# REMARKS ON REGRESSION

*By*

S. D. Wicksell

1. In a paper published twelve years ago[1] I derived a set of formulae for bivariate regression which were found to give good results on unimodal materials of a fairly general nature and which, in the case of moderately skew distributions, were reduced to very simple and easily applicable forms. Two years later I extended the theory also to the case of multiple correlations of similar types[2]. These formulae were deduced on the assumption that the correlation surface could be expressed by a so-called series of type A[3], i. e. that the deviations from the best fitting normal surface could be expressed as a series, developed according to the derivatives of different orders of the Bravais function, expressing that normal surface.

When, after the lapse of so many years, I find that this theory has not received the attention which it seems to me it merits in view of the very simple, and on a fairly large class of curved regressions readily applicable results, I attribute this in part at least to the apparent (not actual) speciality of the assumptions made with regard to the mathematical expression for the correlation surface, and in part also to the rather repellent show of mathematics involved in the deductions. In the hope to give the theory a better chance of coming to the attention of statisticians, I propose here to deduce some of my main results in an entirely different way, bringing the theory back on more simple principles. I believe that by this method of deduction it will be more easy for the reader to see exactly where assumptions come in, and also the nature of the restrictions caused by these assumptions.

2. Let $x$ and $y$ be a pair of correlated variates, our material

1. The correlation function of Type A, and the regression of its characteristics. Kungl. Svenska Vetenskapsakademiens Handlinger Bd. 58, Nr. 3. 1917. Also "Meddelanden fran Lunds Astronomiska Observatorium" Ser. II, Nr. 17.
2. Multiple correlation and non-linear regression. Arkiv. for Matematik, Fysik och Astronomi. Bd. 14 Nr. 10, 1919. Also "Meddelanden fran Lunds Astronomiska Observatorium." Ser. I, Nr. 91.
3. Charlier. Contributions to the mathematical theory of statistics. 6. The correlation function of type A. Arkiv for Matematik, Fysik och Astronomi. Bd. 9, Nr. 26, 1914. Also "Meddelanden fran Lunds Astronomiska Observatorium" Ser. I, Nr. 58.

consisting of $N$ such pairs. Computing the means and central moments, we have

$$M_x = \frac{1}{N}\sum x \; ; \quad M_y = \frac{1}{N}\sum y \; ; \quad \mu_{ij} = \frac{1}{N}\sum (x - M_x)^i (y - M_y)^j$$

The standard deviations of $x$ and $y$ and the coefficient of correlation are then defined by

$$\sigma_x = \sqrt{\mu_{20}} \; ; \qquad \sigma_y = \sqrt{\mu_{02}} \qquad\qquad r = \frac{\mu_{11}}{\sigma_x \sigma_y}$$

Following Yule[1] and Pearson[2] we now treat the problem of regression as a simple problem of graduation, defining the regression of $y$ on $x$ as a parabola of a given degree, which, with $x$ as argument, is fitted to the $y$'s by the method of least squares. The regression may then be written in the form

$$y_x - M_y = a_0 + a_1 (x - M_x) + a_2 (x - M_x)^2 + \cdots + a_p (x - M_x)^p,$$

and the least squares normal equations for determining the parameters $a_0, a_1, a_2, \cdots a_p$ assume the form (Pearson Op. Cit. p. 25).

$$(1) \quad \begin{cases} 0 = a_0 + a_2 \mu_{20} + a_3 \mu_{30} + \cdots + a_p \mu_{p,0} \\ \mu_{11} = a_1 \mu_{20} + a_2 \mu_{30} + a_3 \mu_{40} + \cdots + a_p \mu_{p+1,0} \\ \mu_{21} = a_0 \mu_{20} + a_1 \mu_{30} + a_2 \mu_{40} + a_3 \mu_{50} + \cdots + a_p \mu_{p+2,0} \\ \mu_{31} = a_0 \mu_{30} + a_1 \mu_{40} + a_2 \mu_{50} + a_3 \mu_{60} + \cdots + a_p \mu_{p+3,0} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mu_{p,1} = a_0 \mu_{p,0} + a_1 \mu_{p+1,0} + a_2 \mu_{p+2,0} + a_3 \mu_{p+3,0} + \cdots + a_p \mu_{2p,0} \end{cases}$$

1. On the Theory of Correlation. Jour. Roy. Stat. Soc., Vol. 60, 1897, and On the Theory of Correlation for any number of Variables treated by a new System of Notation. Proc. Roy. Soc., Ser. A, Vol. 79, 1907.
2. Mathematical Contributions to the Theory of Evolution XIV. On the General Theory of Skew Correlation and non-Linear Regression. Drapers Co. Research Memoirs Biometric Series II. Cambridge Univ. Press, 1905.

