

MULTI-STAGE SELECTION FOR GENETIC GAIN

S. S. Y. YOUNG

Division of Animal Genetics, C.S.I.R.O., McMaster Laboratory, Glebe, N.S.W.

Received 29.ix.63

1. INTRODUCTION

IN animal (or plant) selection for more than one character, the use of a single stage selection index (Fairfield Smith, 1936; Hazel, 1943) is known never to be less efficient than the method of independent culling levels (Hazel and Lush, 1942; Young, 1961). However, it is sometimes no more efficient, and in cases where the traits under consideration become evident at different ages, independent culling levels offer the practical advantage of being able to dispose of a proportion of inferior animals earlier. This possibility is of particular importance in male selection, since the final proportion of entire males retained is usually small and early disposal of surplus animals may offer considerable advantage to the breeder.

However, the theory of independent culling levels at present covers only the case when selection is for one character at a time. In practice it is possible that one or more traits may be available at each of several times, and selection could take place sequentially. Such a scheme might be called "multi-stage selection" and methods of applying it are investigated in this paper. In particular, two methods are considered. The first is an extension of the independent culling approach to cover the case when culling at each stage is done according to a number of characters. In the second method part and whole selection indices are used in selection.

2. THEORY

(i) *Single stage selection*

The theory of the selection index will be outlined very briefly below to establish notations used in the subsequent discussion.

Let the phenotypic value of a trait X_i ($i = 1, \dots, n$) be represented by $(G_i + E_i)$, where G_i is the additive genetic effect and E_i the non-additive genetic effect plus the environmental contribution. If the economic weight of X_i be a_i , the total genetic value of an individual may be defined as

$$H = \sum a_i G_i = \sum b_i X_i + e, \quad (1)$$

where the b_i s are the regression coefficients to be calculated and e the residual. The coefficients of the selection index

$$I = \sum_{i=1}^n b_i X_i \quad (2)$$

are calculated by minimising the sum of squares of e s and may be

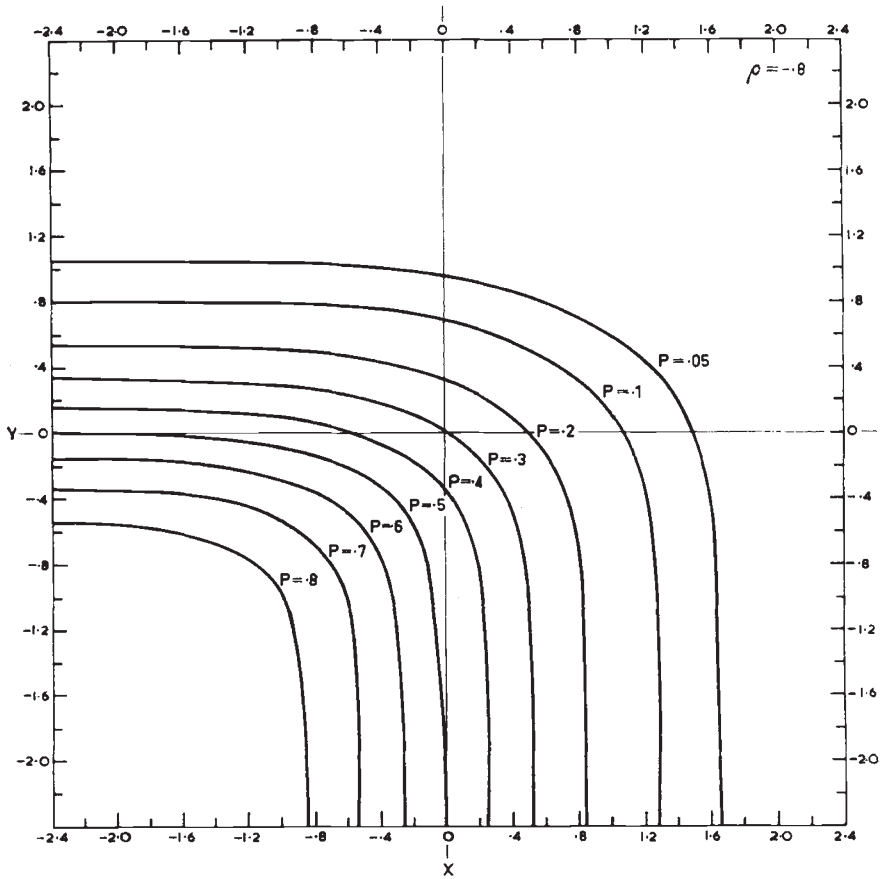


FIG. 1. Culling points when $\rho = -0.8$. Each P curve shows various combinations of truncation points X and Y to give a fixed proportion of animals saved.

considered as the best linear predictor of H. By assuming the independence of G_i and E_i , the vector of bs, \mathbf{b} is given by

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{a} \tag{3}$$

where \mathbf{P} and \mathbf{G} are the phenotypic and genetic variance and covariance matrices respectively and \mathbf{a} the vector of economic weights.

(ii) Multi-character independent culling levels

(a) *The general case.* If we have n characters X_1, X_2, \dots, X_n under consideration, and we wish to select for n_1 characters at the first stage and n_j characters at the j th stage $\left(\sum_{j=1}^s n_j = n\right)$, where the number in each stage is not necessarily equal, we have a situation of multi-stage selection. Since selection is to be carried out in s stages the total index can be written as

$$I = \sum_{i=1}^{k_1} b_i X_i^1 + \sum_{i=k_1+1}^{k_2} b_i X_i^2 + \dots + \sum_{i=k_{s-1}+1}^{k_s} b_i X_i^s = \sum_{j=1}^s I_j. \tag{4}$$

