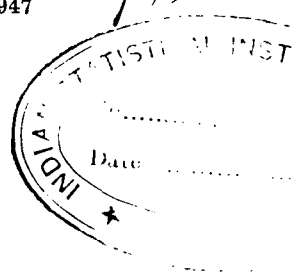
REPRINTED FROM *ANNALS OF EUGENICS*, VOL. 14, PART 1, pp. 35–43, 1947

PRINTED IN GREAT BRITAIN



THE DYSGENIC EFFECT OF INDUCED RECESSIVE MUTATIONS

BY J. B. S. HALDANE

The particles and radiations of high energy produced by radioactivity and nuclear fission can doubtless produce mutations in man, as in other organisms. Such mutations will generally be harmful, though very rarely desirable. Some will be genes or chromosomal rearrangements with wholly or partly dominant effects. These effects will be apparent within a few generations, and independently of inbreeding. In so far as they lower fitness, such genes will disappear relatively rapidly, their frequencies in successive generations forming a geometric series. Thus, the effects of an atomic bomb explosion, in so far as they are due to dominant genes, will appear at once. The delay will be only a little greater in the case of sex-linked recessive genes. Other mutations will be recessive and produce visible effects only as the result of unions between two heterozygotes. It is the object of this paper to show that the visible effects of such induced mutations will reach their maximum after about four or five generations and remain at about the same level for a very long time.

As a preliminary we consider a gene which is completely neutral as regards selective value and which arises by mutation in a very large population. Any fully recessive autosomal gene is neutral until two heterozygotes mate. If one such gene arises by mutation, we have to calculate the probability that it will occur in x loci after n generations. Since Fisher (1930) first posed this problem and solved it formally the distribution of x may be called the Fisher distribution, though of course it is only one of a number described by Fisher.

MOMENTS AND CUMULANTS OF THE FISHER DISTRIBUTION

Let a_r be the probability that a gene carried by one zygote should be present in r zygotes of the nth generation of a large population, generations being supposed to be separate; and let

$$f(t) = a_0 + a_1 t + a_2 t^2 + \ldots.$$

Then $f(1) = 1, f(0) = a_0$, and, since the gene is neutral, then in a stationary population

$$f'(1) = a_1 + 2a_2 + 3a_3 + \ldots = 1.$$

Fisher considered the special case $f(t) = e^{t-1}$, which is particularly appropriate if zygotes are counted in each generation at sexual maturity; however, some information can be obtained in a more general case (Haldane, 1939). Let $p_{n,x}$ be the probability that a mutant gene is present in one locus each of x zygotes in the nth generation. Then Fisher showed that, if

$$\phi(n,t) = p_{n,0} + tp_{n,1} + t^2 p_{n,2} + \ldots,$$

then

$$\phi(n+1,t) = \phi[n, f(t)].$$

If we are concerned with the progeny of a single mutant gene, $\phi(0,t) = t$; if we start with g genes $\phi(0,t) = t^g$.

The mean value of x is unity if $g = 1$, as we shall assume to be the case for the present. After a number of generations by far the commonest value of x is zero, but the mean of the non-zero

values is high. The distribution of the non-zero values will be called the truncated Fisher distribution. Two methods are available for the calculation of their moments. The first is applicable to any admissible function $f(t)$. The nth factorial moment of x about zero is

$$\overline{x^{[n]}} = \left[\left(\frac{d}{dt}\right)^n \phi(n,t)\right]_{(t=1)}$$

Thus $\bar{x} = \phi'(n,1)$, $\overline{x^2} = \phi''(n,1)+\phi'(n,1)$, $\overline{x^3} = \phi'''(n,1)+3\phi''(n,1)+\phi'(n,1)$, etc.

The coefficients are 'difference quotients of powers of zero'. Now

$$\begin{aligned}\phi'(n,t) &= \frac{d}{dt}\phi[n-1,f(t)]\\ &= f'(t)\phi'[n-1,f(t)].\end{aligned}$$

So
$$\begin{aligned}\phi'(n,1) &= f'(1)\phi'(n-1,1)\\ &= [f'(1)]^n\\ &= 1,\end{aligned}$$

if $f'(1) = 1$, as we have assumed to be the case, and shall assume in what follows. Also

$$\phi''(n,t) = [f'(t)]^2\phi''[n-1,f(t)]+f''(t)\phi'[n-1,f(t)].$$

So
$$\begin{aligned}\phi''(n,1) &= \phi''(n-1,1)+f''(1)\\ &= nf''(1),\end{aligned}$$

since $\phi''(0,1) = 0$.