Writing the solution in the form of determinants, we have

$$a_{i-1} = \frac{1}{\Delta} \cdot \Delta_i ,$$

where

$$(3) \qquad \Delta = \begin{vmatrix} 1, & 0, & \mu_{20}, & \mu_{30}, & \cdots & \mu_{p,0} \\ 0, & \mu_{20}, & \mu_{30}, & \mu_{40}, & \cdots & \mu_{p+1,0} \\ \mu_{20}, & \mu_{30}, & \mu_{40}, & \mu_{50}, & \cdots & \mu_{p+2,0} \\ \mu_{30}, & \mu_{40}, & \mu_{50}, & \mu_{60}, & \cdots & \mu_{p+3,0} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \mu_{p,0}, & \mu_{p+1,0}, & \mu_{p+2,0}, & \mu_{p+3,0}, & \cdots & \mu_{2p,0} \end{vmatrix}$$

and $\Delta_i$ is obtained when the $i$'th row in $\Delta$ is exchanged for the left membra of equations (1), i. e. for the series of elements:

$$0, \ \mu_{11}, \ \mu_{21}, \ \mu_{31}, \ \cdots \ \mu_{p,1} .$$

3. Some important general conclusions may at once be derived from this system. Defining as *non-regression of the $p$'th order* the case that all the coefficients $a_1, a_2, a_3, \cdots a_p$ turn out to be practically equal to zero, i. e. that a horizontal straight line is the best parabola of the $p$'th degree that can be fitted to the series of $y$'s, it is first seen, from the first of equations (1), that then also $a_0 = 0$ . Secondly we can draw the conclusion that this can take place only if all the elements $\mu_{11}, \ \mu_{21}, \ \mu_{31}, \ \cdots \ \mu_{p,1}$, are equal to zero. Hence the condition for non-regression of the $p$'th order of $y$ on $x$ is that we have

$$(4) \qquad \mu_{i,0} = 0 \quad for \ i = 1, 2, 3, \cdots p$$

This clearly involves also that the coefficient of correlation, r equals zero.

Defining further as *linear regression of the $p$'th order* the case that the coefficients $a_2, a_3, \cdots a_p$ are equal to zero, i. e. that a non-horizontal straight line is the best parabola of the $p$'th degree that can be fitted to the series of $y$'s, we immediately see, from the two first of equations (1), that then we must have

$$(5) \qquad a_0 = 0 \ ; \qquad a_1 = \frac{\mu_{11}}{\mu_{20}}$$

Referring here to the well-known theorem that any determinant will disappear, when the elements of two rows are proportional (the elements of any one row being obtained by multiplying the corresponding elements of another row by a constant factor) it is easily seen that all the determinants $\Delta_i$ except $\Delta_2$, and hence by (2) all the coefficients $a_0, \cdots a_p$, except $a_1$, will disappear if the quantities $o$, $\mu_{11}$, $\mu_{21}, \mu_{31}, \cdots \mu_{p,1}$, in the left membra of (1) are proportional to the elements $o$, $\mu_{20}, \mu_{30}, \cdots \mu_{p+1,0}$ in the second row of the determinant $\Delta$. Hence the condition for linear regression of the $p$'th order of $y$ on $x$ is that we have

$$(6) \qquad \mu_{11} : \mu_{i+1,0} = \mu_{20} : \mu_{i,1} \qquad \text{for } i = 1, 2, 3, \cdots p.$$

A few considerations will show that this condition is not only sufficient but also necessary. For $p = 3$ these criteria were demonstrated by Pearson.

