

**The significance of the correlation coefficient when applied to Mendelian distributions / by John Brownlee.**

**Contributors**

Brownlee, John, 1868-1927.  
Royal Society of Edinburgh.  
University of Glasgow. Library

**Publication/Creation**

Edinburgh : Robert Grant & Son ; London : Williams & Norgate, 1910.

**Persistent URL**

<https://wellcomecollection.org/works/wagz4v6a>

**Provider**

University of Glasgow

**License and attribution**

This material has been provided by This material has been provided by The University of Glasgow Library. The original may be consulted at The University of Glasgow Library. where the originals may be consulted. Conditions of use: it is possible this item is protected by copyright and/or related rights. You are free to use this item in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s).



Wellcome Collection  
183 Euston Road  
London NW1 2BE UK  
T +44 (0)20 7611 8722  
E [library@wellcomecollection.org](mailto:library@wellcomecollection.org)  
<https://wellcomecollection.org>

REPRINT FROM THE

PROCEEDINGS

OF THE

ROYAL SOCIETY OF EDINBURGH.

SESSION 1909-1910.

---

VOL. XXX.—PART VI.—(No. 34.)

---

The Significance of the Correlation Coefficient when  
applied to Mendelian Distributions.

By John Brownlee, M.D., D.Sc.

EDINBURGH :

PUBLISHED BY ROBERT GRANT & SON, 107 PRINCES STREET, AND  
WILLIAMS & NORGATE, 14 HENRIETTA STREET, COVENT GARDEN, LONDON

MDCCCX.

*Price Two Shillings.*

THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO



XXXIV.—The Significance of the Correlation Coefficient when applied to Mendelian Distributions. By John Brownlee, M.D., D.Sc.

(MS. received February 22, 1910. Read January 24, 1910.)

1. AT the present moment there is much discussion regarding the means by which properties are hereditarily transmitted from a parent organism to its offspring, and of the extent to which the Mendelian theory is capable of accounting for the facts. In this note it is not proposed to discuss the general question but to investigate the conditions under which the theory of correlation may be applied to Mendelian groupings. Two important papers on this subject have already been published: one by Professor Pearson, entitled "A Generalized Theory of Mendelian Inheritance";\* the other, which is largely a criticism of this, by Professor Udny Yule.† In Professor Pearson's paper the results produced when two organisms with any number of pairs of different zygotes mate indiscriminately are fully considered. He finds that such a population once established is stable, and he then deduces the parental and fraternal correlation coefficients. He finds that the parental correlations are independent of the number of zygotes, and also that the coefficients are considerably inferior in value to the numbers actually found by observation. Professor Yule, in criticism, says that the observed value of the coefficients can be obtained if a certain amount of weight is given to the effect of the hybrid and recessive elements, and he gives a formula in which this result is exhibited.

2. Professor Yule's criticism suggests that if the Mendelian theory is true, great care will be required in interpreting the meaning of a correlation coefficient, and the purpose of this paper is to investigate how far values of the latter can be taken as representations of real relationships. As Professor Pearson has shown that the simplest Mendelian formula has the same regression as the more complex, it is unnecessary for me to repeat his mathematical proofs, the case of the mating of two organisms differing in one particular giving the information required.

3. Professor Yule has pointed out there are several varieties of correlation possible on a Mendelian basis. The chief, however, are, (1) where the hybrid has properties of its own differentiating it from either of its parents;

\* *Royal Soc. Trans.*, 1903, p. 53.

† "On the Theory of Inheritance of Quantitatively Compound Character on the Basis of Mendel's Laws," by G. Udny Yule. Report of Conference on Genetics, published by Royal Horticultural Society of London.



and (2) where the dominant includes the hybrid. It is obvious that the correlation between parent and offspring will be much greater in the former case than in the latter. This argument will be made clearer if the elementary Mendelian formula is examined. In the first place, consider a population consisting of two pure races. Let them be denoted by (a, a) and (b, b) respectively and let (a, b) be the hybrid between them. Then the whole population may be expressed by a parentage of both sexes each represented by

$$x^2 (a, a) + 2xy (a, b) + y^2 (b, b),$$

where  $x^2$ ,  $2xy$  and  $y^2$  denote the numbers respectively of each type. If mating is random and fertility equal, we have offspring in the following proportions:—

$x^2 (a, a)$	mating with $x^2 (a, a)$	gives $x^4 (a, a)$
„	„ $2xy (a, b)$	„ $x^3y (a, a) + x^2y^2 (a, b)$
„	„ $y^2 (b, b)$	„ $x^2y^2 (a, b)$
$2xy (a, b)$	„	„ $x^2 (a, a)$ „ $x^3y (a, a) + x^2y^2 (a, b)$
„	„ $2xy (a, b)$	„ $x^2y^2 (a, a) + 2x^2y^2 (a, b) + y^2x^2 (b, b)$
„	„ $y^2 (b, b)$	„ $xy^3 (a, b) + xy^3 (b, b)$
$y^2 (b, b)$	„	„ $x^2 (a, a)$ „ $x^2y^2 (a, b)$
„	„ $2xy (a, b)$	„ $xy^3 (a, b) + xy^3 (b, b)$
„	„ $y^2 (b, b)$	„ $y^4 (b, b)$

Adding together and arranging the terms, we have the population of offspring given by

$$x^2(x+y)^2 (a, a), \quad 2xy(x+y)^2 (a, b), \quad y^2(x+y)^2 (b, b),$$

or the numbers of the offspring are in the same proportions as those of the parents; that is, the population is stable. Stability, then, depends on the number of the hybrid being equal to twice the geometric mean of the number of the pure races. It is also easily shown that even though these proportions are not originally present they at once appear.

4. When these figures are arranged so as to show the correlation from parent to child the following table is formed:—

NUMBER OF PARENTS OF EACH TYPE.

Number of Offspring of each Type.	(a, a).	(a, b).	(b, b).
(a, a) . .	$x^4 + x^3y$	$x^3y + x^2y^2$	
(a, b) . .	$x^3y + x^2y^2$	$x^3y + 2x^2y^2 + xy^3$	$x^2y^2 + xy^3$
(b, b) . .	...	$x^2y^2 + xy^3$	$xy^3 + y^4$

Dividing by the common factor  $x + y$  this becomes

PARENTS.			
Offspring.	(a, a).	(a, b).	(b, b).
(a, a) .	$x^3$	$x^2y$	
(a, b) .	$x^2y$	$xy(x+y)$	$xy^2$
(b, b) .	...	$xy^2$	$y^3$

In this table the regression is linear, and therefore the correlation between parent and offspring may be determined by the product method and is given by

$$r = \cdot 5.$$

This shows that in a stable population the correlation is independent of the relative proportions of purer races. Now in ascertaining the correlation when the hybrid can be distinguished from the dominant the process given above is correct, but when the hybrid has no points of special distinction and must therefore be included in the dominant, the table is condensed to the following:—

PARENTS.		
Offspring.	(a, a) + (a, b).	(b, b).
(a, a) + (a, b) .	$x^3 + 3x^2y + xy^2$	$xy^2$
(b, b) .	$xy^2$	$y^3$

Here the regression is linear as shown by Professor Pearson, so that by the product method

$$r = \frac{y}{x + 2y},$$

or,

$$= \cdot 333 \text{ when } x = y.$$

5. By repeating the above process the correlation of offspring with remoter ancestors can be easily evaluated. The first hypothesis, namely, that the hybrid is independent of the dominant, leads to correlation of  $\cdot 5$ ,  $\cdot 25$ ,  $\cdot 125$ , etc., or, in other words, they are there given by Galton's Law of Ancestral Inheritance.\* On the second hypothesis, the one investigated by

\* Professor Pearson, *Royal Soc. Trans.*, vol. cxv. p. 119, Table IX., "Exclusive Inheritance."