Similarly,
$$\begin{aligned}\phi'''(n,1) &= \phi'''(n-1,1)+3f''(1)\phi''(n-1,1)+f'''(1)\phi'(n-1,1)\\ &= \tfrac{3}{2}n(n-1)[f''(1)]^2+nf'''(1), \quad \text{etc.}\end{aligned}$$

By such methods we find

$$\begin{aligned}\bar{x} &= 1,\\ \overline{x^2} &= nf''(1)+1,\\ \overline{x^3} &= \tfrac{3}{2}n(n-1)[f''(1)]^2+n[f'''(1)+3f''(1)]+1,\\ \overline{x^4} &= 3n(n-1)(n-2)[f''(1)]^3+n(n-1)f''(1)[5f'''(1)+\tfrac{3}{2}\{f''(1)\}^2+9f''(1)]\\ &\quad +n[f^{\mathrm{iv}}(1)+6f'''(1)+7f''(1)]+1, \quad \text{etc.},\end{aligned}$$

whence
$$\left.\begin{aligned}\kappa_1 &= 1,\\ \kappa_2 &= \mu_2 = nf''(1),\\ \kappa_3 &= \mu_3 = \tfrac{3}{2}n(n-1)[f''(1)]^2+nf'''(1),\\ \mu_4 &= 3n(n-1)(n-2)[f''(1)]^3+n(n-1)f'''(1)[5f''(1)+\tfrac{3}{2}\{f''(1)\}^2+3f''(1)]\\ &\quad +n[f^{\mathrm{iv}}(1)+2f'''(1)+f''(1)],\\ \kappa_4 &= 3n(n-1)(n-2)[f''(1)]^3+n(n-1)f''(1)[5f'''(1)+\tfrac{3}{2}\{f''(1)\}^2]\\ &\quad +n[f^{\mathrm{iv}}(1)+2f'''(1)+f''(1)-3\{f''(1)\}^2].\end{aligned}\right\} \quad (1)$$

If $f(t) = e^{t-1}$, then $f'(1) = f''(1) = f'''(1)$, etc. $= 1$. In general this is not the case. Where fertility is very variable, the higher derivatives of $f(1)$ will exceed unity. On the other hand, in a community where there were efficient propaganda and incentives both against sterility and against large families, $f''(1)$, etc., might be found to be less than unity.

ERRATA

Annals of Eugenics, Vol. 14

p. 35, line 7, *delete* 'at once' *insert* 'among the progeny of the survivors'.

p. 38, line 3, *should read*:

whence $$v_n = \tfrac{1}{2}(n + \tfrac{1}{3}\log n + 1{\cdot}79825) + O(n^{-1}\log n).$$

p. 40, *after* '$= g[\exp\{g^{-1}K_{n-1}(\theta)\} - 1]$' *delete paragraph beginning* 'To calculate the cumulants of one generation...' *to line* 5, *p.* 41, *and substitute*:

'Thus $$g^{-1}K_n(\theta) = \exp[g^{-1}K_{n-1}(\theta)] - 1.$$

It follows that successive values of $g^{-1}K_n(\theta)$ are related to one another in exactly the same way as successive values of $K_n(\theta)$ were found to be related when $g = 1$. Since when $n = 1$, $K_1 = g$, it follows that $K_0(\theta) = g\theta$, or $g^{-1}K_0(\theta) = \theta$. Hence, in the nth generation

$$\left.\begin{aligned} \kappa_1 &= g, \\ \kappa_2 &= gn, \\ \kappa_3 &= \tfrac{1}{2}gn(3n-1), \\ \kappa_4 &= \tfrac{1}{2}gn(2n-1)(3n-1), \text{ etc.} \end{aligned}\right\} \tag{6}$$

The moments about zero are:

$$\begin{aligned} \overline{x} &= g, \\ \overline{x^2} &= g^2 + ng, \\ \overline{x^3} &= g^3 + 3ng^2 + \tfrac{1}{2}n(3n-1)g, \\ \overline{x^4} &= g^4 + 6ng^3 + n(9n-2)g^2 + \tfrac{1}{2}n(2n-1)(3n-1)g, \text{ etc.} \end{aligned}$$