4. Thus far there are no other assumptions involved than the principle of least squares, and that the regression of $y$ on $x$ may be described by a whole rational function. The chief difficulty in the application of this theory of regression is that, as seen from equation (1), in order to determine a regression of the $p$'th degree we must compute and use moments (of the series of $x$'s up to the order $2p$). Now, as justly remarked by Pearson, moments of high orders are, on account of their large standard errors, very little to be relied upon, at least in the case of ordinary materials ($N$ not very large). Besides this, the numerical labor involved in computing higher moments is comparatively very great. Hence, Pearson's theory of regression will be practically applicable only in cases when the regression is at the most parabolic of the second degree. Indeed, this is a very serious restriction, because curved regressions often have at least one inflection. Thus in order to meet fairly frequent cases of regression we must needs have recourse at least to cubic parabolas. But this should require the computation of all the moments of $x$ up to the sixth order.

In order to remove, as far as possible, this difficulty, I take refuge in a golden rule expressed by Thiele[1]. Thiele introduces, instead of the moments, a system of coefficients called the semi-invariants. These semi-invariants (here denoted by $\lambda_{i,o}$) are defined in terms of the moments by the identity:

1. Theory of Observations. London 1903, p. 49.

$$\lambda_{20}\frac{x^2}{2!}+\lambda_{30}\frac{x^3}{3!}+\lambda_{40}\frac{x^4}{4!}+\cdots$$
$$=\log_e\left(1+\mu_{20}\frac{x^2}{2!}+\mu_{30}\frac{x^3}{3!}+\mu_{40}\frac{x^4}{4!}+\cdots\right.$$

Developing, we find

$$(7)\quad \begin{aligned}&\lambda_{20}=\mu_{20};\ \lambda_{30}=\mu_{30};\ \lambda_{40}=\mu_{40}-3\mu_{20}^2;\\ &\lambda_{50}=\mu_{50}-10\mu_{30}\mu_{20};\ \lambda_{60}=\mu_{60}-15\mu_{40}\mu_{20}+30\mu_{20}^3-10\mu_{30}^2\end{aligned}$$

Now, the rule indicated by Thiele is the following:

To obtain the first semi-invariants rely entirely on computations. To obtain the intermediate semi-invariants rely partly on computations, partly on theoretical considerations. But to obtain the higher semi-invariants rely entirely on theoretical considerations.

Of course, this rule is just as well applicable to the determination of moments, as any moment may be expressed in terms of the semi-invariants of the same and lower order. In particular we have

$$(8)\quad \begin{aligned}&\mu_{20}=\lambda_{20};\ \mu_{30}=\lambda_{30};\ \mu_{40}=\lambda_{40}+3\lambda_{20}^2;\\ &\mu_{50}=\lambda_{50}-10\lambda_{30}\lambda_{20};\ \mu_{60}=\lambda_{60}+15\lambda_{40}\lambda_{20}+15\lambda_{20}^3+10\lambda_{30}^2\end{aligned}$$

5. A most natural way of applying the rule is afforded by Pearson's celebrated theory of frequency-functions. The moments $\mu_{i,0}$ are the moments of one of the marginal distributions (here the distribution of the $x$'s). Computing $\mu_{20}$, $\mu_{30}$ and $\mu_{40}$ in the ordinary way from the observations, criteria can be formed[1] showing to which of the Pearson Types the frequency curve of $x$ belongs. This being decided, the parameters of the curve may be determined by the aid of the same moments. As the moments of higher order are easily expressed in terms of the parameters we get, in this way, $\mu_{50}$ and $\mu_{60}$ expressed in terms of $\mu_{20}$, $\mu_{30}$ and $\mu_{40}$.

To state the matter in a more general way, we may use the formulae given by Pearson in his memoir on regression, loc. cit. pp. 5 and 6.

1. See W. Palin Elderton: Frequency Curves and Correlation, London 1927, Table VI.

Pearson starts from a differential equation of the form

$$(9)\quad f'(x)(b_0 + b_1 x + b_2 x^2 + b_3' x^3 + \cdots) = (x+a) f(x)$$

where $f(x)$ is the frequency function of $x$.

Multiplying on both sides by $x$ and integrating by parts, he finds the following formulae[1] (placing the origin in the mean)

$$(10)\quad n b_0 \mu_{n-1,0} + (n+1) b_1 \mu_{n,0} + (n+2) b_2 \mu_{n+1,0} + \cdots = -\mu_{n+1,0} - a \mu_{n,0}$$

Now, Pearson remarks that experience shows that for the great bulk of frequency distributions the higher terms, multiplied by $b_3$, $b_4$, etc., may be neglected. In fact, Pearson's system of frequency curves is obtained as a result of putting $b_i = 0$ for $i \geq 3$.