The superscript of X_i^j indicates that the trait X_i is measured at the j th age, so that in each sub-index, I_j , all the n_j traits are measured at the same age. Also, in (4) X_i^j and $X_i^{j'}$ can be a single character measured at different times. Following the least squares procedure mentioned earlier and writing the normal equations in terms of covariance, we have

$$b_1 \text{Cov}(X_i^j X_i^{j'}) + b_2 \text{Cov}(X_i^j X_i^{j''}) + \dots + b_n \text{Cov}(X_i^j X_i^n) = \text{Cov}(X_i^j H). \quad (5)$$

If we write $X_i^j = G_i^j + e_i^j$, where G_i^j is the additive genetic effect of X_i at the j th age then

$$\begin{aligned} \text{Cov}(X_i^j X_i^{j'}) &= P_{ij}^{jj'}, \text{ and since} \\ \text{Cov}(X_i^j G_i^{j'}) &= G_{ij}^{jj'}, \end{aligned} \quad (6)$$

we have

$$\text{Cov}(X_i^j H) = \sum_{j'=1}^n a_{j'} G_{ij}^{jj'}.$$

The phenotypic and genetic variance and covariance matrices (\mathbf{P} and \mathbf{G}) in (5) can be written down according to the number of stages

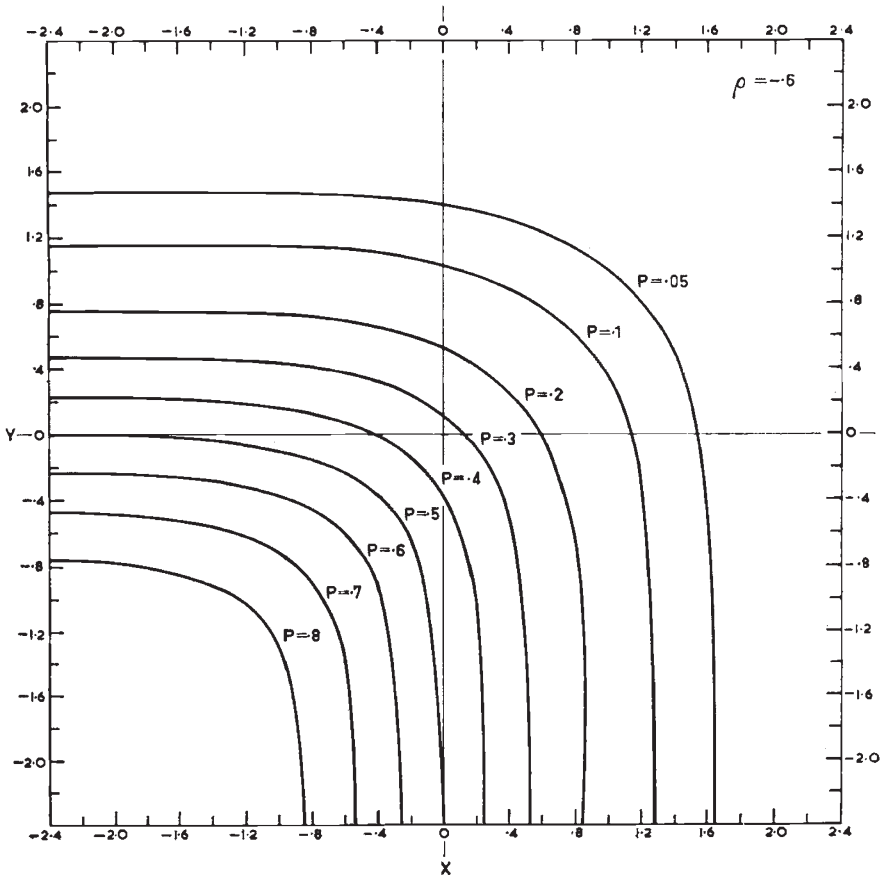


FIG. 2.—Culling points when $\rho = -0.6$.

involved in the selection scheme. As an example, if we wish to select for 5 different traits in 3 stages (at ages 1, 2 and 3 respectively) in which $n_1 = 2, n_2 = 2$ and $n_3 = 1$, with the aim of improving the breeding values at age 3, and if we further assume that the G_s at ages other than age 3 have no economic weights we have,

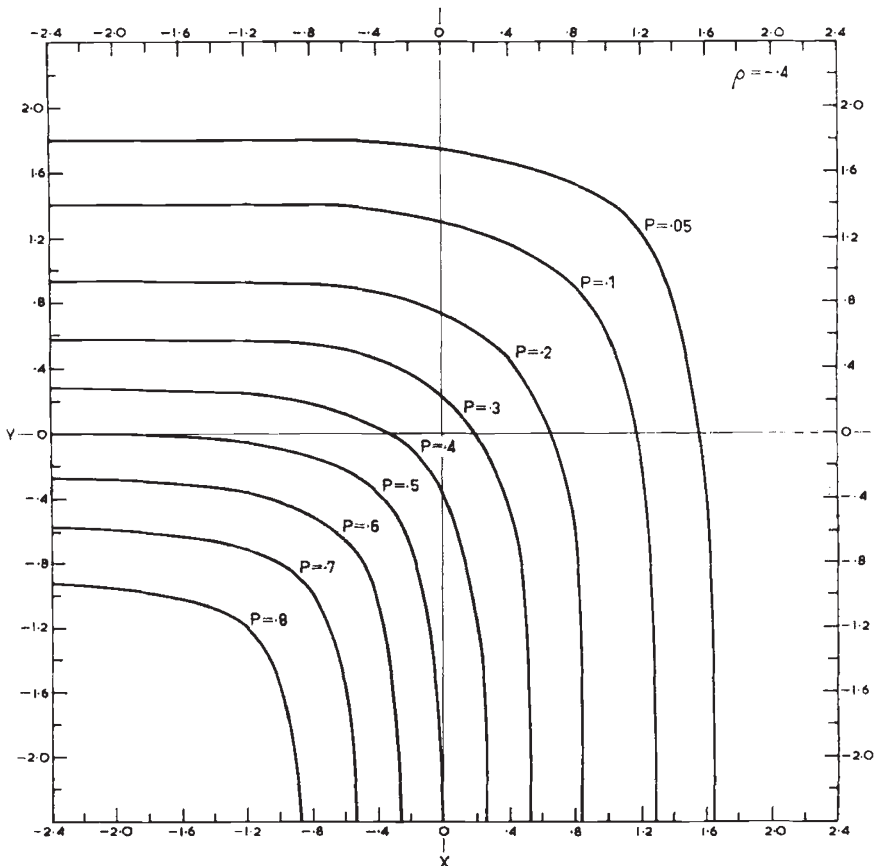
$$P = \begin{bmatrix} P_{11}^{11} & P_{12}^{11} & P_{13}^{12} & P_{14}^{12} & P_{15}^{13} \\ P_{21}^{11} & P_{22}^{11} & P_{23}^{12} & P_{24}^{12} & P_{25}^{13} \\ \hline P_{31}^{21} & P_{32}^{21} & P_{33}^{22} & P_{34}^{22} & P_{35}^{23} \\ P_{41}^{21} & P_{42}^{21} & P_{43}^{22} & P_{44}^{22} & P_{45}^{23} \\ \hline P_{51}^{31} & P_{52}^{31} & P_{53}^{32} & P_{54}^{32} & P_{55}^{33} \end{bmatrix}$$