Now the probability that x will not be zero in the nth generation is $[f^n(0)]^g = \left(1 - \frac{1}{v_n}\right)^g$. Thus the moments of x' about zero are those of x multiplied by w_n, i.e. by $\left[1 - \left(1 - \frac{1}{v_n}\right)^g\right]^{-1}$. But $gw_n = \tfrac{1}{2}(n + \tfrac{1}{3}\log n + g + 2C - 1 - \tfrac{1}{3}\log 2) + O(n^{-1}\log n)$ and we have for the cumulants of the distribution of x',

$$\begin{aligned} \kappa_1 &= \tfrac{1}{2}(n+g) + \tfrac{1}{6}\log n + 0{\cdot}89913 + O(n^{-1}\log n), \\ \kappa_2 &= \tfrac{1}{4}(n+g)^2 + O(\log n)^2, \\ \kappa_3 &= \tfrac{1}{4}(n+g)^3 + O(n^2\log n), \\ \kappa_4 &= \tfrac{3}{8}(n+g)^4 + O(n^3\log n). \end{aligned}$$

It is clear that when n is large, γ_1 and γ_2 tend to the same values as before, though a good deal more slowly.

$$\begin{aligned} r_{n+1} &= \frac{\alpha(gw_n - 1)}{2w_n} + \frac{(1-\alpha)g(n+g)}{N} \\ &= \alpha\left(\tfrac{1}{2}g - \frac{1}{n}\right) + \frac{g(n+g)}{N} \text{ approximately.} \end{aligned}$$

Thus the frequency is, as before, approximately equal to $\frac{1}{2}\alpha$ multiplied by the number of mutant genes over a long period.'

We now consider the case where $f(t) = e^{t-1}$. The moment-generating function of the Fisher distribution in the nth generation is

$$\begin{aligned} M_n(\theta) &= \phi(n, e^\theta) \\ &= \phi[n-1, f(e^\theta)] \\ &= \phi[0, f^n(e^\theta)] \\ &= f^n(e^\theta), \end{aligned}$$

where the index denotes n-fold iteration of the function f. Thus if $f(t) = e^{t-1}$,

$$M_n(\theta) = f^{n+1}(1+\theta),$$

and
$$K_n(\theta) = \log M_n(\theta) = \log f^{n+1}(1+\theta) = f^n(1+\theta) - 1 = M_{n-1}(\theta) - 1,$$

or
$$M_n(\theta) = \exp[M_{n-1}(\theta) - 1].$$

That is to say the cumulants of the Fisher distribution for the nth generation are its moments about zero for the $(n-1)$th generation. They can be calculated from this peculiar property. The formulae for moments in terms of cumulants (Kendall, 1943, p. 62) give us the cumulants or the moments of the nth generation in terms of those of the $(n-1)$th. If $\Delta\kappa_r$ is the difference between the rth cumulants in successive generations, we have, since $\kappa_1 = 1$ in each generation,

$$\left.\begin{aligned} \Delta\kappa_2 &= 1, \\ \Delta\kappa_3 &= 3\kappa_2 + 1, \\ \Delta\kappa_4 &= 4\kappa_3 + 3\kappa_2^2 + 6\kappa_2 + 1, \\ \Delta\kappa_5 &= 5\kappa_4 + 10\kappa_3\kappa_2 + 10\kappa_3 + 15\kappa_2^2 + 10\kappa_2 + 1, \\ \Delta\kappa_6 &= 6\kappa_5 + 15\kappa_4\kappa_2 + 15\kappa_4 + 10\kappa_3^2 + 60\kappa_3\kappa_2 + 20\kappa_3 + 15\kappa_2^3 + 45\kappa_2^2 + 15\kappa_2 + 1, \text{ etc.} \end{aligned}\right\} \quad (2)$$

By summing the right-hand sides of these equations, we find, for the cumulants of the nth generation, or the moments about zero of the $(n-1)$th generation,

$$\left.\begin{aligned} \kappa_1 &= 1, \\ \kappa_2 &= n, \\ \kappa_3 &= \tfrac{1}{2}n(3n-1), \\ \kappa_4 &= \tfrac{1}{2}n(2n-1)(3n-1), \\ \kappa_5 &= \tfrac{1}{6}n(45n^3 - 65n^2 + 30n - 4), \\ \kappa_6 &= \tfrac{1}{24}n(540n^4 - 1155n^3 + 890n^2 - 273n + 22). \end{aligned}\right\} \quad (3)$$

Of course equations (1) lead to the same results. Clearly κ_r tends to infinity with n^{r-1}, and γ_r with $n^{\frac{1}{2}r}$. The distribution is bimodal, and does not approximate to any of the Pearsonian types.

