Biometrika Trust

On the Effect of Differential Fertility of Degeneracy

Author(s): Karl Pearson

Source: *Biometrika*, Vol. 7, No. 3 (Apr., 1910), pp. 258-275

Published by: [Biometrika Trust](#)

Stable URL: <http://www.jstor.org/stable/2345386>

Accessed: 17/06/2014 18:13

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Biometrika Trust is collaborating with JSTOR to digitize, preserve and extend access to *Biometrika*.

<http://www.jstor.org>

ON THE EFFECT OF A DIFFERENTIAL FERTILITY ON DEGENERACY.

By KARL PEARSON, F.R.S.

A NEW YEAR'S GREETING TO FRANCIS GALTON, 1910.

(1) I HAVE indicated in several earlier papers* the very marked changes which are produced when there is correlation between fertility and any character in a species. I have termed the changes which result in the mean and variability of the character *genetic or reproductive selection*, and have shown that if such correlation exists reproductive selection may wholly defeat or largely neutralise the results of natural selection. In actual wild life, it is very difficult to find any character sensibly correlated with fertility, and it would appear that the low value of correlations of this character is an essential condition for rapid progress by natural selection. Nature has in some way—at present not clear to us—prevented this reversal of the survival of the fitter by suppressing all *marked* correlation between the physical characters and the fertility of a species. Half a century ago, I think, this absence of correlation between fertility and physique would have been found to hold for man in this country. It is practically certain that it does not hold to-day. Artificially a differential fertility has been created; the better mental as well as physical characters can be shown to be associated with a lessened fertility and a reproductive selection has been called into play, which not only impedes, but possibly reverses natural selection. The object of the present paper is to obtain—at any rate to a first approximation—some measure of the secular changes in a race which must flow from a correlation of fertility with any character of an organism. The problem in the case of bisexual reproduction is not a wholly straightforward one. It is influenced in the first place by the intensity of assortative mating in the species. In the next place we have to ask: Does the fertility depend on the intensity of the character in one or in both parents? And lastly we must find a reasonable form for the relation between fertility and the intensity of the character. These points will be considered in the following sections.

* See *Phil. Trans.* Vol. 187, p. 258; Vol. 192, pp. 259, 314; *R. S. Proc.* Vol. 59, p. 303. *The Chances of Death*, Vol. I. p. 63 et seq.

(2) We will suppose that some character—e.g. grade of intelligence—influences the fertility. We have then to consider whether this character in the male and female influences equally or unequally the fertility of the pair. It has been asserted by some that academic training in the woman lessens her power of child-bearing. In this case the more intelligent women would have fewer offspring, but owing to assortative mating such women marry the more intellectual men, and whether intelligence was or was not associated with a lesser grade of fertility in the male, assortative mating would handicap the fertility of the intelligent male, and the more intelligent males would be practically less fertile.

Let ξ_1 and ξ_2 be the deviations from their respective sex means of male and female for any character, σ_1 and σ_2 the standard deviations with respect to this character in the two sexes, represented by the subscripts 1 and 2. Let r_{12} be the correlation of ξ_1 and ξ_2 , i.e. the intensity of assortative mating.

Then it is reasonable to assume that the fertility of the pair is some function of x , where x is a linear function of ξ_1 and ξ_2 , or :

$$x = c_1 \frac{\xi_1}{\sigma_1} + c_2 \frac{\xi_2}{\sigma_2} \dots\dots\dots(i).$$

The mean of x is clearly zero, and we can free ourselves from the influence of either parent by putting c_1 or c_2 zero. If fertility were related to x by a simple linear correlation then we should obtain, it is well known, the highest correlation of fertility (y) and x by taking (the subscript 3 denoting fertility):

$$c_1 = \frac{r_{13} - r_{12}r_{23}}{1 - r_{12}^2}, \quad c_2 = \frac{r_{23} - r_{12}r_{13}}{1 - r_{12}^2} \dots\dots\dots(ii),$$

or proportional to these quantities. We need not at present however assume any special values for c_1 and c_2 . Further, it is desirable to suppose that the curve of mean fertility for each value of x is not necessarily linear, but of a more general type, allowing us to make fertility a maximum at other grades than the extreme values of the character.

(3) This leads us to the next point. What law of fertility seems reasonable? We want a law of fertility which will allow closely of the fertility (a) increasing nearly uniformly and at any given rate with the character, (b) decreasing nearly uniformly and at any given rate with the character, and (c) being concentrated with any degree of intensity round any grade of the character, and falling away on both sides of this maximum. All these conditions are fulfilled if we assume the mean fertility y for any grade of the character x , not to be given by a line but by a normal curve, e.g.

$$y = y_0 e^{-\frac{1}{2}(x-k)^2/\sigma_0^2} \dots\dots\dots(iii).$$

Here by a proper choice of σ_0 and k , we can make the distribution of fertility fall or rise with the character (i.e. k positive or negative and both k and σ_0 very large,— k so taken that the centre of fertility lies outside the region of the range of values of x), or concentrate the fertility on any grade of the character (i.e. make

k equal that value, and take σ_0 small). It appears therefore that the distribution indicated, if not ideal, will still serve as a first approximation to the type of cases likely to occur

There is, however, an argument of still greater weight for the suggestion that a fertility curve of this kind cannot be widely removed from the actual state of the case; namely, when we measure almost any human character (i) in parents and (ii) in their offspring we find both distributions very approximately normal. This could not possibly be the case unless the fertility curve were itself approximately normal. We include of course under this the case in which σ_0 is indefinitely great, or the fertility is not associated with the character in question, which is probably what we may describe as the *organic* condition of affairs. Of course in actual working we have to regard many special features. For example, in the matter of intelligence, it is fairly obvious that the abler men and women have fewer children; but is it true that the fertility falls again towards the extremely dull? This is a point which needs careful consideration; I think it probable, especially when we deal with net and not with gross fertility. Extreme mental defect leads to neglect of offspring, to life in institutions and limited marriage. For example, pauper fertility is higher than criminal fertility. I shall illustrate later this point.