FIG. 3.—Culling points when $\rho = -0.4$.

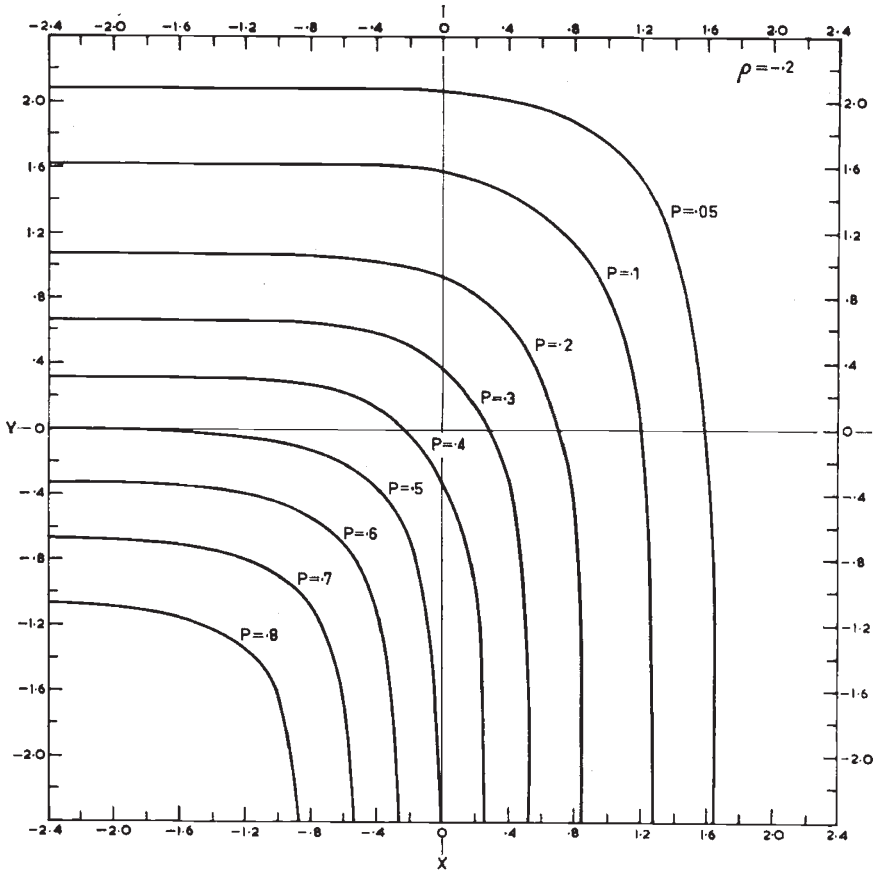


FIG. 4.—Culling points when $\rho = -0.2$.

In each of the diagonal sub-matrices the P_{ii} are phenotypic variances and covariances in animals of the same age, while each of the off-diagonal sub-matrices involves covariances of two traits at different ages. Similarly the matrix G is

$$G = \begin{bmatrix} G_{11}^{13} & G_{12}^{13} & G_{13}^{13} & G_{14}^{13} & G_{15}^{13} \\ G_{21}^{13} & G_{22}^{13} & G_{23}^{13} & G_{24}^{13} & G_{25}^{13} \\ G_{31}^{23} & G_{32}^{23} & G_{33}^{23} & G_{34}^{23} & G_{35}^{23} \\ G_{41}^{23} & G_{42}^{23} & G_{43}^{23} & G_{44}^{23} & G_{45}^{23} \\ G_{51}^{33} & G_{52}^{33} & G_{53}^{33} & G_{54}^{33} & G_{55}^{33} \end{bmatrix}$$

Having constructed the matrices \mathbf{P} and \mathbf{G} with suitable variances and covariances, it is clear that the vector \mathbf{b} can be calculated as in (3), and we have the total index

$$I = \sum_{j=1}^s I_j.$$

The total expected genetic gain is

$$\Delta H = E(H - \mu_H) = \beta(I - \mu_I) = (I - \mu_I) \tag{7}$$

where E denotes the expected value, μ the population mean and β the regression coefficient of H on I . The last quality is well known since

$$\text{Cov}(HI) = \mathbf{b}^T \mathbf{G} \mathbf{a} = \mathbf{b}^T \mathbf{P} \mathbf{P}^{-1} \mathbf{G} \mathbf{a} = \mathbf{b}^T \mathbf{P} \mathbf{b} = \text{Var}(I).$$