The truncated Fisher distribution

Let us consider the distribution of those values of x which are not zero, in other words of the numbers of zygotes which carry a mutant gene. In the nth generation $x = 0$ with frequency $f^n(0)$. Writing $v_n = \dfrac{1}{1 - f^n(0)}$, and using x' to denote non-zero values of x, we find $\overline{x'^r} = \overline{x^r} v_n$. Thus from equations (3) which give the values of $\overline{x^r}$ in the $(n-1)$th generation, we find on substituting $n+1$ for n:

$$\overline{x'} = v_n, \quad \overline{x'^2} = (n+1)\,v_n, \quad \overline{x'^3} = \tfrac{1}{2}(n+1\ \ (3n+2)\,v_n, \text{ etc.}$$

Fisher (1930) investigated the behaviour of v_n when $f(t) = e^{t-1}$ and found that

$$v_n = \tfrac{1}{2}n + \tfrac{1}{6}\log_e v_n + 1{\cdot}01464{,}8607 - \frac{1}{72v_n} + O(v_n^{-2}),$$

whence

$$v_n = \tfrac{1}{2}(n + \tfrac{1}{3}\log n + 1{\cdot}7982515) + O(n^{-1}\log n).$$

Hence we find for the cumulants of the distribution of x',

$$\kappa_1 = v_n,$$
$$\kappa_2 = (n+1)\,v_n - v_n^2,$$
$$\kappa_3 = \tfrac{1}{2}(n+1)\,(3n+2)\,v_n - 3(n+1)\,v_n^2 + 2v_n^3,$$
$$\kappa_4 = \tfrac{1}{2}(n+1)\,(2n+1)\,(3n+2)\,v_n - (n+1)\,(9n+7)\,v_n^2 + 12(n+1)\,v_n^3 - 6v_n^4, \text{ etc.,}$$

or

$$\left.\begin{aligned} \kappa_1 &= \tfrac{1}{2}n + \tfrac{1}{6}\log_e n + 0{\cdot}89913 + O(n^{-1}\log n),\\ \kappa_2 &= \tfrac{1}{4}n(n+2) + \tfrac{1}{6}n\log_e n + O(\log n),\\ \kappa_3 &= \tfrac{1}{4}n^3 + O(n^2),\\ \kappa_4 &= \tfrac{3}{8}n^4 + O(n^3\log n), \end{aligned}\right\}$$

whence

$$\gamma_1 = 2 - n^{-1}\log_e n + O(n^{-1}), \quad \gamma_2 = 6 + O(n^{-1}\log n).$$

Thus as n tends to infinity the mean and standard deviation are closely represented by $\frac{1}{2}n$, while γ_1 and γ_2 approximate to 2 and 6 respectively. In fact, the distribution approximates to a discontinuous Pearsonian distribution of type III, and as all the reduced cumulants γ_r approximate to finite values, this comparison can be made without reservations. Further, the shape parameters γ_1 and γ_2 approximate to those of the χ^2 distribution for two degrees of freedom, though of course the location and scale parameters are different.

A simple example shows how rapidly the truncated Fisher distribution approximates to the Pearsonian form. If $n = 10$, $v_n = 6{\cdot}323$. Thus ten generations after a group of recessive mutations at different loci, 84·2 % of the mutant genes have disappeared by random extinction. But the number of recessive genes in the population is unaltered; and in the remaining 15·8 % of cases, the mean number of heterozygotes is 6·323, and its standard duration 5·44, while $\gamma_1 = 1{\cdot}86$, $\gamma_2 = 5{\cdot}07$. Thus the distribution is already fairly close to the form to which it approximates when n tends to infinity, with $\gamma_1 = 2$, $\gamma_2 = 6$.

In the more general case, Haldane (1939, equation (3)) showed that if μ_2 and μ_3 are the second and third moments about the mean of the fertility distribution whose generating function is $f(t)$, then

$$v_n = \tfrac{1}{2}\mu_2 n + \left(\tfrac{1}{2}\mu_2 - \frac{\mu_3}{3\mu_2}\right)\log n + c + O(n^{-1}\log n).$$

Thus if μ_2 is somewhat larger than unity, as seems to be the case with man, we should have, since $f''(1) = \mu_2$,

$$\kappa_1 = v_n = \tfrac{1}{2}\mu_2 n + \left(\tfrac{1}{2}\mu_2 - \frac{\mu_3}{3\mu_2}\right)\log n + C,$$
$$\kappa_2 = v_n(n\mu_2 + 1 - v_n) = \tfrac{1}{4}\mu_2^2 n^2 + O(n), \text{ etc.,}$$

in place of equations (4).

