AN APPLICATION OF THE DISCRIMINANT FUNCTION FOR SELECTION IN POULTRY

By V. G. PANSE, *Institute of Plant Industry, Indore, India*

INTRODUCTION

To the poultry keeper certain features related to the egg production of laying stocks are of direct economic significance. The annual number of eggs laid and their size determine the cash return from each bird. Other characters such as the size of the bird or its period of sexual maturity, that is, the period from its hatching to the date of laying the first egg, are concerned with the economy of feed consumption. Besides, the size of the bird would affect the price it may fetch when, at the end of its laying career, it is sold for meat. The breeder engaged in the improvement of egg production aims at incorporating such characters in his breed in so far as their genetic or physiological relationship would permit.

Productivity or the capacity of laying a large number of eggs is an inherited character, and at least some of the controlling genes are located in the sex chromosome, so that those factors are transmitted from the sire to the daughter (Pearl, 1912). Similarly, significant differences in the daughters' egg weight from different sires was demonstrated by Hays (1941), who also discusses additional evidence on this point and, in confirmation of his previous view (1929) that inheritance of egg weight is autosomal, concludes that both sires and dams must be selected for factors affecting egg size in order to secure a uniform progeny. Poultry breeders usually devote considerable care in the selection of suitable sires, and the evidence available shows that this selection is at least as important as that of dams, if not more, in breeding better laying stocks. The importance of having selected sires is increased by the fact that only a few cocks are maintained for mating in the breeding pens. The object of the present paper is to describe the application of the discriminant function (Fisher, 1936) for judging the genetic potentiality of the sire.

The discriminant function provides an extremely useful technique for selection when, as in the present case, selection is to be based on several characters simultaneously. The individual characters may show different degrees of variation due to genetic and environmental causes, and their true value in assessing the genotype will also depend upon the strength of their mutual correlations. In evaluating the discriminant function both these factors are taken into account and the relative weight to be attached to the observed values of different characters is so fixed that the resulting compound score is most highly correlated with the true genetic value of a variety or a line. The application of this technique to plant selection is shown by Fairfield Smith (1936). Its use in selection in poultry presents no essentially different features.

MATERIAL

The data analysed relate to a stock of Brown Leghorns maintained at the Institute of Animal Genetics, Edinburgh, and cover three seasons, 1935, 1936 and 1937, in which the pullets from different matings were hatched. Records for the number of eggs laid during five months from November to March, the average egg weight in March, the body weight of the pullet when further growth had ceased, that is, at an age of approximately 40 weeks, and the age at sexual maturity are available over the first pullet year for 218 pullets belonging to the progenies of seventeen sires. These four characters are of primary importance in determining the productivity of the stock and are analysed to obtain a discriminant function. The number of hens from individual sires which were allowed to complete their year's record was too small, and therefore egg production from November to March is used as a measure of the laying rate. The figures are comparable because hatching was so arranged that almost all pullets had started laying in the beginning of November. The egg weight in March, when all pullets were in full production, is considered to be more consistent and more suitable for comparison than the average for the whole year. The pullets were kept in standardized twenty-bird pens and all received the same diet and treatment. The pullets from individual sires used in any one year were not kept separately but were intermixed.

While the matings were not controlled, in the sense that each sire was not mated to the same set of dams, the dams all belonged to the same common stock, and no deliberate selection was exercised in arranging matings. Eighty-seven hens in all were involved, and hens belonging to different age groups were used in each of three seasons. Twenty-three of the hens were mated to different sires in more than one season. Variation within a sire's progeny would be increased by contributions from genetic differences between mothers, but differences between sires are not likely to be biased from this cause, since the matings were apparently random. The size of each progeny ranged between four and twenty-nine, the average being thirteen pullets approximately.

METHOD OF ANALYSIS AND RESULTS

Each season's data were analysed as between sires' progenies and within progenies, and the analyses for the three seasons pooled together. The results are shown in Tables 1 and 2.

Table 1. *Pooled analyses of variance for number of eggs, egg weight, body weight and period of maturity*

Due to	Degrees of freedom	Mean squares and variance ratios: No. of eggs, M.S.	No. of eggs, V.R.	Egg weight, M.S.	Egg weight, V.R.	Body weight, M.S.	Body weight, V.R.	Maturity, M.S.	Maturity, V.R.
Between sires' progenies	14	462·6	1·37	70·15	5·78†	23·94	1·31	2673	1·69*
Within progenies	201	337·3	–	12·14	– –	18·21		1582	

* Significant, v.r. 1·74, $P = 0·05$. † Significant, $P = 0·01$.