After all stages of selection the selected animals will have an average index score of

$$I = \sum_{j=1}^s I_j. \tag{8}$$

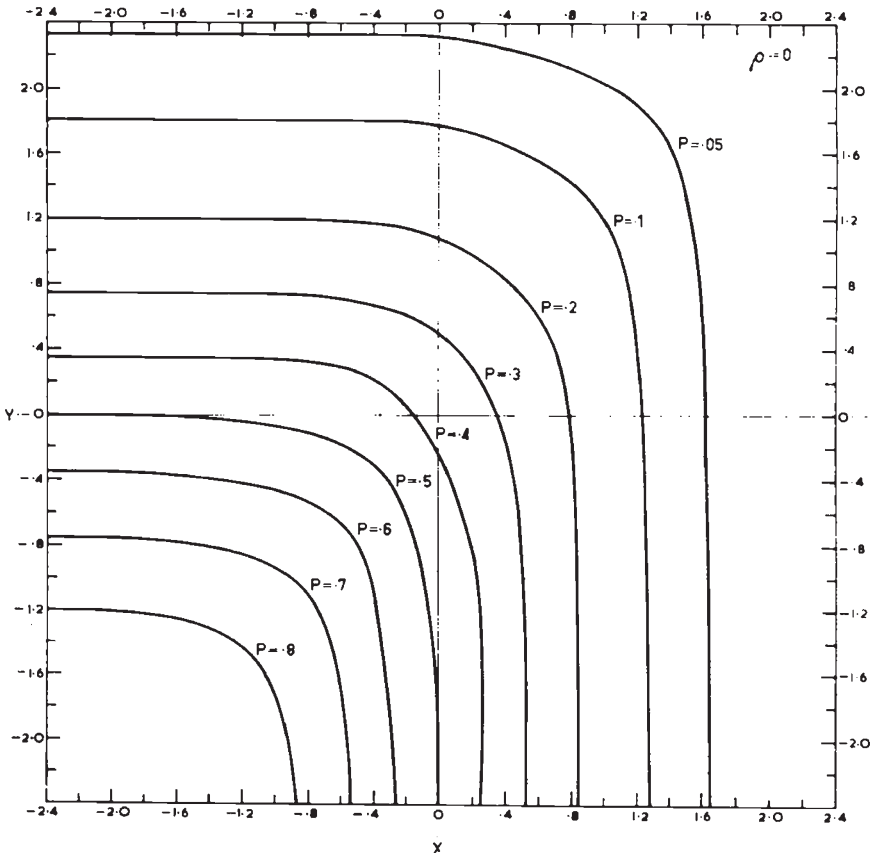


FIG. 5.—Culling points when $\rho = 0$.

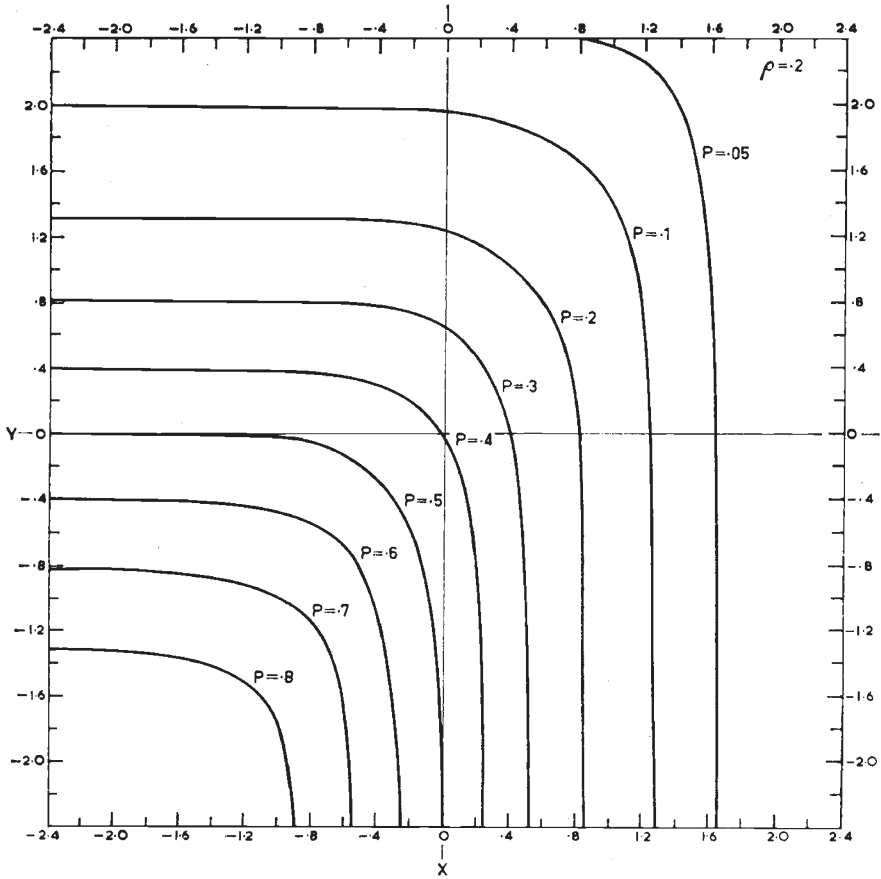


FIG. 6.—Culling points when $\rho = 0.2$.