(4) Let us now turn to the direct problem. A parentage has the deviation x of the character and σ_x is the variability of this deviation; the deviation in the offspring is z with a variability σ_z and a correlation ρ . Then, if there were no weighting with a differential fertility, the frequency surface would be

$$u = u_0 e^{-\frac{1}{2} \frac{1}{1-\rho^2} \left(\frac{x^2}{\sigma_x^2} - \frac{2\rho xz}{\sigma_x \sigma_z} + \frac{z^2}{\sigma_z^2} \right)} \dots\dots\dots(\text{iv})$$

in the usual way. But we must multiply this by the average fertility of each grade to obtain the frequency distribution. This will be

$$v = v_0 e^{-\frac{1}{2} \frac{(x-k)^2}{\sigma_0^2}} \times e^{-\frac{1}{2} \frac{1}{1-\rho^2} \left(\frac{x^2}{\sigma_x^2} - \frac{2\rho xz}{\sigma_x \sigma_z} + \frac{z^2}{\sigma_z^2} \right)} \dots\dots\dots(\text{v})$$

$$= v_0 e^{-\frac{1}{2} \frac{k^2}{\sigma_0^2}} \times e^{-\frac{1}{2} \frac{1}{1-\rho^2} \left\{ x^2 \left(\frac{1}{\sigma_x^2} + \frac{1-\rho^2}{\sigma_0^2} \right) - \frac{2xk(1-\rho^2)}{\sigma_0^2} - \frac{2\rho xz}{\sigma_x \sigma_z} + \frac{z^2}{\sigma_z^2} \right\}} \dots(\text{v}) \text{ bis}$$

$$= v_0' \times e^{-\frac{1}{2} \frac{1}{1-R^2} \left\{ \frac{(x-m_1)^2}{\Sigma_1^2} - \frac{2(x-m_1)(z-m_2)R}{\Sigma_1 \Sigma_2} + \frac{(z-m_2)^2}{\Sigma_2^2} \right\}} \dots\dots\dots(\text{vi}),$$

where v_0 is a fresh constant, i.e. independent of x and z , and m_1 and m_2 are the means of effective parentages (i.e. weighted parentages) and of actual sons, Σ_1 and Σ_2 their standard deviations and R their correlation. We have by finding the centre of the elliptic contour of (v) bis in the usual way:

$$m_1 = k\sigma_x^2/(\sigma_0^2 + \sigma_x^2) \dots\dots\dots(\text{vii}),$$

$$m_2 = m_1\rho\sigma_z/\sigma_x \dots\dots\dots(\text{viii}).$$

The first of these equations shows the shifting of the parental mean, the change being due to differential fertility; the second equation shows that the shift of the

offspring mean exactly follows the shift of the parental mean, or the mean of the offspring is the offspring of the mean parent. This gives at once the secular change due to weighting the parentages with their fertility.

Further we must have

$$\begin{aligned}\frac{1}{(1-R^2)\Sigma_1^2} &= \frac{1}{1-\rho^2} \left(\frac{1}{\sigma_x^2} + \frac{1-\rho^2}{\sigma_0^2} \right), \\ \frac{1}{(1-R^2)\Sigma_2^2} &= \frac{1}{1-\rho^2} \left(\frac{1}{\sigma_z^2} \right), \\ \frac{R}{(1-R^2)\Sigma_1\Sigma_2} &= \frac{\rho}{1-\rho^2} \frac{1}{\sigma_x\sigma_z}.\end{aligned}$$

From these flow readily

$$R = \frac{\rho}{\sqrt{1 + (1-\rho^2) \frac{\sigma_x^2}{\sigma_0^2}}} \dots\dots\dots (ix),$$

$$\Sigma_1 = \frac{\sigma_x}{\sqrt{1 + \frac{\sigma_x^2}{\sigma_0^2}}} \dots\dots\dots (x),$$

$$\Sigma_2 = \sigma_z \sqrt{1 - \frac{\rho^2}{1 + \frac{\sigma_0^2}{\sigma_x^2}}} \dots\dots\dots (xi).$$

These results show us at once that differential fertility: (i) reduces the variability (Σ_2) of the offspring, (ii) reduces the variability of the effective parentages, (iii) reduces the apparent correlation between parentage and offspring. It should be noted that the non-differential fertility is obtained by putting $\sigma_0 = \infty$; thus any differential fertility up to $\sigma_0 = 0$, which denotes breeding from a single grade, will give lowered values for R , Σ_1 and Σ_2 .

While the variabilities and correlation are thus changed, it is worth noting that the regression of offspring on parents, i.e. $R\Sigma_2/\Sigma_1 = \rho\sigma_z/\sigma_x$ and is therefore unchanged. Further the equation to the regression line is

$$z - m_2 = \rho \frac{\sigma_z}{\sigma_x} (x - m_1),$$

or since $m_2 = m_1\rho\sigma_z/\sigma_x$, it follows that

$$z = \rho \frac{\sigma_z}{\sigma_x} x$$

or is identical with the regression line for non-differential fertility. This is, of course, only a case of the general principle found long ago* that selection of a character A does not alter the regression of a second character B on A , but it is interesting to note that the effect of differential fertility is as it were a pushing of the population down its regression line.