Professor Pearson, the same correlation coefficients are represented by  $\frac{1}{3}$ ,  $\frac{1}{6}$ ,  $\frac{1}{12}$ ,  $\frac{1}{24}$ , etc. The well-known correlations found by observation have no obvious relation to either of these sets of figures, and if Mendel's law is proved to be efficient, some means of reconciling theory and observation must be found. In the subsequent pages the various factors which influence correlation will be considered under different heads.

INFLUENCE OF THE DIFFERENT METHODS OF CALCULATING CORRELATION COEFFICIENTS ON THE VALUES DEDUCED IF MENDELIAN PRINCIPLES HOLD.

6. When the typical correlation table for parent and offspring given by Mendelian theory is considered it is evident that it shows several properties. If, say, the population consist of

$$(a, a), \quad (a, b), \quad (b, b),$$

then it may be tabulated in two ways :

Pure (a, a) containing two *a* elements ;  
 Hybrid (a, b)      „      one *a* element ;  
 Pure (b, b)      „      no *a* element ;

or if the hybrid (a, b) resemble (a, a) in appearance we have (a, a)+(a, b) not having a pair of *b* zygotes and (b, b) possessing a pair of *b* zygotes. Both these forms have linear regression, and in consequence the product method of determining correlation is valid. The case already given may be repeated. Taking *x* equal to *y* the correlation of parent and offspring reduces to the following simple form :—

PARENT.

Offspring.	(a, a).	(a, b).	(b, b).	Totals.
(a, a) .	1	1	...	2
(a, b) .	1	2	1	4
(b, b) .	...	1	1	2
Totals .	2	4	2	8

This table shows obvious symmetry, has evidently linear regression, and gives a correlation coefficient between parent and offspring of  $r = .5$ . But if the table is further condensed, that is, if (a, a) and (a, b) are considered as one class we have instead :—

PARENT.

Offspring.	(a, a)+(a, b).	(b, b).	Totals.
(a, a)+(a, b)	5	1	6
(b, b) . .	1	1	2
Totals .	6	2	8

Here again the regression is linear, and as the result we have

$$\cdot 333.$$

So far all is clear. In the last case, however, the distribution is markedly skew, and while the product method is applicable it is only applicable because the regression is linear.

7. It is therefore specially important to consider what happens when other methods of obtaining the correlation are employed. The chief of these is the fourfold division method. In a Mendelian instance such as this, the fourfold table seems specially applicable, but it assumes normality of distribution so that the fourfold table should give a higher correlation than  $r = \cdot 3333$ . As a matter of fact it does. The equation for determining  $r$  is

$$\cdot 62035 = r + \cdot 22747r^2 + \cdot 04951r^3 + \cdot 12279r^4 + \cdot 001898r^5 + \dots$$

which gives

$$r = \cdot 53.$$

That is to say, the correlation is even higher than that obtained when the hybrid is distinguishable from the dominant, and in applying the fourfold method we have returned to or even gone beyond the uncondensed table. The higher coefficients are likewise increased and the series becomes

Parental.	Grand-parental.	Great-grandparental.	Great-great-grandparental.
$\cdot 53,$	$\cdot 29,$	$\cdot 15,$	and $\cdot 073$
as against			
$\cdot 5,$	$\cdot 25,$	$\cdot 125,$	and $\cdot 063.$

8. If the simple Mendelian table be again considered, and if for the moment the distinguishing character of the hybrid and the dominant be assumed somewhat indefinite, we can make several tentative divisions, either bisecting the hybrid or dividing it into such divisions that one-fourth resembles the recessive as follows:—



PARENT.				PARENT.			
M. Offspring.	(a, a).	(a, b).	(b, b).	N. Offspring.	(a, a).	(a, b).	(b, b).
(a, a) .	4	2	2	(a, a) .	4	3	1
(a, b) .	2	2	2	(a, b) .	3	5	1
	2	2	2		1	1	1
(b, b) .		2	2	(b, b) .		3	1
			4				4

giving fourfold distributions,

PARENT.		PARENT.	
M. Offspring.		N. Offspring.	
10	6	15	5
6	10	5	7

leading to correlations

$$\text{M. } r = \cdot 441,$$

$$\text{N. } r = \cdot 501,$$

when calculated by the fourfold method. Thus, again, Mendelian principles do not lead to low correlations but to figures approximately equal to those found by observation.

9. When more complex formulæ are taken the result is nearly the same. Supposing that instead of one pair of zygotes the parents possess two or three, that is, we have

Dominant.	Recessive.
Father.	Mother.
(a, a)	(b, b)
(c, c)	(d, d)
(e, e)	(f, f)

and let mating be random, then the correlation table in the case of two pairs of zygotes becomes

PARENTS.			
Offspring.	Two Pairs of Dominants.	One Pair of Dominants.	No Dominants.
Two pairs .	25	10	1
One pair .	10	12	2
None .	1	2	1

admitting of two fourfold divisions, namely:—

B.		C.	
25	11	57	3
11	17	3	1

The former of these gives

$$r = \cdot 45,$$

and the latter

$$r = \cdot 45,$$

both values much in excess of the .333 given by the product method.

When the three pairs are involved we have:—

PARENT.

Offspring.	Three Pairs.	Two Pairs.	One Pair.	None.
Three pairs	125	75	15	1
Two pairs .	75	105	33	3
One pair .	15	33	21	3
None . .	1	3	3	1

This form is capable of three different fourfold divisions, namely:—

A.		B.		C.	
125	91	380	52	497	7
91	205	52	28	7	1

Giving

$$\begin{aligned} \text{A. } r &= \cdot 42, \\ \text{B. } r &= \cdot 42, \\ \text{C. } r &= \cdot 45. \end{aligned}$$

10. It is evident that when two and three pairs of zygotes are condensed we do not go straight back to the normal distribution. The reason of this is that the normal surface obtained when the elements are considered separately, represents something different from the surface which is condensed into the last tables.



If the parents be  $\begin{vmatrix} a, a \\ c, c \end{vmatrix}$  and  $\begin{vmatrix} b, b \\ d, d \end{vmatrix}$  then the offspring having two elements from the same parents are—

P.	Q.	R.
$\begin{vmatrix} a, a \\ d, d \end{vmatrix}$	$\begin{vmatrix} c, c \\ b, b \end{vmatrix}$	$\begin{vmatrix} a, b \\ c, d \end{vmatrix}$

which represent different things according as dominance exists or not; for if dominance exist R is included among those having apparently two pairs of dominant zygotes, while if the hybrid is distinct it is grouped with P and Q as containing two units from the same parent.

11. In addition to the methods just given Professor Pearson has also discovered two methods of determining correlation by means of what he calls contingency. It is not necessary to go fully into this part of the question. The manner in which the results given by these methods differ from those just considered is illustrated in the subjoined table. They are not in general suitable for simple Mendelian cases, as they depend for success on the number of divisions being much more numerous than these tables give.

TABLE SHOWING THE CORRELATION COEFFICIENTS CALCULATED BY DIFFERENT METHODS WHERE ONE, TWO, OR THREE DOMINANT ZYGOTES OCCUR IN ONE PARENT AND A LIKE NUMBER OF RECESSIVE IN THE OTHER.