Following Pearson's example, we get the recursion formula,

$$(11)\quad n b_0 \mu_{n-1,0} + \left[(n+1) b_1 + a\right] \mu_{n,0} = -\left[(n-2) b_2 - 1\right] \mu_{n+1,0}$$

Putting here $n = 0, 1, 2, 3$, we get four equations to determine $a$, $b_0$, $b_1$, and $b_2$ in terms of the moments $\mu_{20}$, $\mu_{30}$, and $\mu_{40}$. This being done, we get $\mu_{50}$ and $\mu_{60}$ on putting $n = 4$ and 5.

The procedure indicated above leads, in fact, to the theory of skew regression which is the natural consequence of Pearson's theory of skew frequency curves.

6. As the theory just indicated above is at present at my request being worked out in detail by one of my pupils, Mr. Walter Anderson, I refrain from proceeding further into the matter.

It remains, however, to show how the special formulae for cubic regression, given by me twelve years ago, arise out of a somewhat similar procedure.

Instead of starting from Pearson's theory of frequency functions, I now start from Thiele's theory of frequency functions. Just as in the preceding section the coefficients $b_3$, $b_4$ etc. were neglected in the equation (10), given by Pearson, I now neglect the semi-invariants $\lambda_{50}$ and $\lambda_{60}$ in the equations (8), given by Thiele. There is no doubt that the former approximation is of

1. See also Palin Elderton, Op. cit. p. 39.

far more general validity than the latter; still the latter may be justified by the following considerations.

Assuming the variate $x$ to be generated as the sum of a large number of independent, elementary increments, each of which has its own frequency distribution and its own set of semi- invariants, it follows from the theory of Thiele that any semi-invariant $\lambda_{r,0}$ of $x$ is the sum of the elementary semi-invariants of the same order. Supposing the elementary increments to be $s$ in number and denoting by $\lambda'_r$ the mean value of the $r$ elementary semi-invariants of order $r$ we consequently have

$$\lambda_{r,0} = s\lambda'_r$$

Hence we get

$$\gamma_{r,0} = \frac{\lambda_{r,0}}{\lambda_{20}^{r/2}} = \frac{\lambda'_r}{\lambda_2^{r/2}} \frac{1}{s^{\frac{r-2}{2}}}$$

Except under rather special conditions, which it is not necessary to dwell on here, the ratios $\lambda'_r / \lambda'^{r/2}_2$ are not extensively great. Thus if $s$ is a large number we see that the "standardized" semi-invariants $\gamma_{r,0}$ of $x$ are small of the order of magnitude of $\left(\frac{1}{\sqrt{s}}\right)^{r-2}$ In particular we have.

| | | | | |
|---|---|---|---|---|
| $\gamma_{30}$ | of | the | order | $\frac{1}{\sqrt{s}}$ |
| $\gamma_{40}$ | " | " | " | $\frac{1}{s}$ |
| $\gamma_{50}$ | " | " | " | $\frac{1}{s\sqrt{s}}$ |
| $\gamma_{60}$ | " | " | " | $\frac{1}{s^2}$ |

We now have, denoting by

$$\alpha_{r,0} = \frac{\mu_{r,0}}{\mu_{20}^{r/2}}$$

the "standardized" moment of $x$ , by a simple transformation of equation (8).

$$(8') \quad \alpha_{20} = 1\,; \quad \alpha_{30} = \gamma_{30}\,; \quad \alpha_{40} = \gamma_{40} + 3\,;$$
$$\alpha_{50} = \gamma_{50} + 10\gamma_{30}\,; \quad \alpha_{60} = \gamma_{60} + 15\gamma_{40} + 10\gamma_{30}^2 + 15$$

Stopping with quantities of the order $\frac{1}{s}$ we get

$$(13) \quad \alpha_{30} = 10\gamma_{30}\,; \quad \alpha_{40} = 15\gamma_{40} + 10\gamma_{30}^2 + 15$$

In practice we can, of course, not very well know if the hypothesis of elementary increments is valid, but if we have, on computing the moments up to the fourth order, found that $\gamma_{30}$ and $\gamma_{40}$ are rather small, and that $\gamma_{22}$ is of the order of magnitude of $\gamma_{30}^2$, there is a certain plausibility in assuming that $\gamma_{50}$ and $\gamma_{60}$ are still smaller and that they may be neglected as compared to $\gamma_{40}$ and $\gamma_{30}^2$.