* *Phil. Trans.* Vol. 197, pp. 20, 21.

(5) We may next inquire into the mean fertility and variability of fertility in the population.

The fertility of the parentage x is

$$y = y_0 e^{-\frac{1}{2}(x-k)^2/\sigma_0^2} \dots\dots\dots (iii),$$

and the frequency of such parentages is given by

$$w = \frac{N}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}x^2/\sigma_x^2}$$

where N is the total number of such parentages.

Hence, if $\bar{f}$ = the mean fertility :

$$N\bar{f} = \int_{-\infty}^{+\infty} \frac{Ny_0}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left\{\frac{x^2}{\sigma_x^2} + \frac{(x-k)^2}{\sigma_0^2}\right\}} dx,$$

and accordingly :

$$\bar{f} = \frac{y_0\sigma_0}{\sqrt{\sigma_0^2 + \sigma_x^2}} e^{-\frac{1}{2}\frac{k^2}{\sigma_0^2 + \sigma_x^2}} \dots\dots\dots (xii).$$

This is one equation linking up the indeterminate quantities y_0 , σ_0 , k with observable quantities like $\bar{f}$ the mean size of family in the population. We can now find the standard deviation Σ_0 of the size of families. Clearly :

$$\Sigma_0^2 = \text{Sum } (y^2)/N - \bar{f}^2,$$

$$\begin{aligned} N(\Sigma_0^2 + \bar{f}^2) &= \int_{-\infty}^{+\infty} wy^2 dx \\ &= \frac{Ny_0^2}{\sqrt{2\pi}\sigma_x} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{(x-k)^2}{\frac{1}{2}\sigma_0^2}\right)} dx \\ &= \frac{Ny_0^2 \frac{1}{2}\sigma_0}{\sqrt{\frac{1}{2}\sigma_0^2 + \sigma_x^2}} e^{-\frac{1}{2}\frac{k^2}{\frac{1}{2}\sigma_0^2 + \sigma_x^2}} \end{aligned}$$

as before. Hence :

$$\begin{aligned} \Sigma_0^2 &= y_0^2 \left\{ \frac{\sigma_0}{\sqrt{\sigma_0^2 + 2\sigma_x^2}} e^{-\frac{k^2}{\sigma_0^2 + 2\sigma_x^2}} - \frac{\sigma_0^2}{\sigma_0^2 + \sigma_x^2} e^{-\frac{k^2}{\sigma_0^2 + \sigma_x^2}} \right\}, \\ V_{\bar{f}}^2 &= \left(\frac{\Sigma_0}{\bar{f}} \right)^2 = \frac{\sigma_0^2 + \sigma_x^2}{\sigma_0 \sqrt{\sigma_0^2 + 2\sigma_x^2}} e^{\frac{k^2\sigma_x^2}{(\sigma_0^2 + \sigma_x^2)(\sigma_0^2 + 2\sigma_x^2)}} - 1 \dots\dots\dots (xiii). \end{aligned}$$

$V_{\bar{f}}$ is an observable quantity and this equation may be used to find k if σ_0 has been otherwise determined. The previous equation will then give y_0 . We should have to go to the third moment coefficient to determine y_0 , σ_0 and k from the fertility distribution, and it may be doubted whether the result would be of much value; it seems better to endeavour to determine σ_0 and k from observing the average fertilities of those exceeding certain grades of the character.

(6) Determination of the constants of the fertility distribution from the average fertilities of those possessing more than a certain grade of the character.

Let n_h be the number of the population, N , of pairs of parents who exceed the value h of the character x ; then:

$$n_h = \frac{N}{\sqrt{2\pi}\sigma_x} \int_h^\infty e^{-\frac{1}{2}x^2/\sigma_x^2} dx,$$

$$n_h/N = \frac{1}{\sqrt{2\pi}} \int_{h/\sigma_x}^\infty e^{-\frac{1}{2}x'^2} dx'.$$

This equation determines h/σ_x from Sheppard's Tables, since n_h/N is known.

Further if $\bar{f}_h$ be the mean fertility of those over grade h :

$$n_h \bar{f}_h = \int_h^\infty \frac{N}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}x^2/\sigma_x^2} y dx,$$

$$\frac{n_h}{N} \bar{f}_h = \int_h^\infty \frac{y_0}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{(x-k)^2}{\sigma_0^2}\right)} dx$$

$$= \frac{y_0 e^{-\frac{1}{2}\frac{k^2}{\sigma_0^2 + \sigma_x^2}}}{\sqrt{2\pi}\sigma_x} \int_h^\infty e^{-\frac{1}{2}\frac{\sigma_0^2 + \sigma_x^2}{\sigma_x^2 \sigma_0^2} \left(x - \frac{\sigma_x^2 k}{\sigma_0^2 + \sigma_x^2}\right)^2} dx$$

$$= \bar{f} \frac{1}{\sqrt{2\pi}} \int_{H'}^\infty e^{-\frac{1}{2}x'^2} dx' \dots\dots\dots(\text{xiv}),$$

where
$$H' = \frac{\sqrt{\sigma_0^2 + \sigma_x^2}}{\sigma_x \sigma_0} \left(h - \frac{\sigma_x^2 k}{\sigma_0^2 + \sigma_x^2} \right).$$

Now $n_h \bar{f}_h / N \bar{f}$ is clearly known, therefore H' can be found from the tables.