In an increasing population $f'(1)$ exceeds unity. If successive generations increase by a factor $1+k$, then $1 - f^n(0)$ does not tend to zero, but to $2k - \frac{8}{3}k^2 + \frac{28}{9}k^3 + \ldots$.

The first cumulants of the Fisher distribution are

$$\kappa_1 = (1+k)^n, \quad \kappa_2 = k^{-1}(1+k)\,[(1+k)^n - 1],$$

and so on, if $f(t) = e^{k(t-1)}$.

In all cases which are at all likely to occur the distribution of the probabilities of x after n generations consists of a large value for $x = 0$, and a somewhat positively skew distribution over a series of non-zero values. In view of the uncertainty of data on mutation rates, equations (3) will be sufficiently accurate for our purpose.

The appearance of recessives

Suppose that there were x_n heterozygotes **Aa** in the nth generation, we have next to calculate the expected number of recessives. Let α be the mean coefficient of inbreeding, that is to say if q be the frequency of **a** gametes, let the probability that a given **a** gamete unites with another be $\alpha+(1-\alpha)q$. Then Bernstein (1930) showed that α is independent of q, and Haldane & Moshinsky (1939) showed that α is of the order of 10^{-3} for several European populations. Let N be the number of the population.

Then if there are x heterozygotes the expected number of recessives in the next generation of a population of hermaphrodites would be $\frac{1}{2}\alpha x+\frac{(1-\alpha)x^2}{4N}$. However, the human species is bisexual. Assuming a sex ratio of equality, the probability that the x zygotes will consist of y females and $x-y$ males is $\binom{x}{y}2^{-x}$. The term $\frac{1}{2}\alpha x$ must therefore be multiplied by a correction factor of

$$\sum_{y=0}^{x}\frac{4y(x-y)}{x^2}\binom{x}{y}2^{-x}, \quad \text{or} \quad \frac{x-1}{x},$$

so the expected number is $\frac{1}{2}\alpha(x-1)+\frac{(1-\alpha)x^2}{4N}$, provided x exceeds zero. The correction to the term involving x^2 is negligible, since it only becomes relevant when x is of the order of $N^{\frac{1}{2}}$. Thus the expected number of recessives in the $(n+1)$th generation is

$$\begin{aligned} r_{n+1} = \Sigma\left[\frac{\alpha}{2v_n}(x'-1)+\frac{1-\alpha}{4N}x^2\right]p_{n,x} &= \frac{\alpha(v_n-1)}{2v_n}+\frac{(1-\alpha)}{4N}(n+1) \\ &= \frac{\alpha}{2}\left(1-\frac{2}{n+\frac{1}{3}\log n+c}\right)+\frac{(1-\alpha)(n+1)}{4N} \\ &= \frac{\alpha}{2}+\frac{n}{4N} \text{ approximately.} \end{aligned} \tag{5}$$

The second term is negligible, provided N is of the order of 10^6, for before n reaches a value of the order of 1000, we shall have to correct for the effect of selection.

If the homozygotes have a fitness $1-k$, the number of recessive genes will be reduced by a factor $1-\alpha k$ or approximately $e^{-k\alpha}$ in each generation. Thus in n generations they would be reduced to $e^{-k\alpha n}$. For serious conditions $k>\frac{1}{2}$, so although the process will be slightly speeded up when n reaches the order $N^{\frac{1}{2}}$, i.e. after some thousands of generations, the rate of appearance of homozygotes per mutation will be at first very nearly $\frac{1}{2}\alpha$, later approximating to $\frac{1}{2}\alpha e^{-k\alpha n}$.

However, the process does not begin at once. Apart from the correction for bisexuality, inbreeding does not begin immediately. Incest is negligible, first-cousin marriages account for about 72 % of the total value of α according to Haldane & Moshinsky, marriages of first cousins once removed for another 8 %, and so on. Thus r_n will not effectively equal $\frac{1}{2}\alpha$ for about ten generations, but will reach $0{\cdot}2\alpha$ after four generations.