	Product Method.	Mean Square Contingency.	Mean Contingency.	Fourfold Table.		
				A.*	B.*	C.*
One zygote . . .	.333	.32	.37	...	...	.5
Two zygotes . . .	.333	.33	.41	...	.46	.46
Three zygotes . . .	.333	.32	.39	.42	.42	.45

\* See par. 9.

#### RESULTS OF ASSORTIVE MATING.

12. With the same notation as just used the most general form of correlation under a Mendelian system for assortive mating between husband and wife, if the standard deviation of each is equal, is the following:—

##### HUSBANDS.

Wives.	(a, a).	(a, b).	(b, b).	Totals.
(a, a) .	$m$	$2r$	$n$	$m+n+2r$
(a, b) .	$2r$	$4p$	$2r$	$4(r+p)$
(b, b) .	$n$	$2r$	$m$	$m+2r+n$
Totals.	$m+2r+n$	$4(r+p)$	$m+2r+n$	

When the hybrid is distinct from the dominant the value of the correlation coefficient depends only on the value of  $m$ ,  $n$ , or  $r$ , though in the case when the hybrid is not distinct the value of  $p$  exercises an influence on the result. In a typical simple Mendelian distribution of the population  $m+n$  will be equal to  $2r$  and  $p$  to  $r$ . Those values, however, do not give an immediately stable population, the standard deviation of the offspring being higher than that of the parents. This population, however, quickly tends to stability. On the other hand, if the population is immediately stable it is easily seen that  $p$  must be equal to  $n$ , for the first generation gives a parentage and offspring as below:—

PARENT.

Offspring.	(a, a).	(a, b).	(b, b).	Totals.
(a, a) .	$m+r$	$r+p$	...	$m+2r+p$
(a, b) .	$r+n$	$2(r+p)$	$r+n$	$4r+2p+2n$
(b, b) .	...	$(r+p)$	$m+r$	$m+2r+p$
Totals .	$m+2r+n$	$4(r+p)$	$m+2r+n$	$2m+2n+8r+4p$

and as the total is the same whether the addition is made by columns or by rows, the sum of each row must be equal to the sum of the corresponding column if the standard deviation remains the same.

Or,

$$m+2r+p = m+2r+n,$$

which requires that  $n$  shall be equal to  $p$ .

13. In the first place, the varieties of the correlation coefficients when  $m+n=2p$  will be considered. In this case, changing the letters for convenience, the initial correlation table between husband and wife may be taken to be:—

HUSBANDS.

Wives.	(a, a).	(a, b).	(b, b).	Totals.
(a, a) .	$n-a$	$n$	$a$	$2n$
(a, b) .	$n$	$2n$	$n$	$4n$
(b, b) .	$a$	$n$	$n-a$	$2n$
Totals .	$2n$	$4n$	$2n$	$8n$

. . . (a)



If the hybrid is distinct from the dominant the correlation of husband and wife is given by—

$$r = .5 - \frac{a}{n}.$$

If the dominant include the hybrid, then the table condenses to—

HUSBANDS.

Wives.	(a, a)+(a, b).	(b, b).	Totals.
(a, a)+(a, b) .	$5n - a$	$n + a$	$6n$
(b, b). . .	$n + a$	$n - a$	$2n$
Totals . . .	$6n$	$2n$	$8n$

giving a correlation

$$r = \frac{2}{3} \left( .5 - \frac{a}{n} \right).$$

14. If the parentage be as in (a) the correlation table for parent and offspring is—

PARENT.

Offspring.	(a, a).	(a, b).	(b, b).	Totals.
(a, a) .	$\frac{3}{2}n - a$	$n$	...	$\frac{5}{2}n - a$
(a, b) .	$\frac{1}{2}n + a$	$2n$	$\frac{1}{2}n + a$	$3n + 2a$
(b, b) .	...	$n$	$\frac{3}{2}n - a$	$\frac{5}{2}n - a$
Totals .	$2n$	$4n$	$2n$	$8n$

giving a correlation—

$$r_{f.o.}^* = \frac{3n - 2a}{\sqrt{(4n \cdot 5n - 2a)}},$$

reducing if  $a = 0$  to  $r = .5$ , *i.e.* there is no assortive mating, or to

$$r_{f.o.} = .596 \text{ if } r_{f.m.} = .25.$$

15. The population given by the parentage (a) is evidently represented by offspring in the proportion

$$\frac{5}{2}(n - a) (a, a) + (3n + 2a) (a, b) + \frac{5}{2}(n - a) (b, b),$$

\*  $r_{f.o.}$  signifies correlation of father and offspring.

$r_{f.m.}$       "      "      "      "      mother.

which has a higher standard deviation than the parentage, being equal in the latter case to .5 and in the former to

$$\frac{5n - 2a}{8n}, \text{ or, if } a = \frac{1}{4}n, \text{ to } .5625.$$

This latter value is not, however, constant with such mating, but increases gradually up to a limit.

16. In addition to the correlation coefficients the contingency coefficients have also been calculated in some instances to show the degree of correspondence of the two. It is seen that for the parental correlations they fall short of the former, but approach them closely when they arrive at great-grand-parental correlations. Three sets of figures have been calculated for each case.

Case 1. That when there is no assortive mating.

Case 2. That when there is assortive mating with equal fertility and population not immediately stable.

Case 3. That when there is assortive mating with equal fertility and an immediately stable population.

#### 17. TABLES OF PARENTAL, ETC., CORRELATIONS BASED ON DIFFERENT HYPOTHESES.

i. The hybrid separate :—

	No Assortive Mating.		Assortive Mating. $r = .25.$		Assortive Mating : Immediately Stable Population. $r = .125.$ $r = .25.$	
	Product Method.	Contingency Method.	Product Method.	Contingency Method.	Product Method.	Product Method.
Parental . . . . .	.5	.487	.589	.576	.563	.625
Grandparental . . . .	.25	.242	.366	.343	.316	.391
Great-grandparental . .	.125	.124	.234	.223	.178	.244
Grt.-grt.-grandparental .	.0625	.0624	.143	.142	.100	.153

ii. The dominant including the hybrid :—

	No Assortive Mating.	Assortive Mating. $r = .25.$	Assortive Mating : Immediately Stable Population. $r = .1875.$
Parental . . . . .	.3333	.495	.548
Grandparental . . . .	.1667	.307	.351
Great-grandparental . .	.0833	.203	.236
Grt.-grt.-grandparental .	.0417	.141	.161



It is to be noted that column 2 in Table ii. gives almost exactly the figures found by observation and would thus appear a possible expression of the facts, though it is more probably a mere coincidence, as will be shown later.

18. Two more cases of importance remain to be considered: that where like mates unlike, and that where the dominant includes the hybrid. Taking that where like mates unlike and reversing the mating given in par. 13, we have:—

HUSBANDS.

Wives.	(a, a).	(a, b).	(b, b).
(a, a) .	$a$	$n$	$n - a$
½ (a, b) .	$n$	$2n$	$n$
½ (b, b) .	$n - a$	$n$	$a$

If the population of offspring be then found and the correlation calculated we find that—

$$r = \frac{1 + 2\frac{a}{n}}{2\left(3 + 2\frac{a}{n}\right)^{\frac{1}{2}}}.$$

TABLE OF VALUES. (*Hybrid distinct.*)

Value of $\frac{a}{n}$ .	Correlation, Husband and Wife.	Correlation, Parent and Offspring.
·000	— ·500	·289
·125	— ·375	·347
·250	— ·250	·401
·375	— ·125	·452
·500	0	·500
·675	·125	·546
·750	·250	·593
·875	·375	·631
1·000	·500	·671

This table also gives the effect of assortive mating when it is positive as well as negative.