We will suppose two grades of the character to be h_1 and h_2 and the corresponding average fertilities $\bar{f}_{h_1}$ and $\bar{f}_{h_2}$, then H'_1 and H'_2 will be known; and there results

$$H'_1 \frac{\sigma_x \sigma_0}{\sqrt{\sigma_0^2 + \sigma_x^2}} = h_1 - \frac{k \sigma_x^2}{\sigma_0^2 + \sigma_x^2},$$

$$H'_2 \frac{\sigma_x \sigma_0}{\sqrt{\sigma_0^2 + \sigma_x^2}} = h_2 - \frac{k \sigma_x^2}{\sigma_0^2 + \sigma_x^2},$$

$$\frac{H'_1 - H'_2}{h_1 - h_2} = \frac{\sqrt{\sigma_0^2 + \sigma_x^2}}{\sigma_x \sigma_0}$$

Thus:
$$\sigma_0 = \frac{\sigma_x}{\sqrt{\sigma_x^2 \left(\frac{H'_1 - H'_2}{h_1 - h_2} \right)^2 - 1}} \dots\dots\dots(\text{xv}).$$

Further:
$$\frac{H'_1 (h_1 - h_2)}{H'_1 - H'_2} - h_1 = -k \frac{\sigma_x^2}{\sigma_0^2 + \sigma_x^2}$$

$$= -k \frac{\sigma_x^2 (H'_1 - H'_2)^2 - (h_1 - h_2)^2}{\sigma_x^2 (H'_1 - H'_2)^2}$$

Or, finally :

$$k = \frac{H_1'h_2 - H_2'h_1}{H_1' - H_2'} \frac{1}{1 - \frac{(h_1 - h_2)^2}{\sigma_x^2 (H_1' - H_2')^2}} \dots\dots\dots(\text{xvi}).$$

Thus σ_0 and k are determined, and $\bar{f}$ being known, we can find y_0 . In fact

$$y_0 = \bar{f} \frac{H_1' - H_2'}{(h_1 - h_2)/\sigma_x} e^{\frac{\frac{1}{2} \{ (H_1'h_1 - H_2'h_1)^2/\sigma_x^2 \}}{(H_1' - H_2')^2} - \frac{(h_1 - h_2)^2/\sigma_x^2}{(H_1' - H_2')^2}} \dots\dots\dots(\text{xvii}).$$

Thus the constants of the fertility distribution are fully determined.

(7) Determination of the constants of the fertility distribution, from a knowledge of the fertility of certain grades of the population. For example, when $x = -m_1\sigma_x$, let the fertility be y_1 , and when $x = m_2\sigma_x$, let the fertility be y_2 . Then

$$y_1 = y_0 e^{-\frac{1}{2} \left(\frac{m_1\sigma_x + k}{\sigma_0} \right)^2},$$

$$y_2 = y_0 e^{-\frac{1}{2} \left(\frac{m_2\sigma_x - k}{\sigma_0} \right)^2}$$

Let $\sigma_x/\sigma_0 = \lambda$, $k/\sigma_0 = \mu$; we have :

$$\sqrt{\frac{2(\log y_0 - \log y_1)}{\log e}} = m_1\lambda + \mu,$$

$$\sqrt{\frac{2(\log y_0 - \log y_2)}{\log e}} = m_2\lambda - \mu,$$

$$\left. \begin{aligned} \frac{1}{m_1 + m_2} \left\{ \sqrt{\frac{2(\log y_0 - \log y_1)}{\log e}} + \sqrt{\frac{2(\log y_0 - \log y_2)}{\log e}} \right\} &= \lambda \\ \frac{1}{m_1 + m_2} \left\{ m_2 \sqrt{\frac{2(\log y_0 - \log y_1)}{\log e}} - m_1 \sqrt{\frac{2(\log y_0 - \log y_2)}{\log e}} \right\} &= \mu \end{aligned} \right\} \dots\dots(\text{xviii}).$$

The signs of the roots are arbitrary and must be selected so as to accord with the needs of the problem. For example, if $m_1 = m_2 = m$, either λm or μ may be equal to

$$\sqrt{\frac{2(\log y_0 - \log y_1)}{\log e}} \pm \sqrt{\frac{2(\log y_0 - \log y_2)}{\log e}},$$

and which we choose depends on whether the maximum fertility lies between the two grades or outside and beyond the grade of larger fertility.

The above equations assume that y_0 is known, it is the maximum average fertility of any grade in the community. A third equation to determine y_0 from the average fertility of the whole community is provided on p. 262, Equ. (xii). But the equations then become troublesome and can only be solved by approximation. It will be found also that small changes in the fertility curve are not very influential in modifying the main results to be drawn from the equations set forth in this paper.

(8) Numerical Illustrations. I propose in this section to illustrate the main results of the above investigation by selecting possible values for some of the constants. I am far from asserting that these are actually existing values, but they appear not improbable values—say, for the case of ability. We are only at present acquiring knowledge of the relationship between ability and fertility in the case of man, and the examples I give are merely illustrative and intended to indicate how the algebraical formulae can be dealt with. At the same time, I think, they throw some light on rather urgent national problems. It is desirable to bring home to the minds of the thinking classes what it really does mean, if the fittest in any character have not a third of the fertility of the least fit.

Illustration I. The upper decile of a population has an average fertility rate of 2, the 50% showing the lower values of the character an average fertility rate of 6. The fertility of the whole population has a mean value of 5. Find the changes in the population during one generation.

Probability integral tables give at once $h_1 = 1.2816\sigma_x$.

