limit $P > \cdot 8600$. Thus the measure of approach does not seem very close in this case until we reach higher values of λ.

On the whole we must express disappointment at the results of the Tchebycheff process. We had found Tchebycheff's own limit based only on the second moment of small practical value, although it is to be found occupying a prominent position in many continental works on probability. By extending it to higher moments and product-moments we have reached results which are great improvements on the original Tchebycheff limit, but the method still lacks the degree of approximation (except for probabilities over ·99, say) which would make the result of real value in practical statistics. It is, however, conceivable that some more ingenious application of Tchebycheff's idea may lead to a limit more close to the actual value of the probability.