19. When the dominant includes the hybrid and the assortive mating is confined to the mixture we have then a correlation table as the following:—

HUSBANDS.

Wives.	(a, a).	(a, b).	(b, b).
(a, a) .	$m$	$2m$	$n$
(a, b) .	$2m$	$4m$	$2n$
(b, b) .	$n$	$2n$	$m$

which gives the parent and offspring table:—

PARENT.

Offspring.	(a, a).	(a, b).	(b, b).
(a, a) .	$2m$	$2m$	...
(a, b) .	$m+n$	$3m+n$	$2n$
(b, b) .	...	$m+n$	$m+n$

reducing to—

PARENT.

Offspring.	(a, a)+(a, b).	(b, b).
(a, a)+(a, b)	$8m+2n$	$2n$
(b, b) . . .	$m+n$	$n+m$

or

PARENT.

Offspring.	(a, a)+(a, b).	(b, b).
(a, a)+(a, b)	$8+2\alpha$	$2\alpha$
(b, b) .	$1+\alpha$	$1+\alpha$

$$\text{if } \alpha = \frac{n}{m}.$$



TABLE OF VALUES OF THE CORRELATION OF PARENT AND OFFSPRING FOR DIFFERENT  
VALUES OF  $\frac{n}{m}$  BY FOURFOLD TABLE METHOD.

Values of $\alpha$ .	Assortive Mating.	Parent-Offspring Correlation.
·500	·454	·666
·667	·287	·621
·750	...	·593
1·000	·000	·539
1·5	— ·315	·454
2	— ·525	·397

20. The values of the grandparental coefficients can likewise be evaluated, but the labour is somewhat greater than in the previous sections, and does not seem to promise any results beyond what can be surmised from the previous argument. In this case a moderate degree of assortive mating in the parents has apparently little effect on the correlation coefficients.

21. In general it is to be noted that a large variety of different values of the correlation coefficients arises on different hypotheses, and also that the correlation of parent and offspring differs greatly according to the kind of assortive mating of the parents, so that the value of the coefficient of assortive mating gives very little guide to the value of correlation between parent and offspring. It is also to be noted that the successive heredity correlation coefficients are not in an exact geometrical progression.

#### EFFECT OF PARENTAL SELECTION ON THE CORRELATION COEFFICIENT.

22. The effect of parental selection has been investigated by Professor Pearson on the basis of the normal curve of error. On this basis it is shown that the higher the parental selection the lower the correlation coefficients. This, however, does not seem to follow on a Mendelian mechanism. Three cases occur on this basis which require to be considered separately: (1) Where the dominant is present in excess or defect; (2) where the hybrid is present in excess or defect; (3) where the recessive is present in excess or defect. These are very easily evaluated.

The correlation tables here, however, are different from those which go before. Regression is not linear, so that the product method does not give an exact but only an approximate value of the correlation coefficient.

23. CASE I.—Let  $m$   $(a, a) + 2(a, b) + (b, b)$  be the population of the selected parent and  $p$   $\{(a, a) + 2(a, b) + (b, b)\}$  of the non-selected parent.

These are equal if  $m+3=4p$ ; but  $p$  may be neglected as occurring in every term and therefore not affecting the result.

The correlation table for the selected parents and offspring, if mating be random, is then the following:—

SELECTED PARENTS.

Offspring.	(a, a).	(a, b).	(b, b).	Totals.
(a, a) .	$m$	1	...	$m+1$
(a, b) .	$m$	2	1	$m+3$
(b, b) .	...	1	1	2
Totals .	$2m$	4	2	$2m+6$

This gives

$$r_{s.o.}^* = \sqrt{\frac{6m+2}{m^2+14m+17}}$$

if the hybrid be distinct, or to

$$r'_{s.o.} = \frac{m+1}{2(m+2)}$$

if the dominant include the hybrid;

reducing if

$$m=1 \text{ to } r=.5,$$

and to

$$r'=.333,$$

respectively, as before seen to be the case.

The correlation table for the non-selected parent and the offspring is:—

NON-SELECTED PARENT.

Offspring.	(a, a).	(a, b).	(b, b).	Totals.
(a, a) .	$m+1$	$m+1$	...	$2m+2$
(a, b) .	2	$m+3$	$m+1$	$2m+6$
(b, b) .	...	2	2	4
Totals .	$m+3$	$2m+4$	$m+3$	$4m+12$

\*  $r_{s.o.}$  signifies the correlation of the selected parent and offspring.

$r_{n.o.}$  " " " non-selected parent and offspring.



This gives

$$r_{n.o.} = \frac{m+3}{\sqrt{2(m^2+14m+17)}}$$

if the hybrid be distinct;

$$r'_{n.o.} = \frac{1}{\sqrt{3m+2}}$$

if the dominant includes the hybrid;

reducing if

$$m = 1 \text{ to } r = .5,$$

and

$$r' = .333.$$

24. CASE II.—In like manner, if the parentage be such that hybrid is in excess or defect we have, if the population of selected parents be

$$(a, a) + 2m, (a, b) + (b, b),$$

and the population of non-selected parents

$$p\{(a, a) + 2, (a, b) + (b, b)\},$$

$$r_{s.o.} = \frac{1}{\sqrt{2(m+1)}}$$

if the hybrid be distinct;

$$r'_{s.o.} = \frac{1}{3(2m+1)}$$

if the hybrid be included in dominant.

The correlations in this case of the non-selected parents and offsprings are constant and identical with those where there is no selection, namely,  $r = .5$  and  $r' = .333$ , for the correlation table for the non-selected parents when written out is as follows:—

NON-SELECTED PARENT.

Offspring.	(a, a).	(a, b).	(b, b)
(a, a) .	$m+1$	$m+1$	...
(a, b) .	$m+1$	$2m+2$	$m+1$
(b, b) .	...	$m+1$	$m+1$

and  $m+1$  being a factor throughout, the result is not affected.

25. CASE III.—If the recessive be in excess or defect we take again,

the population of the selected parent,  $(a, a) + 2, (a, b) + m (b, b)$ ,

and                      „                      „ non-selected parent,  $p\{(a, a) + 2, (a, b) + (b, b)\}$ .

From this the correlations, if the hybrid is distinct, are obviously the same as in Case I., but if the hybrid be included in dominant then we have—

$$r'_{s.o.} = \sqrt{\frac{4m}{3(m+1)(m+5)}}$$

$$r'_{n.o.} = \frac{(m+1)^{\frac{1}{2}}}{\sqrt{3(m+5)^{\frac{1}{2}}}}$$

26. The values of the correlation coefficients on these bases as  $m$  varies are given in the following tables.

TABLE I.—CORRELATION OF PARENTS (SELECTED AND NON-SELECTED) AND OFFSPRING WHERE THE HYBRID IS DISTINCT FROM THE DOMINANT.