Table 2. *Pooled analyses of covariance between number of eggs, egg weight, body weight and period of maturity*

	Degrees of freedom	Mean products and correlation coefficients: Egg no. × egg weight, M.P.	Egg no. × egg weight, r	Egg no. × body weight, M.P.	Egg no. × body weight, r	Egg no. × maturity, M.P.	Egg no. × maturity, r
Between progenies	14	48·44	0·269	53·61	0·510	431·7	0·388
Within progenies	201	1·83	0·029	0·86	0·011	178·6	0·244

	Degrees of freedom	Mean products and correlation coefficients: Egg wt. × body wt., M.P.	Egg wt. × body wt., r	Egg wt. × maturity, M.P.	Egg wt. × maturity, r	Body wt. × maturity, M.P.	Body wt. × maturity, r
Between progenies	14	17·73	0·433	58·99	0·136	50·58	0·290
Within progenies	201	6·20	0·415	2·01	0·014	6·82	0·040

For $n = 13$, $r = 0·514$ for $P = 0·05$, and for $n = 200$, $r = 0·138$ and $0·181$ for $P = 0·05$ and $P = 0·01$ respectively.

There are no significant differences between sires for the body size of their daughters and the number of eggs laid. This is probably due to these two characters having been subjected to intensive selection over a long period. Egg weight, on the other hand, shows striking differences, and the period of maturity between progenies also differs almost significantly. This indicates that there is some scope for increasing the egg size, and, to a less extent, for decreasing the period of maturity in this stock by selecting suitable sires.

Correlation coefficients between progenies are not significant except between the number of eggs laid and body weight, which approaches the 5 % level. This correlation, particularly in the absence of any correlation in these two characters within progenies, would be interpreted to mean that sires which affected the body weight of their daughters would also affect their productivity in the same direction. There is a significant negative correlation between the number of eggs and the period of sexual maturity within the progenies. This is to be expected. Knox, Jull & Quinn (1935) also observed a significant correlation of $-0{\cdot}27$ in these two characters within a flock of White Leghorns. The significant positive correlation between body weight and egg weight within progenies indicates that birds with a bigger body produce larger eggs. This is the information obtained from simple correlations. If partial correlations are worked out, additional information useful in developing criteria for selection would become available. The discriminant function which is based on the variances and covariances between all pairs of characters utilizes all such information.

If the genotype of sire could be accurately measured for the various characters related to the egg production of its daughters, and if the relative weights to be attached to these measures were given, the correct formula for assessing the genetic value of the sire would become immediately available. In the present case, with four characters, x_1, x_2, x_3 and x_4, and with their weights, a_1, a_2, a_3 and a_4 respectively, this formula would be

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4.$$

Actually the sire's genotype is only estimated through the mean performance of the progeny, and these observed values are subject to extraneous causes of variation and may also be mutually correlated. This necessitates an adjustment to the weights to be attached to the observed values. The new formula in terms of these values and modified weights may be written down as

$$X = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4.$$

With the help of the discriminant function technique, the coefficients $b_1 \ldots b_4$ are so evaluated that the compound score X is most highly correlated with the true genetic value of the sire.

The value of the discriminant function scores will depend partly on the weights, a, initially chosen. For the present material the following three alternative sets of weights were suggested:

	a_1 (egg no.)	a_2 (egg weight)	a_3 (body weight)	a_4 (age at maturity)
Set I	7	6	5	-2
Set II	7	5	5	-3
Set III	8	5	5	-2

These coefficients must be interpreted as giving the relative values on a common scale of a unit change in the genetic measures of the four characters.

The two covariance matrices necessary for setting up the simultaneous equations in coefficients, b, are obtained from the mean squares and mean products given in Tables 1 and 2. The details of the calculation need not be shown here as the method is fully described by Fairfield Smith (1936). The solution of the equations gave the following values for b:

	b_1 (egg no.)	b_2 (egg weight)	b_3 (body weight)	b_4 (age at maturity)
Set I	0·7166	6·5628	2·3837	1·4984
Set II	1·4362	5·4173	1·7517	1·7266
Set III	0·2997	5·9595	3·1198	−1·7381

Since the coefficients are adjustable, we may put one of them, say b_1, equal to unity and rewrite them as follows:

	b_1	b_2	b_3	b_4
Set I	1	9·16	3·33	−2·09
Set II	1	3·77	1·22	−1·20
Set III	1	19·89	10·41	−5·80

In correspondence with the wide differences existing between progenies for egg weight, this character has the highest selective value in the discriminant function. The coefficients for egg number and body weight are considerably reduced when compared with the theoretical weight given to these characters. A numerical score can now be assigned to each sire by substituting the mean values of its progeny in the discriminant formula, and this score will best indicate the selective potentiality of the sire.