To determine H'_1 , we have

$$n_{h_1} \times \bar{f}_{h_1} / N \times \bar{f} = .1 \times 2 / 5 = .04,$$

therefore

$$H'_1 = 1.7500.$$

Clearly $h'_2 = 0$, $n_{h_2} \bar{f}_{h_2} / N \times \bar{f} = .5 \times 6 / 5 = .6$, from which the tables give us $H'_2 = 2533$.

A rough diagram indicates at once the sign of the quantities and shows that we must have

$$H'_1 = h_1 + \sigma_x^2 k / (\sigma_x^2 + \sigma_0^2)$$

$$H'_2 = \sigma_x^2 k / (\sigma_x^2 + \sigma_0^2) - h_2$$

or in this case

$$1.7500\sigma_x\sigma_0/\sqrt{\sigma_0^2 + \sigma_x^2} = 1.2816\sigma_x + \sigma_x^2 k / (\sigma_0^2 + \sigma_x^2),$$

$$.2533\sigma_x\sigma_0/\sqrt{\sigma_0^2 + \sigma_x^2} = \sigma_x^2 k / (\sigma_0^2 + \sigma_x^2);$$

or, subtracting:

$$1.4967\sigma_x\sigma_0/\sqrt{\sigma_0^2 + \sigma_x^2} = 1.2816\sigma_x,$$

whence

$$\sigma_0 = 1.6579\sigma_x,$$

and on substituting

$$k = .8131\sigma_x.$$

The equation

$$\bar{f} = \frac{y_0\sigma_0}{\sqrt{\sigma_0^2 + \sigma_x^2}} e^{-\frac{1}{2} \frac{k^2}{\sigma_0^2 + \sigma_x^2}}$$

will now give us y_0 . On substituting the values of σ_0 and k and 5 for $\bar{f}$ we find

$$y_0 = 6.439.$$

The distribution of fertility is accordingly given by

$$y = 6.439 e^{-\frac{1}{2} \left(\frac{x - .8131\sigma_x}{1.6579\sigma_x} \right)^2}$$

It is clear therefore that the fertility rises to a maximum at $\cdot8131\sigma_x$ and then falls again, so that the lowest values of the character have a lowered fertility, but are still in excess of the corresponding high values. Thus m_1 , the mean of the effective parentages, is given by

$$m_1 = k\sigma_x^2/(\sigma_0^2 + \sigma_x^2) = \cdot2169\sigma_x,$$

measured *below* the old mean, and

$$\Sigma_1 = \text{standard deviation of effective parentages}$$

$$= \sigma_0\sigma_x/\sqrt{\sigma_x^2 + \sigma_0^2} = \cdot8563\sigma_x,$$

or, the effective parentages have 16 per cent. less variability than individual parentages, i.e. parentages unweighted with their offspring. We can now find the average fertility of the lowest decile of the population:

$$\begin{aligned} \frac{n_{-h_1}\bar{f}_{-h_1}}{N\bar{f}} &= \frac{1}{\sqrt{2\pi}} \int_q^\infty \frac{1}{\Sigma_1} e^{-\frac{1}{2}\left(\frac{x''}{\Sigma_1}\right)^2} dx'' \\ &= \frac{1}{\sqrt{2\pi}} \int_{q'}^\infty e^{-\frac{1}{2}x'^2} dx', \end{aligned}$$

where

$$\begin{aligned} q' &= q/\Sigma_1 = \frac{h_1 - m_1}{\Sigma_1} \\ &= \frac{1\cdot2816 - \cdot2169}{\cdot8563} = 1\cdot2434. \end{aligned}$$

Hence

$$\frac{n_{-h_1}\bar{f}_{-h_1}}{N\bar{f}} = \frac{1}{10} \frac{\bar{f}_{-h_1}}{\bar{f}} = \cdot10686,$$

or,

$$\bar{f}_{-h_1} = 5\cdot34;$$

that is to say, that while the average value of the uppermost decile's fertility is only 2, that of the lowermost decile is 5·34, i.e. is still in excess of the average fertility, 5, of the population. Indeed for only about 6 per cent. of the population with the very lowest values of the character does the fertility fall below some 4 per cent.

If we take N as before for the number of pairs of parents, then the distribution of effective parentages is given by

$$w = \frac{5N}{\sqrt{2\pi}} \frac{1}{\cdot8563\sigma_x} e^{-\frac{1}{2}\left(\frac{x - \cdot2169\sigma_x}{\cdot8563\sigma_x}\right)^2}.$$

We have now to consider the distribution of offspring from these effective parentages. We have to consider ρ of p. 260 in relation to σ_x and σ_z . Four hypotheses are possible:

(i) We may suppose that the character is equally influential in the case of the fertility of both sexes; thus brain capacity and resulting intellectual energy might mark equal reduction of fertility in both sexes. In this case the proper value for x is*

$$x = \sigma_1 \left(\frac{\xi_1}{\sigma_1} + \frac{\xi_2}{\sigma_2} \right).$$

* Here as later it must be borne in mind that a constant factor is of no importance.

This gives at once

$$\sigma_x = \sigma_1 \sqrt{2 \sqrt{1 + r_{12}}}$$

$$\rho_{xz} = \frac{1}{\sqrt{2}} (r_{1z} + r_{2z}) / \sqrt{1 + r_{12}}.$$

Here if the population were stable we should have $\sigma_1 = \sigma_z$, 1 and 2 representing as before mother and father. According to our present knowledge $r_{1z} = r_{2z} = \cdot 4$ to $\cdot 5$, = say, $\cdot 45$, and r_{12} lies between $\cdot 15$ and $\cdot 25$, say, $\cdot 2$. We have then

$$\sigma_x = \sigma_z \times 1\cdot 5492$$

and $\rho_{xz} = \rho$ of our formulae on pp 260—1, = $\cdot 5809$.