$m$ .	Dominant or Recessive in Excess or Defect.		Hybrid in Excess or Defect.	
	$r_{s.o.} = \sqrt{\frac{6m+2}{m^2+14m+17}}$	$r_{n.o.} = \frac{2(m+3)}{2(m^2+14m+17)}$	$r_{s.o.} = \frac{\sqrt{2}}{2(m+1)}$	$r_{n.o.} \cdot 5$ .
0	343	515	702	5
25	413	507	632	5
5	454	503	577	5
75	481	501	534	5
100	500	500	500	5
15	523	502	447	5
20	536	505	408	5
25	540	510	378	5
30	542	516	342	5
4	541	525	316	5
5	534	534	289	5
6	526	544	267	5
$\infty$	0	702	0	

TABLE II.—CORRELATION OF PARENTS (SELECTED AND NON-SELECTED) AND OFFSPRING WHERE THE HYBRID IS INCLUDED IN THE DOMINANT.

$m$ .	Dominant in Excess or Defect.		Hybrid in Excess or Defect.		Recessive in Excess or Defect.	
	$r'_{s.o.} = \frac{m+1}{2(m+2)}$	$r'_{n.o.} = \frac{1}{\sqrt{3(m+2)^3}}$	$r'_{s.o.} = \frac{1}{\sqrt{3(2m+1)^3}}$	$r'_{n.o.} = .333$	$r'_{s.o.} = \sqrt{\frac{4m}{3(m+1)(m+5)}}$	$r'_{n.o.} = \frac{(m+1)^{\frac{1}{3}}}{\sqrt{3(m+5)^{\frac{1}{3}}}}$
0	250	408	578	333	0	258
25	278	385	471	333	225	281
50	300	365	408	333	284	301
75	318	348	365	333	319	320
100	333	333	333	333	333	333
150	357	308	289	333	350	357
200	375	289	258	333	356	378
250	388	273	235	333	356	394
300	400	257	218	333	353	408
400	411	236	192	333	344	430
500	429	218	179	333	333	447
600	437	204	160	333	236	460
$\infty$	5	0	0	...	0	577

27. Considering the values of the correlation coefficients in these tables, we see that uni-parental selection except when large makes little



difference in the correlation. Selection may raise or lower the correlation. In some cases there is a maximum and in others a minimum, these points being in general not far distant from the points of normal Mendelian distribution of the population. Selective mating is not, then, likely to interfere with the correlation coefficients to any appreciable extent except when the selection is stringent.

28. There are a few other cases which demand attention, some of which will be referred to when the actual figures are discussed, while some others are added in this place.

29. CASE (A).—If both parents be equally selected and if the parentage is given being

$$\begin{array}{lll} m(a, a) & 2(a, b) & (b, b) \\ m(a, a) & 2(a, b) & (b, b), \end{array}$$

we have as the correlation of either parent and offspring,

$$= \frac{1}{2} \sqrt{\frac{3m+1}{2(m+1)}}$$

when the hybrid is distinct.

TABLE OF VALUES.

<i>m.</i>	<i>r.</i>		<i>m.</i>	<i>r.</i>
0	·353		1·5	·522
·5	·456		2	·577
1·0	·500		∞	·612

30. CASE (B).—If the hybrid be present in normal numbers but the recessive present in defect and the dominant in corresponding excess. In other words, both parental populations consist of

$$(1+m)(a, a) \quad 2(a, b) \quad (1-m)(b, b).$$

This gives a correlation coefficient when the hybrid is distinct of

$$r = \sqrt{\frac{2-m^2}{2(4-m^2)}}$$

*m* being always less than unity.

TABLE OF VALUES.

<i>m.</i>	<i>r.</i>		<i>m.</i>	<i>r.</i>
0	·500		·6	·474
·2	·498		·8	·450
·4	·490			

31. CASE (C).—Let the race be made up of such a population that a part only of the hybrid assumes dominant characters, that is, let it consist of parental populations of  $(1+m)$  (apparently dominant),  $(2-m)$  (hybrid),  $(1)$  (recessive), and let it mate indiscriminately.

CASE (a).—Let the hybrid offspring be distinguishable at birth, developing the resemblance to the dominant later, a condition frequently seen. The correlation table is as follows:—

PARENT.			
Offspring.	(a, a).	(a, b).	(b, b).
(a, a) .	$2 + m$	$2 - m$	...
(a, b) .	$2 + 2m$	$4 - 2m$	2
(b, b) .	$m$	$2 - m$	2

which gives a correlation coefficient between parents and offspring at birth of the latter,

$$r = \frac{\sqrt{2}}{(8 + 4m - m^2)^{\frac{1}{2}}}.$$

TABLE OF VALUES.

$m = 0$	$r = .500$	$m = 1$	$r = .426$
$m = .5$	$r = .453$	$m = 1.5$	$r = .412$

32. CASE (b).—Let a normal population (a, a), 2 (a, a), (a, b), mate at random, and let dominance appear among the offspring later. The normal correlation table,

PARENT.			
Offspring.	(a, a).	(a, b).	(b, b).
(a, a) .	2	2	...
(a, b) .	2	4	2
(b, b) .	...	2	2

then becomes—

PARENT.			
Offspring.	(a, a).	(a, b).	(b, b).
(a, a) .	$2 + m$	$2 + 2m$	$m$
(a, b) .	$2 - m$	$4 - 2m$	$2 - m$
(b, b) .	...	2	2



giving the same correlation coefficient as before, namely,

$$r = \frac{\sqrt{2}}{(8 + 4m - m^2)^{\frac{1}{2}}}.$$

With the increase of  $m$  the correlation becomes less. The values of  $r$  are given under Case (a).

#### CORRELATION COEFFICIENTS WHEN MORE THAN TWO RACES MIX.

33. So far, a mixture of two races alone has been considered. Many stocks of cattle, etc., are supposed to be derived from more than two, so that a brief consideration of how this affects the correlation values is necessary. With the same notation let the original races be—

$$(a, a) \quad (b, b) \quad (c, c).$$

Then the stable population with random mating is as before,

$$(a, a) + (b, b) + (c, c) + 2(a, b) + 2(a, c) + 2(b, c).$$

A correlation table is then easily written down and is as follows:—

		PARENT.					
Offspring.		(a, a).	(a, b).	(b, b).	(b, c).	(c, c).	(c, a).
(a, a)	.	3	3				3
(a, b)	.	3	6	3	3		3
(b, b)	.		3	3	3		
(b, c)	.		3	3	6	3	3
(c, c)	.				3	3	3
(c, a)	.	3	3		3	3	6

To evaluate the correlation the product method hitherto used is inapplicable, and the method of contingency must be employed. In the first place, on the supposition that all hybrids are distinct, we have  $r = .597$ , which is considerably higher than the value  $r = .487$ , found by contingency when only two types of parent are considered.

Secondly—

34. If  $a$  be dominant over  $b$ ,  $b$  over  $c$ , and  $c$  over  $a$  (indicated in table by dotted lines), the coefficient when estimated by mean square contingency falls in value to .425. This case is very suitable for a fourfold division,

and if  $a$  and  $b$  be gathered against  $c$ , allowing for dominance, the correlation coefficient rises to a value of  $r = .51$ , and when the mean contingency is used to  $r = .56$ .

Thirdly—

35. If  $b$  and  $c$  are both dominant over  $a$  and the hybrid ( $b, c$ ) is distinct, the correlation becomes .460.

Thus in all cases we have a higher figure than in the case where only two types intermingle. As Professor Pearson has shown, the figures in the latter case are quite independent of the number of zygotes, and the like will probably hold here.

TABLE SHOWING THE CORRELATION BETWEEN PARENT AND OFFSPRING IN TWO AND THREE RACES.