So far we have only considered single mutations at different loci. Two other cases must be considered. In the first place several mutations may occur simultaneously at the same locus. Secondly, one or more mutations may occur giving genes which are already present in the population. Doubtless both events occurred among the survivors of atomic bomb explosions.

The Fisher distribution with several initial mutations

Suppose that the same mutation has occurred in one generation in g different zygotes, so that $\phi(0,t) = t^g$. Then the moment-generating function for the nth generation becomes

$$\begin{aligned} M_n(\theta) &= \phi[0, f^n(e^\theta)] \\ &= [f^n(e^\theta)]^g, \\ K_n(\theta) &= g \log [f^n(e^\theta)] \\ &= g[\{M_{n-1}(\theta)\} g^{-1} - 1] \\ &= g[\exp\{g^{-1} K_{n-1}(\theta)\} - 1]. \end{aligned}$$

To calculate the cumulants of one generation from those of the preceding one we have therefore to multiply each product of r cumulants by g^{1-r}. Thus $\kappa_1 = g$ and equations (2) become

$$\begin{aligned} \Delta\kappa_2 &= g^{-1}\kappa_1^2, \\ \Delta\kappa_3 &= 3g^{-1}\kappa_2 + g^{-2}\kappa^3, \\ \Delta\kappa_4 &= g^{-1}(4\kappa_3\kappa_1 + 3k_2^2) + 6g^{-2}\kappa_1^2\kappa_2 + g^{-3}\kappa_1^4, \\ \Delta\kappa_5 &= 5g^{-1}(\kappa_4\kappa_1 + 2\kappa_3\kappa_2) + 5g^{-2}(2\kappa_3\kappa_1^2 + 3\kappa_2^2\kappa_1) + 10g^{-3}\kappa_2\kappa_1^3 + g^{-4}\kappa_1^5, \\ \Delta\kappa_6 &= g^{-1}(6\kappa_5\kappa_1 + 15\kappa_4\kappa_2 + 10\kappa_3^2) + 15g^{-2}(\kappa_4\kappa_1^2 + 4\kappa_3\kappa_2\kappa_1 + \kappa_2^3) + 5g^{-3}(4\kappa_3\kappa_1^3 + 9\kappa_2^2\kappa_1^2) \\ &\quad + 15g^{-4}\kappa_2\kappa_1^4 + g^{-5}\kappa_1^6. \end{aligned}$$

Hence, in the nth generation,

$$\left.\begin{aligned} \kappa_1 &= g, \\ \kappa_2 &= ng, \\ \kappa_3 &= ng + \tfrac{3}{2}n(n-1), \\ \kappa_4 &= \tfrac{1}{2}ng(2n^2 + 7n - 7) + 2n(n-1)(n-2). \end{aligned}\right\} \tag{6}$$

The moments about zero are

$$\begin{aligned} \bar{x} &= g, \\ \overline{x^2} &= g^2 + ng, \\ \overline{x^3} &= g^3 + 3ng^2 + ng + \tfrac{3}{2}n(n+1), \\ \overline{x^4} &= g^4 + 6ng^3 + n(3n+4)\,g^2 + \tfrac{1}{2}n(2n^2 + 19n - 19)\,g + 2n(n-1)(n-2), \text{ etc.} \end{aligned}$$

Now the probability that x will be zero in the nth generation is $[f^n(0)]^g = \left(1 - \frac{1}{v_n}\right)^g$. Thus the moments of x' about zero are those of x divided by $1 - \left(1 - \frac{1}{v^n}\right)^g$. If $w_n = \frac{v_n^g}{v_n^g - (v_n - 1)^g}$, then $w_n = \frac{n}{2g} + O(\log n)$, and we have for the cumulants of the distribution of x',

$$\begin{aligned} \kappa_1 &= \tfrac{1}{2}(n+g) + O(\log n), \\ \kappa_2 &= \tfrac{1}{2}n(\tfrac{1}{2}n + g) + \ldots, \\ \kappa_3 &= \tfrac{1}{4}n^3\left(\frac{3}{g} - 2\right) + \ldots, \\ \kappa_4 &= n^4\left(\frac{7}{8} - \frac{n}{g}\right) + \ldots, \end{aligned}$$

The expressions are complicated, but κ_3 is clearly negative for large n:

$$r_{n+1} = \alpha g \frac{(w_n - 1)}{2w_n} + \frac{(1-\alpha)}{N} g(n+g)$$

$$= \frac{\alpha g}{2}\left[1 - \frac{2g}{n} + O(n^{-2})\right] + \frac{g(n+g)}{N} \text{ approximately.}$$

Thus the frequency is, as before, equal to $\frac{1}{2}\alpha$ multiplied by the number of mutant genes over a *long* period.