We can write the average expected genetic gain as

$$\begin{aligned} \Delta \bar{H} &= \bar{I} - \mu_I = \sum_{j=1}^s (I_j - \mu_{I_j}) \\ &= \sum_{j=1}^s \sigma_{I_j} i_j, \end{aligned} \tag{9}$$

where σ_{I_j} denotes the standard deviation of the j th sub index and i_j the selection differential resulting from all s stages of selection with respect to the j th index, in standard deviation units. The variances and covariances of the sub-indices are given by

$$\begin{aligned} \text{Var}(I_j) &= \sum_{i=m=k_{j-1}+1}^{k_j} \sum_{m=k_{j-1}+1}^{k_j} b_i b_m P_{im}^{jj}, \\ \text{Cov}(I_j I_{j'}) &= \sum_{i=k_{j-1}+1}^{k_j} \sum_{m=k_{j'-1}+1}^{k_{j'}} b_i b_m P_{im}^{jj'}, \end{aligned}$$

and these values are useful in the calculation of the correlation coefficients between sub-indices. For a given proportion (p) of animals retained, many culling plans, with different culling points for each stage, will all give p , only one plan, however, gives the maximum genetic gain. If we regard the σ_{I_j} s in (9) as the weights for the various i_j s, $\Delta\bar{H}$ can be maximised for a fixed p . The maximisation of $\Delta\bar{H}$ when the I_j s are independently normally distributed is the same as the

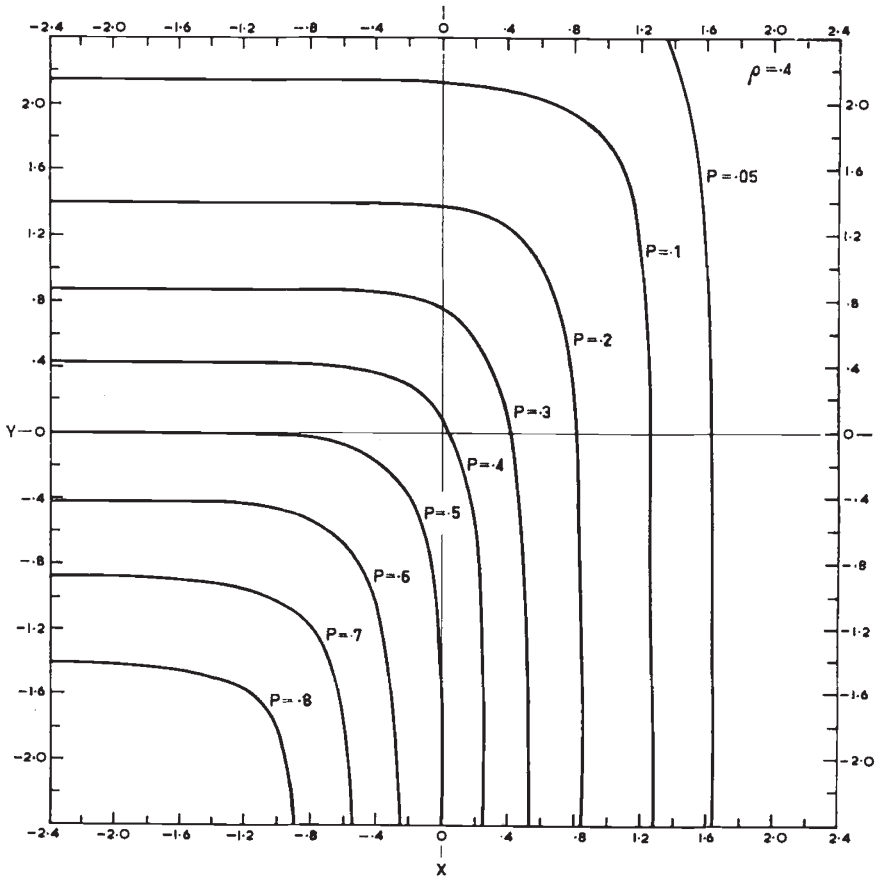


FIG. 7.—Culling points when $\rho = 0.4$.

method discussed by Hazel and Lush (1942). If there are only two stages of selection and if I_1 and I_2 are binormally distributed, the maximisation procedure has been discussed in some detail by Young and Weiler (1960). For three or more stages of selection the calculations of i_j for a given set of truncation points, in multinormal distribution were given by Birnbaum and Myer (1953) and Tallis (1961), while formulæ for the calculation of p for a set of truncation points were given by Plackett (1954) and the values of p in the trivariate case has been tabulated by Steck (1958). The overall computations in these cases, however, become very tedious indeed.

(b) *Single character case.* When only one character is selected at each stage, we have the case of independent culling levels.

In this case we have $\sigma_{I_j} = b_j \sqrt{P_{jj}}$, and after all stages of selection

$$\begin{aligned} \Delta \bar{H} &= \sum_{i=1}^s b_j \sqrt{P_{jj} i_j} \\ &= \sum W_j i_j \text{ (say)} \end{aligned} \tag{10}$$

where W_j is the weighting factor for i_j .

Considering the case when two characters, measured at the same age, are under selection, we have

$$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} \sqrt{P_{11}} & 0 \\ 0 & \sqrt{P_{22}} \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}^{-1} \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{11}$$

and

$$\begin{aligned} W_1 &= \sqrt{P_{11}} (P_{11} P_{22} - P_{12}^2)^{-1} \{ a_1 (P_{22} G_{11} - P_{12} G_{21}) + a_2 (P_{22} G_{12} - P_{12} G_{22}) \} \\ &= (1 - \rho^2)^{-1} \{ a_1 \sqrt{P_{11}} (h_1^2 - \rho h_{12}) + a_2 \sqrt{P_{22}} (h_{12} - \rho h_2^2) \}, \end{aligned} \tag{12}$$

$$\begin{aligned} W_2 &= \sqrt{P_{22}} (P_{11} P_{22} - P_{12}^2)^{-1} \{ a_1 (P_{11} G_{12} - P_{12} G_{11}) + a_2 (P_{11} G_{22} - P_{12} G_{21}) \} \\ &= (1 - \rho^2)^{-1} \{ a_1 \sqrt{P_{11}} (h_{12} - \rho h_1^2) + a_2 \sqrt{P_{22}} (h_2^2 - \rho h_{12}) \}, \end{aligned}$$

where $h_i^2 = \frac{G_{ii}}{P_{ii}}$, $h_{ij} = \frac{G_{ij}}{\sqrt{P_{ii} P_{jj}}}$ and $\rho = \frac{P_{12}}{\sqrt{P_{11} P_{22}}}$.