	Two Races.		Three Races.	
	Correlation.	Contingency.	Contingency.	Fourfold Division.
Hybrid distinct . . . .	.500	.487	.597	
Hybrid included in Dominant	.333	.316	.425	.51

The same effects will also be produced in this case by assortive mating and parental selection as in the previous cases.

#### FRATERNAL CORRELATION.

36. The question of fraternal correlation remains to be considered. As we have seen, uni-parental selection does not in general affect seriously the values of the correlation coefficients. Assortive mating is more powerful. The effect of the latter on fraternal correlation can be estimated as follows.

Consider a parentage of the following arrangement :—

#### HUSBANDS.

Wives.	(a, a).	(a, b).	(b, b).
(a, a) .	3	4	1
(a, b) .	4	8	4
(b, b) .	1	4	3

This gives a correlation of .25 between husbands and wives. With Professor Pearson let the average families be  $4n$ , and we get a family grouping as follows :—



## CHILDREN.

Number of Times each Fraternal Group occurs.	(a, a).	(a, b).	(b, b).
Father (a, a) . . . $\left\{ \begin{array}{l} 3 \\ 4 \\ 1 \end{array} \right.$	$4n$ $2n$ $\dots$	$\dots$ $2n$ $4n$	
Father (a, b) . . . $\left\{ \begin{array}{l} 4 \\ 8 \\ 4 \end{array} \right.$	$2n$ $n$ $\dots$	$2n$ $2n$ $2n$	$n$ $2n$
Father (b, b) . . . $\left\{ \begin{array}{l} 1 \\ 4 \\ 3 \end{array} \right.$	$\dots$ $\dots$ $\dots$	$4n$ $2n$ $\dots$	$2n$ $4n$

On re-arranging we have each group occurring as in the table.

## BRETHREN.

Number of Times a Group occurs.	(a, a).	(a, b).	(b, b).
3	$4n$		
8	$2n$	$2n$	
2	$\dots$	$4n$	
8	$n$	$2n$	$n$
8	$\dots$	$2n$	$2n$
3	$\dots$	$\dots$	$4n$

So that we can write the correlation table for brothers as follows:—

## FIRST BROTHER.

Second Brother.	(a, a).	(a, b).	(b, b).
(a, a) .	$3 \cdot 4n(4n-1) + 8 \cdot 2n(2n-1) + 8n(n-1)$	$8(4n^2 + 2n^2)$	$8n^2$
(a, b) .	$8(4n^2 + 2n^2)$	$2 \cdot 4n(4n-1) + 24 \cdot 2n(2n-1)$	$8(4n^2 + 2n^2)$
(b, b) .	$8n^2$	$8(4n^2 + 2n^2)$	$3 \cdot 4n(4n-1) + 8 \cdot 2n(2n-1) + 8n(n-1)$

Or dividing by  $4n$ ,

## FIRST BROTHER.

Second Brother.	(a, a).	(a, b).	(b, b).
(a, a) . .	$22n - 9$	$12n$	$2n$
(a, b) . .	$12n$	$32n - 14$	$12n$
(b, b) . .	$2n$	$12n$	$22n - 9$

This gives a correlation as below :—

Size of Family.		Assortive Mating $r_{f.m.} = .25.$	No Assortive Mating.
4	$n = 1$	.407	.333
8	$n = 2$	.508	.428
16	$n = 3$	.515	.454
$\infty$	$n = \infty$	.555	.500

That is, if the hybrid be distinct from the dominant, and an assortive mating of the parents equivalent to .25 is assumed, the correlation coefficients quickly approach the figures given by observation.

37. Taking the dominant to include the hybrid we require a different parental grouping to give the necessary correlation, namely :—

## HUSBANDS.

Wives.	(a, a).	(a, b).	(b, b).
(a, a) . .	7	8	1
(a, b) . .	8	16	8
(b, b) . .	1	8	7

This has a correlation of  $= \frac{2}{3}(\frac{5}{8} - \frac{1}{8}) = .25$  when the dominant includes the hybrid. Proceeding as before, we obtain the table of fraternal correlation :—

## FIRST BROTHER.

Second Brother.	(a, a).	(a, b).	(b, b).
(a, a) . .	$(48n - 19)$	$24n$	$4n$
(a, b) . .	$24n$	$56n - 26$	$24n$
(b, b) . .	$4n$	$24n$	$48n - 19$



Or condensing,

FIRST BROTHER.

Second Brother.	(a, a) + (a, b).	(b, b).
(a, a) + (a, b)	$152n - 45$	$28n$
(b, b) . . .	$28n$	$48n - 19$

Which gives the correlation coefficients as in the following table:—

Size of Family.	$n =$	Correlation Coefficients with Assortive Mating. $r = .25.$	Correlation as calculated by Prof. Pearson with no Assortive Mating.
4	1	.317	
8	2	.401	.333
16	3	.429	.364
$\infty$	$\infty$	.476	.407

The fraternal correlation is not therefore increased so much by assortive mating as the parental-offspring correlation is. The resulting figure is still in defect of observation.

38. The same process may be applied to ascertain the correlation coefficients when three races mix.

If a standard population is taken and the method just outlined applied we get the following correlation table:—

FIRST BROTHER.

Second brother.	(a, a).	(a, b).	(b, b).	(b, c).	(c, c).	(c, a).
(a, a) . . .	$16n - 9$	$8n$	$n$	$2n$	$n$	$8n$
(a, b) . . .	$8n$	$36n - 18$	$8n$	$9n$	$2n$	$9n$
(b, b) . . .	$n$	$8n$	$16n - 9$	$8n$	$n$	$2n$
(b, c) . . .	$2n$	$9n$	$8n$	$36n - 18$	$8n$	$9n$
(c, c) . . .	$n$	$2n$	$n$	$8n$	$16n - 9$	$8n$
(c, a) . . .	$8n$	$9n$	$2n$	$9n$	$8n$	$36n - 18$

Then if  $n = 1$  the contingency coefficient is  $r = .449$ , and when  $n = 2$ ,  $r = .569$ , much higher values, which will be further increased if assortive mating

exists in addition; and even when reduced by the inclusion of the hybrid with the dominant they must approach those given by observation.

39. The effect of parental selection on fraternal correlation remains to be considered. Referring to the parentages before given with reference to the correlation of offspring and parent, the two chief cases are given.

Case I. Let the parentage on both sides be  $m(a, a) + 2(a, b) + (b, b)$ , and let the pure zygotes  $(a, a)$  be in excess or defect; and,

Case II. Let the parentage on both sides be  $(a, a) + 2m(a, b) + (b, b)$ , and let the hybrid  $(a, b)$  be in excess or defect.

Then if  $n=2$ , *i.e.* if the family be 8 on an average, we have the fraternal correlation as in the accompanying table:—

FRATERNAL CORRELATION. (*Hybrid distinct.*)

Value of $m$ .	Case I.	Case II.
·5	·333	·514
1·0	·428	·428
2·0	·523	·347
3·0	·572	·314

40. Thus, such selection as that when the dominant is in excess or the hybrid is in defect tends to raise the correlation, while the opposite condition tends to lower it. If both conditions exist, and if the parentage be such that the dominant is twice as numerous and the hybrid half as numerous as in the stable population, we have  $r=.70$  when hybrid is distinct and  $r=.40$  (product method) when dominant includes the hybrid.