In the present data, the genetic component of various characters is not estimated with much accuracy, being based on only 14 degrees of freedom. The analysis for body weight appears particularly suspicious; because in the covariance matrix for the genetic component, obtained by subtracting the mean squares and mean products within progenies from those between progenies, the mean square for body weight is very low and the correlation coefficient between body weight and egg number exceeds unity! For this reason body weight was entirely omitted from the data, and a fresh discriminant function based on the other three characters was calculated. Its coefficients are:

	b_1 (egg no.)	b_2 (egg weight)	b_3 (age at maturity)
Set I	1	5·07	−0·81
Set II	1	3·47	−0·86
Set III	1	4·70	−0·82

The difference between the three sets of coefficients is now much less. Egg weight continues to have the highest selective value, but less disproportionate to egg number than in the original function.

The discriminant function can be altered by modifying the theoretical weights. It may be considered that one or more of the characters included in the discriminant function are of no direct value. The coefficients, a, for such characters may consequently be equated to zero. The selective value of these characters in the discriminant function will then depend only on their correlation with the other characters. As an illustration we recalculate the new function after putting the a coefficient for the period of maturity equal to zero, on the assumption that this character is of no direct monetary value. The coefficients of the function are altered as shown below:

	b_1 (egg no.)	b_2 (egg weight)	b_3 (age at maturity)
Set I	1	9·80	−0·61
Set II	1	9·19	−0·63
Set III	1	8·38	−0·61

The selective value of the period of maturity is further reduced as it now depends only on its correlation with the other two characters. The coefficients for egg weight are nearly doubled.

The object of selecting sires is to increase the mean productive value of the stock. It is of considerable interest to study what advantage the use of the discriminant function has over straight selection in this respect. Straight selection is equivalent to the use of theoretical weights, that is coefficients, a, in conjunction with the observed values for different characters. In other words, the value of the sire is scored by means of the formula

$$a_1x_1+a_2x_2+a_3x_3+a_4x_4$$

in place of the discriminant formula

$$b_1x_1+b_2x_2+b_3x_3+b_4x_4,$$

x_1, x_2, x_3 and x_4 representing in both formulae the mean performance of the progeny for the four characters.

Fairfield Smith (1936) has given the expectation of genetic advance by selection over the mean of the original population as

$$\frac{z}{q}BV^{\frac{1}{2}},$$

where z and q are determined by the degree of selection applied, B is the coefficient of regression of the true genetic value on the compound score used for selection, and V is the variance of this score between progenies.

With the intensity of selection fixed, the expectation of genetic advance when selection is made with the help of different formulae will be proportionate to the quantity $BV^{\frac{1}{2}}$. This quantity was calculated for the two formulae given above. The results are shown in Table 3.

Table 3. *Comparison of genetic advance with discriminant function and straight selection*

	All four characters				Index of genetic advance ($BV^{\frac{1}{2}}$)		% excess of advance by discriminant function
	a_1	a_2	a_3	a_4	Discriminant function	Straight selection	
Set I	7	6	5	−2	121·5	109·7	10·8
Set II	7	5	5	−3	126·9	120·8	5·1
Set III	8	5	5	−2	130·0	113·1	14·9

	Three characters (body weight omitted)					
	a_1	a_2	a_3			
Set I	7	6	−2	108·1	98·8	9·4
Set II	7	5	−3	118·6	113·3	4·6
Set III	8	5	−2	112·1	101·3	10·7
Set I′	7	6	0	91·2	71·0	28·5
Set II′	7	5	0	85·2	65·1	30·8
Set III′	8	5	0	93·3	70·3	32·7

The expected genetic advance is greater throughout when selection is based on the discriminant formula. This is a consequence of the discriminant formula being most highly correlated with the genetic value of the sire among the various alternative formulae that might be used. The magnitude of actual difference between the two is not appreciable here, being only of the order of 10 %; but when a character which on its own account is not considered to be of any value is included, the difference would be enlarged. This is illustrated in the last group of sets in Table 3, where the expectation of genetic advance through the use of the discriminant formula is 30 % greater than with the other formula. Selection by means of the discriminant function should prove definitely more profitable in such cases.