The curve of cubic regression of $y$ on $x$ we may write in the form

$$t_y = c_0 + c_1 t_x + c_2 t_x^2 + c_3 t_x^3$$

where we have put

$$t_x = \frac{x - M_x}{\sqrt{\mu_{20}}} \quad ; \qquad t_y = \frac{y_x - M_y}{\sqrt{\mu_{20}}}$$

and it is evident that equation (1) now takes the form

$$\begin{aligned}
0 &= c_0 + c_2 + c_3 \alpha_{30} \\
r &= c_1 + c_2 \alpha_{30} + c_3 \alpha_{40} \\
\alpha_{21} &= c_0 + c_1 \alpha_{30} + c_2 \alpha_{40} + c_3 \alpha_{50} \\
\alpha_{31} &= c_0 \alpha_{30} + c_1 \alpha_{40} + c_2 \alpha_{50} + c_3 \alpha_{60}
\end{aligned}$$

We get

$$(4)\ \Delta = \alpha_{60}(\alpha_{40} - \alpha_{30}^2 - 1) - \alpha_{50}(\alpha_{50} - 2\alpha_{30}\alpha_{40} - 2\alpha_{30}) - \alpha_{40}(\alpha_{40}^2 - \alpha_{40} + 3\alpha_{30}^2) + \alpha_{30}^4$$

$$\Delta_1 = r(\alpha_{30}\alpha_{60} - \alpha_{30}\alpha_{40}) - \alpha_{21}(\alpha_{60} - \alpha_{40}^2) + \alpha_{31}(\alpha_{50} - \alpha_{40}\alpha_{30}) - r\alpha_{30}(\alpha_{30}\alpha_{50} - \alpha_{40}^2) + \alpha_{21}\alpha_{30}(\alpha_{50} - \alpha_{30}\alpha_{40}) - \alpha_{31}\alpha_{30}(\alpha_{40} - \alpha_{30}^2)$$

$$\Delta_2 = r(\alpha_{40}\alpha_{60} - \alpha_{50}^2 - \alpha_{60} + 2\alpha_{30}\alpha_{50} - \alpha_{30}^2\alpha_{40}) - \alpha_{21}(\alpha_{30}\alpha_{60} - \alpha_{40}\alpha_{50} + \alpha_{30}\alpha_{40} - \alpha_{30}^3) + \alpha_{31}(\alpha_{30}\alpha_{50} - \alpha_{40}^2 + \alpha_{40} + \alpha_{30}^2)$$

$$\Delta_3 = -r(\alpha_{30}\alpha_{60} - \alpha_{50}\alpha_{40} + \alpha_{40}\alpha_{30} - \alpha_{30}^3) + \alpha_{21}(\alpha_{60} - \alpha_{40}^2 - \alpha_{30}^2) - \alpha_{31}(\alpha_{50} - \alpha_{30}\alpha_{40} - \alpha_{30})$$

$$\Delta_4 = r(\alpha_{50}\alpha_{30} - \alpha_{40}^2 + \alpha_{40} - \alpha_{30}^2) - \alpha_{21}(\alpha_{50} - \alpha_{30}\alpha_{40} - \alpha_{30}) + \alpha_{31}(\alpha_{40} - \alpha_{30}^2 - 1)$$

And the coefficients are

$$c_0 = \frac{\Delta_1}{\Delta} \qquad c_1 = \frac{\Delta_2}{\Delta} \qquad c_2 = \frac{\Delta_3}{\Delta} \qquad c_4 = \frac{\Delta_4}{\Delta}$$

We now introduce the semi-invariants by (8'), taking for $\alpha_{50}$ and $\alpha_{60}$ the approximate formulae (13). For $\alpha_{21}$ and $\alpha_{31}$ we put

$$(15) \qquad \alpha_{21} = \gamma_{21} \quad ; \quad \alpha_{31} = \gamma_{31} + 3r$$

The coefficients $\gamma_{21}$ and $\gamma_{31}$ are then the standardized correlation semi-invariants, according to a generalized theory of semi-invariants for bi-variate distributions.