#### CONSIDERATION OF ACTUAL CASES.

We have seen that many different factors affect the value of the correlation coefficients. What effect these have practically can only be estimated in a few cases. Professor Pearson has considered three cases of colour inheritance, namely:—

1. Coat colour in horses.\*
2. Coat colour in cattle.†
3. Coat colour in greyhounds.‡

Each of these cases will be briefly discussed and the divergences of value in the correlation coefficients explained as far as possible on the basis of what has gone before.

\* *Roy. Soc. Trans.*, vol. cxcv. p. 92. *Biometrika*, vol. i. p. 361; vol. ii. p. 230 *et seq.*

† *Biometrika*, vol. iii. p. 245 *et seq.*

‡ *Ibid.*, vol. iv. p. 427 *et seq.*



## COAT COLOUR IN HORSES.

This case may well be considered first, as the data are large and probably accurate. Stud books giving the colour and pedigree of the horse have been in existence for many years, while the value of the animals and the great interest which exists in breeding combine to give the facts authority.

To find the correlation Professor Pearson has divided the parents and offspring into groups of Bay and Darker, and Chestnut and Lighter, and calculated the coefficients by the fourfold method; the coefficients as determined by him are as follows:—

## INHERITANCE OF COAT COLOUR IN HORSES.

Parental . . . . .	·5216
Grandparental . . . . .	·2976
Great-grandparental . . . . .	·1922
Great-great-grandparental . . . . .	·1469

Now brown and bay seem both dominant over chestnut and white,\* at least to all intents and purposes. Chestnut with chestnut breeds true, and brown or bay mating with chestnut breeds in the first instance dark. The relations of brown and bay do not concern us, being both dominant. The number of pale horses not chestnut is so small that it may be neglected as not affecting the result to any appreciable extent. The proportion of these colours present is roughly that of three dark horses to one chestnut, though it must be borne in mind that this has nothing directly to do with Mendelism, but represents simply the proportions which find favour at present among those who breed horses.

It is worth while reproducing the fourfold tables. That of parent and offspring is as follows †:—

	Bay or Darker.	Chestnut or Lighter.	Totals.
Bay or darker . . . . .	631	125	756
Chestnut or lighter . . . . .	147	147	294
Totals . . . . .	778	272	1000

This table at once reminds us of that already found from Mendel's theory, namely (pars. 6 and 7):—

\* Bateson, *Mendel's Principles of Heredity*, p. 124.

† *Roy. Soc. Trans.*, vol. clxxv. p. 35.



PARENT.

Offspring.	Dominant.	Recessive.	Totals.
Dominant .	5	1	6
Recessive .	1	1	2
Totals .	6	2	8

which when evaluated by the fourfold method gives  $r = .53$  as the correlation. As a matter of fact the table just quoted gives  $r = .54$ .

If the highest ancestral coefficient is now examined we find some difference. The table for great-great-grandparental inheritance \*—

GREAT-GREAT-GRANDPARENTS.

Offspring.	Bay and Darker.	Chestnut and Lighter.	Totals.
Bay and darker . . .	497	252	749
Chestnut and lighter .	130	99	229
Totals . . . . .	627	351	978

is marked by the presence of a great excess in chestnut horses. As before shown (par. 25),† this tends to raise the correlation of parent and offspring. The effect of this, however, on succeeding generations may be here inquired into. In the case in point we have approximately one-third of the parentage recessive. The remaining two-thirds may be divided in two ways: it may be taken as of pure Mendelian composition, that is, we have one case of pure dominant and two of hybrid dominant; on the other hand, considering that the pure horse may be a better animal than the hybrid, and therefore more likely to be chosen for breeding purposes, we may assume that the number of pure and of hybrid dominants is equal. The parentages on this hypothesis will then be:—

$$2 (a, a) \quad 4 (a, b) \quad 3 (b, b) \quad (A.)$$

and

$$2 (a, a) \quad 2 (a, b) \quad 2 (b, b) \quad (B.)$$

The former (A) will probably give the dominant in defect and the latter (B) in excess, so that some value between the results obtained on these two hypotheses may be taken as true.

\* *Biometrika*, vol. ii. p. 255.

† Cf. also par. 4.



The first generation of parentage (A) mating freely gives offspring in the following proportions:—

PARENT.

Offspring.	(a, a).	(a, b).	(b, b).	Totals.	
(a, a) .	8	8	...	16	(a).
(a, b) .	10	18	12	40	
(b, b) .	...	10	15	25	
Totals .	18, or $2 \times 9$	36, or $4 \times 9$	27, or $3 \times 9$	81	

Which shows that the hybrid offspring are in number twice the geometric mean of the pure races as should be (par. 3). To obtain the next generation with a like parentage an increase in the number of the pure races is required, so that the last table becomes:—

PARENTS.

Offspring.	(a, a).	(a, b).	(b, b).	
(a, a) .	10	10	...	20 ( $2 \times 10$ )
(a, b) .	10	18	12	40 ( $4 \times 10$ )
(b, b) .	...	12	18	30 ( $3 \times 10$ )

Let these mate freely and we have for the correlation table of grandparents and grand-offspring the following distribution:—

GRANDPARENTS.

Grand-offspring.	(a, a).	(a, b).	(b, b).	Totals.	
(a, a) .	300	380	120	800	(b).
(a, b) .	475	895	630	$2000\{\sqrt[2]{(800 \times 1250)}\}$	
(b, b) .	125	525	600	1250	
Totals .	900	1800	1350		

The correlations in these two cases are respectively (a)  $r = .578$ , and (b)  $r = .330$ .\* This process may be continued indefinitely, and likewise results may be obtained on the second hypothesis.

If the excess of recessive be maintained for four generations the correlations are as follows:—

	A.	B.
Parental . . . . .	.578	.638
Grandparental . . . . .	.330	.451
Great-grandparental . . . . .	.200	.309
Great-great-grandparental . . . . .	.110	.216

But the heredity has not been quite this. The proportion of dark and light horses in each generation when means of each parentage are taken has altered in the following manner:—

	Dark.	Light.	Total.
Great-great-grandparents . . . . .	641	359	1000
Great-grandparents . . . . .	664	336	1000
Grandparents . . . . .	712	288	1000
Parents . . . . .	728	272	1000

So that for two generations the proportion of recessive is one-third and above, and for the last two approaching the ratio of one-quarter, though, as before remarked, this is not of a Mendelian origin. If we calculate, then, the correlations for the great-great-grandparents on the hypothesis that the recessive is equal to one-third of the total for two generations and to one-quarter of the total during the next two generations, and for the grandparents that on the hypothesis that the recessive numbers one-third for

\* These and the subsequent correlations have been obtained by the fourfold method though not by the full process. They have been calculated by the formula,

$$r = \sin \frac{\pi}{2} \frac{2}{\sqrt{1+k^2}} \text{ when } k^2 = \frac{4abcdN^2}{(ad-bc)^2(a+d)(b+c)}$$

and where the fourfold division is

a	b
c	d

This formula gives results very near the truth. When those coefficients, previously calculated in this paper by the full method, were checked by the method here referred to, the result has been so close that in the present instance where many coefficients are required the extra labour of calculation has not seemed necessary.



two generations and one-quarter thereafter, the following correlations are obtained.