The addition of mutants to previously existing recessives

Suppose that the gene frequency of a recessive in a population is q, then the number of recessives expected per generation is $[\alpha q + (1-\alpha)q^2]N$. If q is increased to $q + \delta q$, this number is increased by $[\alpha + 2(1-\alpha)q]\,\delta q N$, or approximately $(\alpha + 2q)\,\delta q N$. Now a single mutant gene increases q to $q + 1/2N$ so $\delta q = 1/2N$ and $r_{n+1} = \frac{1}{2}\alpha + q$. It follows that if q is of the order of magnitude of α the dysgenic effect begins immediately, though it rises for a few generations. It is also appreciably larger, that is to say, it is over sooner. This argument depends on the assumption that qN is of such magnitude that it does not fluctuate appreciably as the result of the sampling process by which one generation is derived from another.

Discussion

To sum up, the conclusion is that the rate at which recessive zygotes are to be expected will rise in 200 years or less to a level of $\frac{1}{2}\alpha$ per generation for each mutation produced in a gamete which gives rise to a viable zygote. This will be somewhat exceeded if the recessive gene in question is already frequent in the population. In the long run each harmful recessive mutation will involve the elimination of half a zygote on the average. This elimination may occur in different ways. If the gene is lethal at a very early stage, like yellow in mice, the only effect will be an occasional missed period, and perhaps a lowering of fertility. If it is lethal at later stages it will cause abortions, stillbirths, and infantile or juvenile mortality. Or it may lower the fitness drastically without being completely lethal. Such genes are perhaps the most serious, as the elimination of a single pair involves the production of more than one person destined to chronic invalidism from a condition such as one of the recessive nervous diseases. Finally, it may slightly lower the fitness of a large number of individuals, for example reducing the mean fitness of ten people by 5 % on an average. If an expanding population is irradiated, more than half a life must be sacrificed per harmful mutation, if a diminishing population, less. It must be emphasized that this applies not merely to lethal and sublethal mutations, but to all mutations which appreciably lower the fitness of homozygotes, and are therefore ultimately eliminated.

The rate depends on the coefficient of inbreeding α. The value of this quantity increases with every generation of ancestors considered, and may approach unity if we reckon back to palaeolithic times (Haldane & Moshinsky, 1939). How it will behave in the next ten thousand years depends on unpredictable features of social structure. It may be that 0·001 is a serious underestimate of its effective value. If so the rate of production of homozygotes may rise very gradually for some thousand years.

We have now to consider the probable rate of production of new mutations. Consider an atom bomb explosion in which some 50,000 persons are killed at once or die within a few weeks, while another million are exposed to appreciable doses of γ-radiation. The lethal dose is probably of the

order of 500 roentgens, so we may suppose that the million survivors have received a mean dose of between 10 and 50 r., say 20. Those who receive just sublethal doses may be permanently sterilized. Mutations will occur in gametes and in spermatogonia and oogonia. On an average, in a stationary population each individual will produce two gametes which give rise to zygotes. So we have to consider 2×10^6 gametes derived from cells which have received a mean dose of 20 r.

Now the probability that a particular gene will mutate with a dose of 1 r. varies from about $0{\cdot}5 \times 10^{-8}$ to 16×10^{-8} in *Drosophila*, and (if the phenomenon is comparable) in bacteria and viruses (Lea, 1946, p. 144). The median value is about 2×10^{-8}. Thus with a dose of 4×10^7 r. we should expect the same locus to be affected twice in a minority of cases. Human genes cannot be much more sensitive to radiation than those of *Drosophila*, otherwise the rate of spontaneous mutation in man would probably be higher than is the case.

A dose of 1000 r. of X- or γ-rays produces recessive lethal mutations in about 3 % or rather less of the spermatozoa of irradiated *D. melanogaster*. And the number is roughly proportional to the dose (Lea, 1946, p. 151). However, about one-third of these are associated with gross structural changes in chromosomes, and these increase more than proportionally to the dose. So probably a dose of 20 r. would induce lethals in about 4×10^{-4} rather than 6×10^{-4} of the *X*-chromosomes. However, the total number of genes at risk is about five times the number of sex-linked genes. So about 2×10^{-3} of the irradiated gametes would carry a recessive lethal of some kind. Thus the number of recessive lethal mutations expected in 2×10^6 would be about 4000. This number may well have to be increased by a factor of 10 or 50 if man has more genes per nucleus than *Drosophila*, or if on account of their larger size or some other reason they are more sensitive to radiation. 4000 mutations would involve 2000 deaths in all, spread over very many generations and occurring at a rate of the order of two per generation.