It is now a consequence of our principle of approximation that all powers and products $\gamma_{ij}, \gamma_{kl}, \gamma_{mn} \ldots$, of which the sum $i+j+k+l+m+n+\ldots$ of the indices exceeds 6, shall be neglected as compared to powers and products of lower order. Observing this, the determinants reduce to the following:

$$\Delta = 12(1 - 2\gamma_{30}^2 + 2\gamma_{40}),$$

or

$$\frac{1}{\Delta} = \frac{1}{12}(1 + 2\gamma_{30}^2 - 2\gamma_{40}),$$

$$\Delta_1 = 6(r\gamma_{30} - \gamma_{21}),$$

$$\Delta_2 = 12r + 6(r\gamma_{40} - \gamma_{31}) + 24r\gamma_{40} - 24r\gamma_{30}^2 - 12\gamma_{30}(r\gamma_{30} - \gamma_{21}),$$

$$\Delta_3 = -6(r\gamma_{30} - \gamma_{21}),$$

$$\Delta_4 = -2(r\gamma_{40} - \gamma_{31}) + 6\gamma_{30}(r\gamma_{30} - \gamma_{21}).$$

Using the same rule of approximation on multiplying by $\frac{1}{\Delta}$, we finally get

$$
\begin{aligned}
c_0 &= \tfrac{1}{2}(r\gamma_{30} - \gamma_{21}) , \\
c_1 &= r + \tfrac{1}{2}(r\gamma_{40} - \gamma_{31}) - \gamma_{30}(r\gamma_{30} - \gamma_{21}) , \\
c_2 &= -\tfrac{1}{2}(r\gamma_{30} - \gamma_{21}) , \\
c_3 &= -\tfrac{1}{6}(r\gamma_{40} - \gamma_{31}) + \tfrac{1}{2}\gamma_{30}(r\gamma_{30} - \gamma_{21}) .
\end{aligned}
\tag{16}
$$

In my cited memoir of twelve years ago I put[1]

$$r_{30} = \tfrac{1}{2}(r\gamma_{30} - \gamma_{21}); \quad r_{40} = \tfrac{1}{6}(r\gamma_{40} - \gamma_{31}),$$

Using this notation, we get

$$
\begin{aligned}
c_0 &= r_{30} , \\
c_1 &= r - 3r_{40} - 2\gamma_{30} r_{30} \\
c_2 &= -c_0 = -r_{30} , \\
c_3 &= r_{40} + \gamma_{30} r_{30} .
\end{aligned}
\tag{17}
$$

These coefficients are exactly the same as in equation (34*, II) of my former memoir. As shown in that memoir on several numerical examples, the regression formula in question applies very well in cases of moderately skew correlations.

It is seen that the coefficients $r_{30}$ and $r_{40}$ determine the curvature of the regression. If $r_{30} = r_{40} = 0$ the regression is linear (of the third order). I have called these coefficients the correlation coefficients of higher order. If the correlation surface is approximately normal we have the following formulae for the standard errors of the coefficients involved:

1. In Pearson's notation we have $r_{30} = \frac{1}{2}\bar{\epsilon}$ and $r_{40} = \frac{1}{6}\bar{\zeta}$.

$$\sigma_{(\gamma_{30})}=\sqrt{\frac{6}{N}}\,;\quad \sigma_{(\gamma_{40})}=\sqrt{\frac{24}{N}}\,;\quad \sigma_{(\gamma_{21})}=\sqrt{\frac{2+4r^2}{N}}\,;\quad \sigma_{(\gamma_{31})}=\sqrt{\frac{6+18r^2}{N}}$$

$$(18)\quad \sigma_{(r)}=\frac{1-r^2}{\sqrt{N}}\,;\quad \sigma_{(r_{30})}=\sqrt{\frac{1-r^2}{2N}}\,;\quad \sigma_{(r_{40})}=\sqrt{\frac{1-r^2}{6N}}\,;\quad \sigma_{(c_0)}=\sqrt{\frac{1-r^2}{2N}}\,;$$

$$\sigma_{(c_1)}=\sqrt{\frac{(5-2r^2)(1-r^2)}{2N}}\,;\quad \sigma_{(c_2)}=\sqrt{\frac{1-r^2}{2N}}\,;\quad \sigma_{(c_3)}=\sqrt{\frac{1-r^2}{6N}}$$

Lund (Sweden).