	A. One Genera- tion '33 and two '25 Recessive.	A. Two Genera- tions '33 and two '25 Recessive.	B. One Genera- tion '33 and two '25 Recessive.	B. Two Genera- tions '33 and two '25 Recessive.
Parental . . . . .	·578	·578	·638	·638
Grandparental . . . .	·315 (·299) *	·340	·351	·451
Great-grandparental . .	·163 (·159) *	·171	·185	·241
Great-great-grandparental .	...	·089	...	·124

It is thus seen that high ancestral coefficients may arise simply from the kind of mating, and when it is noted that even in recent years chestnut horses are present in excess of one-quarter it will be seen that the values given in this table should be exceeded. In fact, the whole is capable of explanation as a result of Mendelism and of method of calculation. In addition to the effects ascertained assortive mating must be considered. If this consists of an excess of like mating like it will raise the correlation (par. 13). Such is the probable mating, and so it is not necessary to assume that the ancestral coefficients are high because of the nature of inheritance; a simple zygote formula is quite sufficient to explain the facts.

#### COAT COLOUR IN CATTLE.

In considering the value of the coefficients of inheritance in coat colour among cattle it is first necessary to see how far the coat changes can be expressed by a Mendelian law. In this instance we have the dominant group apparently divided into three classes: (1) red, (2) red with a little white, and (3) red and white, all of which seem for present purposes the same. In the accompanying table all the matings are given on the assumption that the red class is uniform.

The red class when mated with the white give in general roan, so that in this case the hybrid is distinct. If we represent red by (R, R), white by (W, W), roan will be (R, W). Considering further the mating of red and white we get out of 135 cases 128 roan calves, while the remaining 7 are red. Such a result might be expected on a Mendelian basis. All reds cannot be alike, nor all whites. There must be some variation among them;

\* According to the method in which the population of parents is adjusted.







Returning now to the mating of red and white we have the red parentage

$$a (R, R) + h (R, W);$$

and the white parentage

$$(a + h) (W, W)$$

should give

$$\frac{1}{2}h(a + h) (W, W).$$

In this case  $(a + h)^2 = 135$ , so that 1.45 white cattle should occur. When red mates with roan in like manner about sixteen white calves should occur though only four are found.

These figures are, of course, based on the first group, and if all the groups were given equal weight the number of white to be expected in groups two and three would be less, but then likewise also the numbers of roan in group one.

Two more matings require to be considered, roan and roan, and roan and white. In both these cases it is to be noted that only about half the white turns up which might be expected. This is in line with what has been observed as regards expected whites. It is not necessarily against Mendelism. Many extracted races are comparatively sterile, and if such be proved with regard to white shorthorns it would explain not only the defect in expected whites but also their unpopularity from a breeder's point of view.

Apart altogether from refined theories the general aspect can be explained roughly on a Mendelian basis, and if it is so then the correlation coefficients may be calculated on the principles already enunciated. As the hybrid is distinct the correlation should be (par. 6) .5. Several factors, however, lower this; the dominant is greatly in excess and the recessive in defect (par. 30). This makes a marked difference. Also the recessive only appears in half the number expected when roan and roan, etc., are mated; this also lowers the correlation. For let the population be 2 (a, a), 4 (a, b), 2 (b, b), and let the recessive only appear only in half numbers, and we get a correlation table:—

PARENTS.

Offspring.	(a, a).	(a, b).	(b, b).
(a, a) .	8	8	
(a, b) .	8	16	8
(b, b) .	...	4	4

Which gives  $r = .454$  instead of  $r = .5$ .



Again the mating is unusual. The table of sires and dams for an offspring of colts is as follows:—

SIREs.

Dams.	Red.	Roan.	White.
Red .	197 (217)	277 (276)	45 (27)
Roan .	221 (210)	271 (268)	12 (26)
White .	34 (26)	28 (33)	0 (3)

Alongside the actual figures are placed within brackets those required by random mating. It is seen at once that red matings with white or dominant with recessive are much more numerous than required by chance, a further cause of low correlation (par. 17).

The values of the coefficient may now be considered. By the product method  $r = \cdot 363$ . By the fourfold table when normality is assumed  $r = \cdot 46$ . If, however, the parentage is taken first, the expected Mendelian population of offspring calculated and the correlation evaluated, we find the aberrance of type among the offspring (the absence of sufficient whites) has lowered the correlation to some extent. A typical offspring for the parentage gives  $r = \cdot 383$ . Professor Pearson by the contingency method gets  $r = \cdot 40$ , not greatly in excess of  $\cdot 363$ , and probably arrived at because he has made the calculation with the red group divided into three classes, and with this increase of division the contingency may be expected to give higher figures. In using the contingency method it is clearly not legitimate to break up one class without breaking up others, especially if one class, as seems here, is arbitrarily divided.

One point remains to be considered: What is the correlation between parent and offspring among the different divisions of the dominants, as these, though of the same strain, have considerable variation of colour?

The table for sires and colts is as follows:—

SIREs.

Colts.	Red.	Red with little White.	Red and White.
Red . . . . .	95	27	13
Red with little White .	14	6	6
Red and White . . .	8	4	11



Calculated by the product method, though, this does not seem specially applicable here;  $r = \cdot 337$ . By the fourfold division  $r = \cdot 393$ , if the reds be taken on the one hand and the reds with white on the other. The parent-ages, however, are very unequal. If each be raised to 100 the correlation falls to  $r = \cdot 269$ . So that even among the dominant class there is a considerable hereditary influence. On what basis it is to be explained there are not sufficient facts to indicate. There must be great variation in the zygote constitution and a certain amount of dominance in the dominant class, but to what it amounts would require much investigation.

#### COLOUR INHERITANCE IN GREYHOUNDS.

(*Biometrika*, vol. iii. p. 245. Barrington and Pearson.)

The colour of greyhounds is somewhat complex; the classes used by Barrington and Pearson are: (1) red, (2) brindle, (3) white, (4) fawn, (5) pure black, and (6) mixed black. The exact relationship of these colours is not easily seen from the nature of the offspring. None are clearly dominant, and the hybrid must be largely separated from both dominant and recessive. The fanciers seem to derive the present stock from a mixture of at least three races, red, black, and white, and thus on a Mendelian mechanism the correlation coefficients should be high. The parental assortive mating obtained by the mean square contingency method is about  $r = \cdot 20$ , but its nature is unknown, so that its effect on raising or lowering the correlation coefficient cannot be estimated. The actual correlations obtained by the mean square contingency are as follows:—

CORRELATION BETWEEN PARENT AND OFFSPRING.

	Unselected Offspring.	Offspring selected for Record.
Sire and Dog . . .	$\cdot 512$	$\cdot 474$
Sire and Bitch . . .	$\cdot 579$	$\cdot 404$
Dam and Dog . . .	$\cdot 505$	$\cdot 485$
Dam and Bitch . . .	$\cdot 532$	$\cdot 499$
Mean . . .	$\cdot 532$	$\cdot 466$

Here two classes of correlation are given: one for the whole litters taken at birth and the second for the offspring selected for record. The fall in the correlation is noticeable. The value in the first case is not far from

that given in Case I. (par. 33). The second is nearer that given in Cases II. or III.

That is, the height of the first can be explained on the ground that three races mix with the production of distinct hybrids, that of the second on the ground that dominance of some sort manifests itself with growth, an explanation as possible as that of the authors who attribute the fall in the correlation coefficient to the selection of puppies. The fraternal correlations are also high, being  $r = .676$  for brethren of same litter and  $r = .559$  for the selected record, both much higher than the Mendelian formulæ given in par. 38 warrants. A cause of such high coefficients has been shown in par. 40, where it is noted that if three races mix fraternal correlation is raised.

*(Issued separately August 1, 1910.)*