It is seen that the W_i s are the same as those derived by Young and Weiler (1961) from a different consideration.

(iii) *The use of part and whole indices—a two-stage case*

A more efficient way to carry out multi-stage selection is to select animals at stage 1 according to I_1 , and at stage 2 according to $(I_1 + I_2)$ and so on. The final stage is based on I the total index. The method is more efficient because more and more information about the animals is used in selection during subsequent stages. The computation involved in this scheme, however, is slightly more difficult than the previous method, where only one sub-index is used in each stage. It is worthwhile, however, to develop formulæ for a case of two stage situation.

Consider sub-indices I_1 and I_2 binormally distributed with standardised variates

$$u = \frac{I_1 - \mu_{I_1}}{\sigma_{I_1}}, \quad v = \frac{I_2 - \mu_{I_2}}{\sigma_{I_2}}$$

where μ_{I_1} and μ_{I_2} are the means of I_1 and I_2 . The method of selection is to retain animals with $u \geq X$ at stage 1 and $(u + v) \geq Y$ at stage 2.

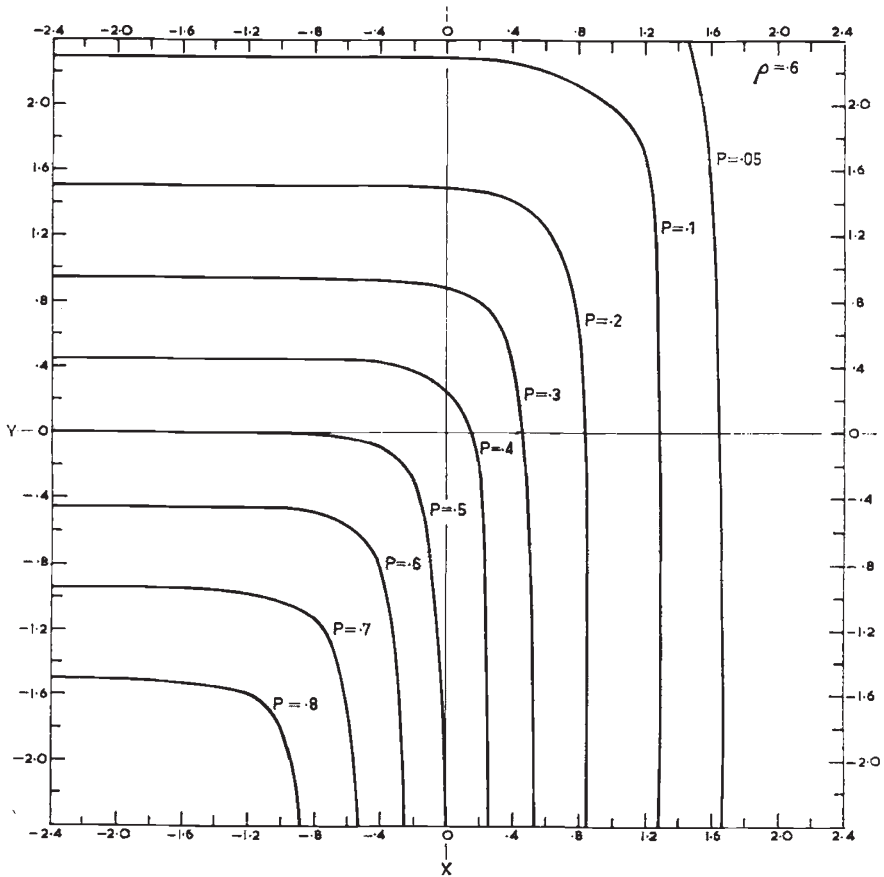


FIG. 8.—Culling points when $\rho = 0.6$.

The selected population then has the probability density $\phi(uv; \rho)/p$, where

$$p = p(X, Y-u; \rho) = \int_X^\infty \int_{Y-u}^\infty \lambda(2\pi)^{-1} \exp\{-\frac{1}{2}\lambda^2(u^2 - 2\rho uv + v^2)\} dv du \tag{13}$$

and $\rho =$ the correlation coefficient between u and v .

$$\lambda = (1 - \rho^2)^{-\frac{1}{2}}$$

To calculate the genetic gain after selection, we have to calculate first the mean of u , $E(u)$ and of v , $E(v)$, in the selected population. If we denote

$$\begin{aligned} S(u) &= \int_{Y-u}^\infty \exp\{-\frac{1}{2}\lambda^2(v^2 - 2\rho uv)\} dv \\ E(u) &= \lambda(2\pi\rho)^{-1} \int_X^\infty uS(u) \exp\{-\frac{1}{2}\lambda^2u^2\} du \end{aligned} \tag{14}$$

writing

$$dh = u \exp\{-\frac{1}{2}\lambda^2u^2\}, g = S(u), \frac{dg}{du} = \frac{d}{du} S(u),$$

(14) can be integrated by parts. We then have

$$\lambda^2 PE(u) = \phi(X)Q(A) + P\rho\lambda^2 E(v) + \phi K(YK)Q(B) \tag{15}$$