This may be a serious underestimate. No doubt an atomic bomb exploding over a crowded area of a large city might cause considerably more mutations than the number calculated above. But it would also cause more immediate deaths. The relevant quantity is twice the dose in roentgens summed over the number of individuals irradiated but surviving, and weighted downwards in so far as they are wholly or partially sterilized. This has been taken as 4×10^7. It might be ten times as great. But even this would only give about twenty deaths per generation; and many of these might be in early embryonic life, and therefore negligible either from a humanitarian point of view or from the point of view of mere population size.

The effect of dominant and semi-dominant mutations will certainly be more spectacular, as these will all appear and mostly be eliminated in a few generations. If it is possible to discover their approximate frequency from observations on Japanese populations this may lead to a revision of the estimates given above. It should, however, be remembered that many dominants and semi-dominants are due to structural changes, and that these increase approximately with the square of the dose, and are not likely to be very numerous with doses below 100 r.

Muller (1941) has made calculations similar to the above, and which are in substantial agreement with them. However, his work has to the best of my knowledge only appeared in a summary, which is perhaps liable to misinterpretation. He writes of a 'latent period' of 5000 years for the inbreeding effect. If this denotes the mean time before a recessive zygote appears, it is in substantial agreement with my own calculation. However, the frequency of appearance of recessives should approach its maximum in a century or two, so the term 'latent period' may be rather misleading.

It is realized that the calculation of the expected mutation frequency may require drastic revision when, on the one hand, more facts are available concerning atomic bomb explosions and, on the other, the results of at present secret experiments conducted on mice in the U.S.A. are published. There is, however, a very simple reason for comparative optimism regarding the genetic effect of atomic bombs, namely, that the dose of X- or γ-rays which is lethal when given instantaneously is probably not much more than tenfold, and almost certainly not more than a hundredfold, the dose of natural radiation received during a normal lifetime. The situation in *Drosophila* is wholly different. Mitosis is not essential to the life of an imago. An insect imago can therefore survive about twenty times the dose of X- or γ-rays which would kill a man if absorbed in a short period. On the other hand, the mean age of a man at reproduction is about 500 times that of *D. melanogaster*. Thus the ratio of the maximum tolerated dose of radiation to that normally received is about 10,000 times greater in *Drosophila* than in man. Any attempt to argue from one species to the other, which does not include an allowance for this very great difference, is therefore likely to be misleading.

SUMMARY

Expressions are found for the moments and cumulants of the distribution of the probability that a neutral mutant gene originally present in one individual will have x descendants after n generations. Assuming a Poisson distribution for the number of offspring of an individual, it is found that when n is large the gene has disappeared in all but about $2/n$ cases, but in the remainder the numbers of genes give a type III distribution.

The number of recessive homozygotes expected per generation from a single induced recessive mutation is half the mean coefficient of inbreeding of the population concerned. The total number of deaths from recessive mutations to be expected as the result of an atomic bomb explosion is a small fraction of the number immediately killed, and is spread out over thousands of generations.

REFERENCES

F. BERNSTEIN (1930). Fortgesetzte Untersuchungen aus des Theorie der Blutgruppen. *Z. indukt. Abstamm.- u. VererbLehre*, **56**, 233.

R. A. FISHER (1930). The distribution of gene ratios for rare mutations. *Proc. Roy. Soc. Edinb.* **50**, 205.

J. B. S. HALDANE (1939). The equilibrium between mutation and random extinction. *Ann. Eugen., Lond.*, **9**, 400.

J. B. S. HALDANE & P. MOSHINSKY (1939). Inbreeding in Mendelian populations with special reference to human cousin marriage. *Ann. Eugen., Lond.*, **9**, 321.

M. G. KENDALL (1943). *The Advanced Theory of Statistics*. London.

D. E. LEA (1946). *Actions of Radiations on Living Cells*. Cambridge.

H. J. MULLER (1941). The role played by radiative mutation in mankind. *Science*, **93**, 438.