(ii) We may suppose that the character is equally influential with the fertility of both sexes, but that there is no assortative mating. In this case $r_{12} = 0$, the coefficients of $\frac{\xi_1}{\sigma_1}$ and $\frac{\xi_2}{\sigma_2}$ must still be equal and we have at once

$$\sigma_x = \sigma_z \times 1\cdot 4142, \quad \rho = \cdot 6364.$$

(iii) It is conceivable that the character affects the fertility in one sex only and that there is no assortative mating. In this case we may take $x = \xi_1$, and if the population is stable $\sigma_x = \sigma_z$, while $\rho = \cdot 45$.

(iv) The character may be supposed to affect one sex only, but there may be assortative mating. In this case, returning to our formula on p. 259, $r_{23} = 0$ and x may be taken proportional to

$$\frac{r_{13}}{1 - r_{12}^2} \frac{\xi_1}{\sigma_1} - \frac{r_{12}r_{13}}{1 - r_{12}^2} \frac{\xi_2}{\sigma_2},$$

or we may write

$$x = \sigma_1 \left(\frac{\xi_1}{\sigma_1} - r_{12} \frac{\xi_2}{\sigma_2} \right),$$

which leads to

$$\sigma_x = \sigma_1 \sqrt{1 - r_{12}^2}, \quad \text{and} \quad \rho_{xz} = \frac{r_{1z} - r_{12}r_{2z}}{\sqrt{1 - r_{12}^2}},$$

or on the same assumptions as before :

$$\sigma_x = \sigma_z \times \cdot 9798, \quad \text{and} \quad \rho = \cdot 3674.$$

In all these cases σ_z and ρ are of course the variability of offspring for the given character and their correlation with the parentage character on the assumption of one child to the parentage. We now proceed to discuss the constants of the offspring distribution as resulting from effective parentages, i.e. the quantities m_2 , Σ_2 and R of our p. 261. Here m_2 is the reduced mean value of the character,

Constants of Offspring Distribution. Illustration I.

Hypothesis ...	(i)	(ii)	(iii)	(iv)
m_2/σ_z ...	$\cdot 1260$	$\cdot 1380$	$\cdot 0976$	$\cdot 0797$
Σ_2/σ_z ...	$\cdot 9539$	$\cdot 9445$	$\cdot 9747$	$\cdot 9818$
R/ρ ...	$\cdot 8976$	$\cdot 9065$	$\cdot 8805$	$\cdot 8721$

i.e. the regression owing to differential fertility, Σ_2 is the variability of the offspring generation and should be compared with σ_z or σ_1 , while R is the correlation between offspring and effective parentages.

An examination of this table shows the following general results: (a) the mean value of the character in the population is reduced 8 to 13 per cent. of its standard deviation according to the hypothesis we make with regard to the influence of the character on fertility. That is to say, the mean character of the population stands one-third to a half-decile lower with each new generation. (b) In the next place the variability is lowered two to five per cent. The probability therefore of extreme deviations is sensibly reduced; but this reduction tells only on excess of the character because the mean of the distribution has been markedly shifted towards defect of the character; in other words, if extremes of the character in excess correspond to noteworthy men, they are much less likely to occur in a population where this differentiation of fertility has set in. Suppose we term one man in a thousand "noteworthy," then such a man corresponds to the grade $3\,090\sigma_1$ in our original population. His position in the offspring generation will be

$$m_2 + 3\,090\sigma_1 = 3\cdot216\sigma_1 = \frac{3\cdot216}{9539}\Sigma_2 = 3\cdot37\Sigma_2.$$

But this corresponds to four such individuals in 10,000, or 1 in 2500; or there will be less than half as many noteworthy men. This is on the first hypothesis; the second hypothesis shows only 1 in 3300, and the fourth or less influential hypothesis 1 in 1700. It will therefore be obvious that if mental powers are hereditary, any differentiation of fertility will not only alter rapidly the average intelligence of the community, but will produce that dearth of noteworthy ability, or of genius, which is fatal to the continued existence under modern competitive conditions of a great national life. The whole system of curves corresponding to this first Illustration is given in Diagram I.

Illustration II. In the previous illustration we considered a case where the maximum fertility did not coincide with the minimum of the character, although the lowest values of the character had markedly higher fertility than the highest. It will be of interest to determine what, if any, differences arise in our general conclusions, if we fix on the fertility of individuals high and low in the scale, instead of on average fertility values. We will suppose that the individual parentage which stands sixth in the 1000 for the given character has only a fertility of 1.5, and that which stands last but six has a fertility of 6.5; the maximum fertility of any grade shall be 7, and this shall occur *below the grade* of the individual parentage with a fertility of 6.5. Using the notation of p. 264, we may take $m_1 = m_2 = 2.5$, say. Hence

$$\begin{aligned} \pm \sqrt{\frac{2(\log 7 - \log 1.5)}{.434,2945}} &= 2.5\lambda + \mu = \pm 1.75525. \\ \pm \sqrt{\frac{2(\log 7 - \log 6.5)}{.434,2945}} &= 2.5\lambda - \mu = \pm .38503. \end{aligned}$$

Since λ must be positive, these give us

$$2.5\lambda = .6851, \quad \mu = 1.0701, \quad \text{or} \quad 2.5\lambda = 1.0701, \quad \mu = .6851.$$

In order that the maximum fertility may be outside the range $-2.5\sigma_x$ to $2.5\sigma_x$, we must take the first pair of roots. We easily find

$$\sigma_0 = 3.6490\sigma_x, \quad k = 3.9049\sigma_x,$$

and from equation (xii) $\bar{f} = 4.2613$.