$$\lambda^2 PE(v) = \rho P\lambda^2 E(u) + K\phi(YK)Q(B) \tag{16}$$

where

$$K = \{2(1 + \rho)\}^{-\frac{1}{2}},$$

$$A = \lambda\{Y - X(1 + \rho)\},$$

$$B = (2X - Y)\{2(1 - \rho)\}^{-\frac{1}{2}},$$

$$\phi(t) = (2\pi)^{-\frac{1}{2}} \exp\{-\frac{1}{2}t^2\},$$

$$Q(t) = \int_t^\infty \phi(x)dx.$$

Solving (15) and (16) we have

$$E(u) = \phi(X)Q(A) + \sqrt{\frac{1+\rho}{2}} \phi(YK)Q(B)/\rho, \tag{17}$$

$$E(v) = \rho\phi(X)Q(A) + \sqrt{\frac{1+\rho}{2}} \phi(YK)Q(B)/\rho. \tag{18}$$

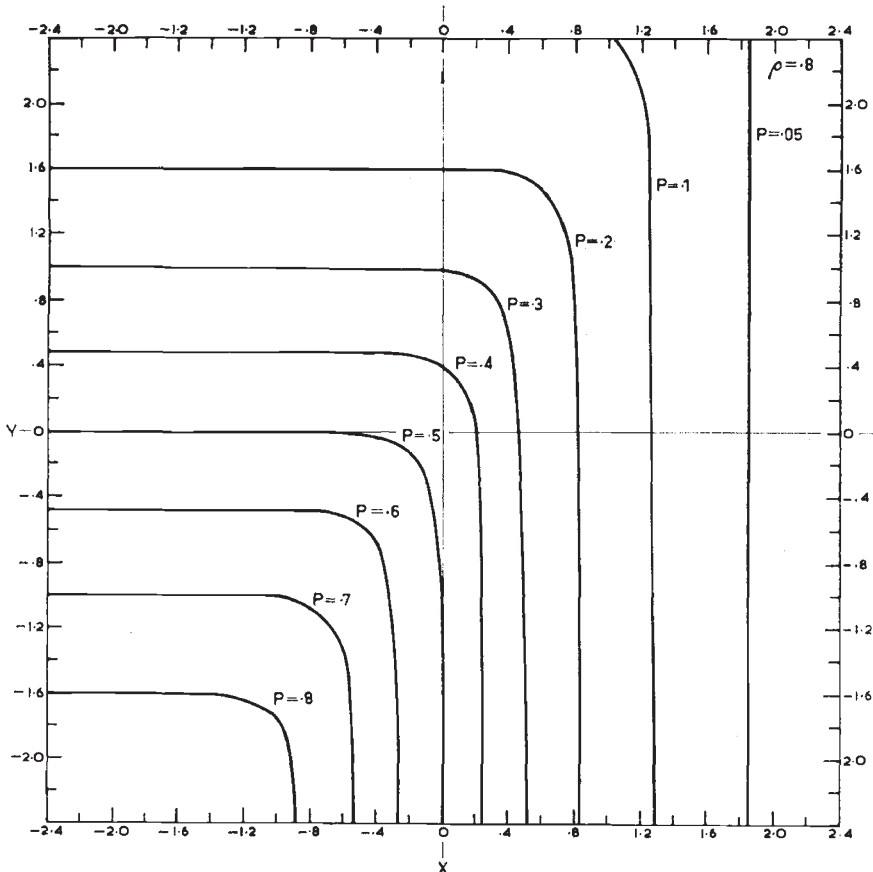


FIG. 9.—Culling points when $\rho = 0.8$.

The calculations of $E(u)$ and $E(v)$ thus involve the values of areas and ordinates of truncated normal distributions and p . Many tables for values of ϕ and Q are available but the writer is unaware of any tables for values of p when truncation is carried out in the present manner. For this reason a number of nomograms are presented in this paper (figs. 1-9). In each figure various combinations of (X, Y) for different values of p and a fixed ρ are plotted. The p values in each figure are selected to cover the range of values most likely to be useful to breeders. The 9 figures give sets of curves for $\rho = -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6$ and 0.8 respectively. The computational procedure of p for given values of X and Y are discussed by Mr G. Brown in the Appendix.

When $E(u)$ and $E(v)$ are known then the total expected genetic gain after selection is

$$\Delta\bar{H}' = \sigma_{I_1}E(u) + \sigma_{I_2}E(v). \quad (19)$$

Since the second stage, in this case, is a full index selection, the expected gain is greatest when no culling is done at stage one. However as genetic gain is often not the only consideration, formula (19) is useful in calculating the different amount of gain for a number of possible schemes involving early culling. These gains can then be considered in conjunction with other factors to derive an optimum scheme.

3. REMARKS

When greater accuracy is not required previous records of the animals may be used to calculate the approximate truncation points by the method of trial and error. In particular when, for husbandry reasons, fixed proportions of animals have to be discarded at various ages equations (9) and (19) can be used to determine which group of characters should be selected at various stages for greater genetic gain. The efficiency of the selection plan, as a result of this approximation, will not be at a maximum but may be quite useful as a guide in practice.

Of the two methods proposed here, the first method, the multi-character independent culling scheme, involves less computation but requires a maximisation procedure. The alternative one, the part and whole index selection, while requiring somewhat more calculation, has a greater efficiency. The latter has the added advantage that no maximisation process is required since, as far as genetic gain is concerned, the culling rate at stage I should be kept as low as possible.

The relative efficiency, in terms of genetic gain, of the two methods probably cannot be expressed in simple expressions and numerical calculation for each situation is required.

The methods discussed in the present paper can also be used when selection is for lifetime production. The sub-indices, in this case, can be calculated according to the formula developed by Young and Tallis (1961).

An alternative approach to the solution of the multi-character independent culling levels was pointed out to the writer by Dr G. M. Tallis. He suggested that $\Delta\bar{H}$ can be written as the mean of the truncated multinormal distribution, which in turn is a function of the unknown b_i s. The coefficients b_i can then be obtained by maximising $\Delta\bar{H}$ with respect to b_i and the numerical values of b_i can be obtained by iteration. The author was unable to establish the equivalence or otherwise of these two approaches.

4. SUMMARY

Two methods of multi-stage selection, when selection at each stage is for an arbitrary number of characters, are discussed. The first method is an extension of the independent culling levels method to more than one character at each stage and in the second method the use of part and whole selection indices in a two-stage selection scheme is considered.

Acknowledgments.—The use of additional information in subsequent stages was first suggested to me by Prof. J. B. S Haldane and later independently by Dr R. N. Curnow, who also read the manuscript and made helpful comments.