The question whether the estimate of genetic advance is significant is dependent on another, namely, whether there is a significant discrimination between sires, using the particular discriminant formula. In the absence of a significant discrimination there is no possibility of a significant advance by selection. The general problem of the tests of significance of discriminant functions is discussed by Fisher (1940), and an analysis of variance of the discriminant score, X, between progenies and within progenies would appear to provide a direct method for making such a test. When this test was applied to the discriminant scores based on all four characters, a significant difference between sires on 5 % level was observed for sets II and III of the a coefficients. In such an analysis, however, an adjustment of the degrees of freedom between the two components is necessary, by transferring p-1 degrees for p variates from 'within progenies' to 'between progenies' (Fisher, 1938). The present case is not strictly analogous, since the discriminant function is chosen not to maximize the differences between sires as such, but to discriminate between them in respect of their potential genetic value. It is not clear what bearing this point would have on the question of adjustment. Reallocating the degrees of freedom, however, it was found that the discriminant function corresponding to set II only just showed a significant difference on the 5 % level. These results indicate that there is a basis, though not appreciable, for selection among the sires examined. A similar analysis of the other discriminant functions considered would be useful.

DISCUSSION

The present results are of interest chiefly as an illustration of a very useful technique for making selection for several characters simultaneously. Their value in providing practical guidance to the breeder in selecting suitable sires is limited. To obtain discriminant formulae of greater practical utility it is necessary to analyse data from a larger number of sires mated either to the same set of hens or to randomly selected groups so that genetic differences between dams are either eliminated from the comparisons between sires or are randomly distributed among them.

The appropriateness of the theoretical weights employed in evaluating the discriminant functions also requires a closer examination. These weights should be proportional to the monetary value or value on some other common scale of a unit change in the various characters included in the function, the units being those in which the measurements of the characters are analysed. For the present analysis the weights suggested by poultry experts have been employed; and the four weights, a_1, a_2, a_3 and a_4, must be taken as being proportionate to the monetary value of an egg, the premium obtainable per gramme increase in egg weight, the premium for a 10 g. increase in body weight of the hen after making an allowance for the increase in feed consumption that this increase in body weight may mean, and the feed value for each day by which the start of laying is delayed. The weights will have to be altered if they do not conform to this interpretation, and this would result in a corresponding change in the coefficients of the discriminant function.

SUMMARY

Egg production in poultry is controlled by genetic factors at least partly inherited from the sire. Selection of sires is therefore important in ensuring highly productive progenies. A discriminant function score based on different characters of the daughters related to egg production would be the best criterion for this purpose, since the coefficients of the function are so obtained as to maximize the regression of the genetic value of the sire on his phenotype as reflected in the mean performance of the progeny.

Data on the number of eggs laid, egg weight, body weight of the pullet and her period of maturity were examined for 218 Brown Leghorn pullets belonging to progenies of 17 sires. There were significant differences between progenies in egg weight and period of maturity. None of the simple correlations between progenies was significant except that for egg number and body weight which approached significance. The discriminant function obtained showed that egg weight had the highest selective value. Selection with the help of this function was found to register a somewhat higher genetic advance than straight selection. The application of the analysis of variance for testing the significance of discrimination between sires is pointed out.

I am thankful to Dr F. Yates, Statistician, Rothamsted Experimental Station, for suggesting the present investigation, and to Dr A. W. Greenwood of the Institute of Animal Genetics, University of Edinburgh, for kindly placing the necessary data at my disposal.

REFERENCES

Fisher, R. A. (1936). *Ann. Eugen., Lond.*, **7**, 179.
Fisher, R. A. (1938). *Statistical Methods for Research Workers*, 7th ed. London: Oliver and Boyd.
Fisher, R. A. (1940). *Ann. Eugen., Lond.*, **10**, 422.
Hays, F. A. (1929). *J. Agric. Res.* **38**, 511.
Hays, F. A. (1941). *Poult. Sci.* **20**, 217.
Knox, C. W., Jull, M. A. & Quinn, J. P. (1935). *J. Agric. Res.* **50**, 573.
Pearl, R. (1912). *J. Exp. Zool.* **13**, 153.
Smith, H. Fairfield (1936). *Ann. Eugen., Lond.*, **7**, 240.