Hence the fertility distribution is

$$y = 7e^{-\frac{1}{2}\left(\frac{x - 3.9049\sigma_x}{3.6490\sigma_x}\right)^2}.$$

The maximum fertility 7 occurs so far from the mean value of the character, that there are practically no parentages with a low value of the character and a low fertility, i.e. the number of *individual* parentages with an extremely low value of the character and a fertility under 7 is only 1 in the 20,000. We may therefore look upon our curve as one which sensibly describes a fertility constantly decreasing as the value of the character increases.

Turning to the distribution of effective parentages we find :

$$m_1 = .2728\sigma_x, \quad \Sigma_1 = .9644\sigma_x.$$

Thus there is rather a greater shift than in Illustration I. of the average effective parentage, but the parentages are more variable. The distribution of effective parentages is :

$$w = \frac{4.261N}{\sqrt{2\pi} \times .9644\sigma_x} e^{-\frac{1}{2}\left(\frac{x - .2728\sigma_x}{.9644\sigma_x}\right)^2}.$$

The number of *effective* parentages with low values of the character and fertility less than the maximum will be found from determining the area of this w curve for values of x greater than $x = 3.9049\sigma_x - .2728\sigma_x = 3.6321\sigma_x$, or by looking up in the Tables $x' = 3.6321/.9644 = 3.77$; this gives 1 in 12,500, a negligible frequency.

In view of our previous results for Illustration I, it may be of interest to note the average fertilities of the top and bottom deciles of our individual parentages, they are to be found from :

$$\frac{n_h \bar{f}_h}{N \bar{f}} = \frac{1}{\sqrt{2\pi}} \int_{q'}^{\infty} e^{-\frac{1}{2}x'^2} dx',$$

where

$$q' = \frac{h - m_1}{\Sigma_1} = \frac{\pm 1.2816 - .2728}{.9644},$$

whence by aid of the tables we find 2.28 for the average fertility of the top decile of the individual parentages, and 6.30 for the bottom decile. These values may be compared with the corresponding 2 and 5.34 of Illustration I.

We now turn to the distribution of offspring. We have :

$$m_2 = \cdot 2728 \rho \sigma_z, \quad \Sigma_2 = \sigma_z \sqrt{1 - \cdot 06986 \rho^2},$$

$$R/\rho = 3\cdot 6490 / \sqrt{14\cdot 3152 - \rho^2}.$$

We can now determine these in precisely the same manner as they were tabled for the four hypotheses in Illustration I.

Constants of Offspring Distribution. Illustration II.

Hypothesis ...	(i)	(ii)	(iii)	(iv)
m_2/σ_z ...	·1585	·1736	·1228	·1002
Σ_2/σ_z ...	·9881	·9858	·9929	·9953
R/ρ ...	·9759	·9783	·9713	·9689

We see at once from this table that whichever hypothesis we take the variability of the new generation is only lessened about 1 per cent. and the correlation about 3 per cent., values markedly less than in the case of Illustration I. This is obviously due to the fact that the fertility distribution is markedly less concentrated. But the shift of the population average is now 10 to 17 per cent. of the standard deviation, or roughly the average has receded through $\frac{1}{2}$ to $\frac{2}{3}$ of a decile in the character. We may measure the extreme effect of this by inquiring how far "noteworthy" individuals, the individual one in a thousand of the old generation, have been reduced in number. What we have to find is the value of $(m_2 + 3\cdot 090\sigma_1)/\Sigma_2$; for the first hypothesis this is 3·288, which corresponds to 1 in 2000, or the halving of these noteworthy individuals in the second generation, and we have practically the same rate on the second hypothesis. We thus see that on the numbers of this second Illustration, the average population will degenerate considerably faster, 25 per cent. faster approximately, but the dearth of noteworthy individuals will be slightly less. The reader must not, however, conclude from these numbers that our second Illustration represents a less undesirable type of degeneracy. Beside the dearth of the noteworthy, the multiplication of the extremely unfit has to be measured. Let us suppose that the individual whose character has the least value in the 1000 of the original population is by virtue of this deviation a non-useful member of the community—physically or mentally unfit. We must clearly then find out in the second generation to what number of the population this standard applies, i.e. we want $(3\cdot 090\sigma_1 - m_2)/\Sigma_2$. This is equal to 2·967 on the first and 2·958 on the second hypothesis, corresponding in both cases to 1 in 667 instead of 1 in 1000. On the other hand it will be found that in the first illustration the corresponding numbers are 3·107 and 3·125 on hypotheses (i) and (ii) respectively, giving practically 1 in 1000, and the same standard of unfitness repeating itself in the second generation. Thus Illustration II, with a slightly reduced rate of decrease of noteworthy individuals, has an increased rate of unfit individuals, when compared with

Illustration I, and a general or average decadence greater by 25 per cent. It would be idle to assert that one or other corresponds to a greater racial degeneracy until we have settled whether it is more important to a nation to have in the case of a desirable character (i) a high average value, (ii) a larger percentage of noteworthy individuals, or (iii) a smaller percentage of unfit individuals. We do not know at present what relative weight is to be given to these categories; but we can see that in any quantitative theory of the socially desirable characters in a community, they naturally arise from the analysis of a differential fertility. Such factors as dearth of leaders, lowered general intelligence of the community, multiplication of the unfit, are seen, even by a preliminary analysis of this kind, not to be unrelated causes, by aid of which the historian accounts in a vague manner for racial and political changes. They are associated factors, explicable and measurable when we know the outlines of a theory of differential fertility.