Thanks are also due to Mr G. Brown of this Division for checking equations (17) and (18) and for the calculation of p values.

5. REFERENCES

- BIRNBAUM, Z. W., AND MEYER, P. L. 1953. On the effect of truncation in some or all co-ordinates of a multi-normal population. *J. Indian Soc. Agric. Statist.*, 5, 17-28.
- HARRIS, F. 1957. *Numerical Integration Constants*. Mathematical tables and other aids to Computation, Vol. 11.
- HASTINGS, C. 1955. *Approximation for Digital Computers*. Princeton University Press.
- HAZEL, L. N. 1943. The genetic basis for constructing selection indexes. *Genetics*, 28, 476-490.
- HAZEL, L. N., AND LUSH, J. L. 1942. The efficiency of three methods of selection. *J. Hered.*, 33, 393-399.
- OWEN, D. B. 1957. *The bivariate normal probability distribution*. Sandia Corporation System Analysis Sc. 383 A(T.R.).
- PLACKETT, R. L. 1954. A reduction formula for normal multivariate integrals. *Biometrics*, 41, 351-360.
- SMITH, H. FAIRFIELD. 1936. A discriminant function for plant selection. *Annals of Eugenics*, 7, 240-252.
- STECK, G. P. 1958. A table for computing trivariate normal probabilities. *Ann. Math. Stat.*, 29, 780-800.
- TALLIS, G. M. 1961. The moment generating function of the truncated multi-normal distribution. *J. R. Statist. Soc. B.*, 23, 223-229.
- YOUNG, S. S. Y. 1961. A further examination of the relative efficiency of three methods of selection for genetic gains under less-restricted conditions. *Genet. Res.*, 2, 106-121.
- YOUNG, S. S. Y., AND TALLIS, G. M. 1961. Performance index for lifetime production. *J. Anim. Sci.*, 20, 506-509.
- YOUNG, S. S. Y., AND WEILER, H. 1960. Selection for two correlated traits by independent culling levels. *J. Genetics*, 57, 329-338.

APPENDIX

Evaluation of Truncated Bivariate Normal Volumes

G. BROWN

Consider the standardised bivariate normal distribution $\phi(0,0,1,1;\rho)$ truncated at

$$\begin{aligned}u &= X \\ u+v &= Y.\end{aligned}$$

We wish to evaluate the volume P, cut off by

$$\begin{aligned}u &\geq X \\ u+v &\geq Y.\end{aligned}$$

A transformation $m = u$, $n = u+v$ followed by a scaling transformation $r = m$, $s = n/\sqrt{2(1+\rho)}$ will reduce the problem to the more usual orthogonal truncation and

$$P = \int_{\mathbf{X}}^{\infty} \int_{\mathbf{Y}/\sqrt{2(1+\rho)}}^{\infty} \phi(0,0,1,1;\rho') dr ds$$

where

$$\rho' = \sqrt{\frac{1+\rho}{2}}$$

Thus the tables of Owen (1957) or the graphs of Young and Weiler (1960) could be used. Adopting Owen's notation we have

(a) $X \neq 0$, $Y \neq 0$

$$\begin{aligned}P = C - \frac{1}{2} \left[\Phi(X) + \Phi\left(\frac{Y}{\sqrt{2(1+\rho)}}\right) \right] - T \left[X, \frac{Y-(1+\rho)X}{\sqrt{1-\rho^2}} \right] \\ - T \left[\frac{Y}{\sqrt{2(1+\rho)}}, \frac{2X-Y}{Y} \sqrt{\frac{1+\rho}{1-\rho}} \right] \quad (i)\end{aligned}$$

where C is 1 or $\frac{1}{2}$ according to whether XY is greater or less than zero respectively.

(b) $X \neq 0$, $Y = 0$

$$P = \frac{1}{2} - \frac{1}{2} \Phi(X) + T \left(X, \sqrt{\frac{1+\rho}{1-\rho}} \right) \quad (ii)$$

(c) $X = 0$, $Y \neq 0$

$$P = \frac{1}{2} - \frac{1}{2} \Phi\left(\frac{Y}{\sqrt{2(1+\rho)}}\right) + T \left(\frac{Y}{\sqrt{2(1+\rho)}}, \sqrt{\frac{1+\rho}{1-\rho}} \right) \quad (iii)$$

(d) $X = Y = 0$

$$\begin{aligned}P &= \frac{1}{4} + T \left(0, \sqrt{\frac{1+\rho}{1-\rho}} \right) \\ &= \frac{1}{4} + \frac{\arctan \sqrt{\frac{1+\rho}{1-\rho}}}{2} \quad (iv)\end{aligned}$$

To evaluate even a single volume using the above formulæ is complex since it involves tedious arithmetic and interpolation from existing tables. Hence it seems worthwhile to evaluate P on a computer and present the results as graphs showing contours of equal volume for various parent correlations ρ . The user then may look up the volumes directly, given the two truncation values X and Y together with ρ . Alternatively, given P the values of X and Y may be chosen to comply with the user's wishes. The function

$$T(h,a) = \frac{1}{2\pi} \int_0^a \frac{\exp[-\frac{1}{2}h^2(1+x^2)]}{1+x^2} dx$$

was evaluated using an eight-point Gaussian quadrature technique Harris (1957) and the results agree with Owen's table (given to six decimal places).

An early attempt was made to evaluate the volume by one direct quadrature, namely

$$P = 1 - \Phi(x) - \int_a^\infty \phi(x) \Phi \left[\frac{Y - (1+\rho)x}{\sqrt{1-\rho^2}} \right] dx$$

but the method was inaccurate for various values of the parameters and was consequently discarded.

The above method was programmed for the IBM 7090 computer at the Weapons Research Establishment and tables for various ρ , $X[-2.4(0.2)2.4]$ and $Y[-2.4(0.2)2.4]$ were obtained from which the charts were made up.