No special stress must be laid on the particular numbers chosen above to illustrate this discussion. They are round numbers not wholly inappropriate, perhaps, to what we know of intelligence and fertility in man. They are used to suggest the main outlines of a discussion of such problems. Considering the immense masses of statistical, especially demographic data now being accumulated in all civilised countries, it seems to me that the historian of the future will have the means of testing in a way, never yet feasible, the changing fitness or unfitness of nations. In that future absolute and differential fertility, the abundance or dearth of leaders, the average intelligence of a nation, the reduction or multiplication of its unfit, will be recognised as the basal factors in racial progress, the biological sources of evolution in political history. Then history becoming biological will for the first time be admitted as a branch of scientific inquiry, and will also for the first time provide the requisite training for the statesman.

If we turn to examine the possibilities of immediate work with regard to differential fertility, there are several directions in which investigations can at present be carried out with reasonable hope of success. If we confine ourselves to the artisan class, wages are a very reasonable measure of capacity, and the correlation between these and fertility or fecundity* might fairly easily be ascertained. Intelligence might be measured by craft classification extending from engineer to general labourer. Much may be done, if we pass from the artisan class alone, by comparing the size of families of members of various professions, of those following mercantile pursuits of retail traders and of artisans. Further within a narrow class like the graduates of the Universities, it ought to be feasible actually to tabulate fertility against a scale of intelligence as measured by academic distinction. Fertility of sane and insane stocks, of mentally defective and normal stocks, of alcoholic and sober parentages can be ascertained, although the task, of course, needs careful social inquirers. Lastly grades of physical fitness might well be investi-

* Fertility might be measured by nett or gross family when complete, or for marriages of 15 or 20 years' duration. It would, I think, be possible to measure fecundity by some function of the intervals between the births of the first three or four children.

gated; the family of the athletic might be compared with that of the non-athletic; and those of the smith and the navy with those of the cobbler, tailor, and street hawker. It would, I think, be quite possible even without a reformed census, for which every sociologist should work, to obtain many results of value bearing on differential fertility by inquiries of the above nature.

Diagrams.

Each diagram consists of five normal curves. They are worked out on Hypothesis (i), i.e. that the fertility of the parentage is equally affected by the character in both parents, and that there is an assortative mating in the community. The coefficient of parental heredity is taken as .45 and that of assortative mating as .20. x is measured from the mean of the character in the parentage of the first generation. Expressing everything in terms of σ_1 the variability, say, of males in the first generation we have the following curves:

Diagram I. Representing Illustration I.

Curve I. Distribution of Males in first generation:

$$y = \frac{n}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2} \frac{x^2}{\sigma_1^2}}.$$

Curve II. Distribution of Individual Parentages in first generation:

$$y = \frac{n}{\sqrt{2\pi} 1.55\sigma_1} e^{-\frac{1}{2} \frac{x^2}{(1.55\sigma_1)^2}}.$$

Curve III. Curve of Fertility of Individual Parentages of first generation:

$$y = 6.439 e^{-\frac{1}{2} \left(\frac{x - 1.2603\sigma_1}{2.5697\sigma_1} \right)^2}.$$

(N.B. Only part of this curve is drawn.)

Curve IV. Distribution of Effective Parentages in first generation:

$$y = \frac{5n}{\sqrt{2\pi} 1.3273\sigma_1} e^{-\frac{1}{2} \left(\frac{x - .3362\sigma_1}{1.3273\sigma_1} \right)^2}.$$

Curve V. Distribution of Males in second generation:

$$y = \frac{5n}{\sqrt{2\pi} .9539\sigma_1} e^{-\frac{1}{2} \left(\frac{x - .1260\sigma_1}{.9539\sigma_1} \right)^2}.$$

Diagram II. Representing Illustration II.

Curve I. Distribution of Males in first generation:

$$y = \frac{n}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2} \frac{x^2}{\sigma_1^2}}.$$

Curve II. Distribution of Individual Parentages in first generation:

$$y = \frac{n}{\sqrt{2\pi} 1.55\sigma_1} e^{-\frac{1}{2} \frac{x^2}{(1.55\sigma_1)^2}}.$$

DIAGRAM I. Fertility has its maximum well within range of parental variation.

In this diagram the value of the character is supposed to increase from left to right, but x is positive when it is a defect from the mean value.

DIAGRAM II. Fertility has an almost uniform gradient within the range of parental variation, its maximum lying practically outside that range. In this diagram the value of the character is supposed to increase from left to right, but x is positive when it is a defect from the mean value.

Curve III. Curve of Fertility of Individual Parentages of first generation:

$$y = 7e^{-\frac{1}{2}\left(\frac{x - 6.0526\sigma_1}{5.6560\sigma_1}\right)^2}.$$

(N.B. Only one half of this curve is drawn.)

Curve IV. Distribution of Effective Parentages in first generation:

$$y = \frac{4.261n}{\sqrt{2\pi} \cdot 1.4948\sigma_1} e^{-\frac{1}{2}\left(\frac{x - .4228\sigma_1}{1.4948\sigma_1}\right)^2}.$$

Curve V. Distribution of Males in second generation:

$$y = \frac{4.261n}{\sqrt{2\pi} \cdot .9881\sigma_1} e^{-\frac{1}{2}\left(\frac{x - .1585\sigma_1}{.9881\sigma_1}\right)^2}.$$