TESTS FOR DISCRIMINANT FUNCTIONS

E. J. WILLIAMS.

(received 7 November 1960)

1. Introduction

In previous papers (Bartlett 1951, Williams 1952a, b, c; 1953, 1955) certain exact tests for discriminant functions have been discussed. These papers have also indicated how the tests may usefully be applied to a wide range of problems: the significance of discriminant function coefficients, the interpretation of interactions in factorial experiments, the concurrence of regression lines with the same independent variables and the proportionality of regression lines with different independent variables, the comparison of sets of ratios, and the testing of arbitrary scores applied to frequency data.

For all of these problems, the tests required are generally of two kinds. First are required tests of the adequacy of a single discriminant function (or some analogous linear combination of variables) to represent the variation among the multivariate populations from which the data have been drawn; these tests readily generalize to tests of the adequacy of a set of such linear functions. Secondly are required tests of, and corresponding fiducial limits for, the coefficients in the given discriminant functions.

The main object of this paper is to present the tests in a more explicit and general form than has been done previously, and to give an alternative derivation. Generalizations of the earlier work to the testing of a set of r discriminant functions will be given.

If r discriminant functions are adequate to specify the variation in a p-variate population, this implies that the population means satisfy $p - r$ linear relations. The determination of such linear relations has important applications in instrument calibration and in some econometric problems (Williams 1959). In an earlier paper (Williams 1955) it was erroneously stated that with p independent variables, the test of s linear relations was equivalent to the test for $p-s$ discriminant functions. The error was pointed out by Bartlett (1957). The fact is that, if the variation among populations with p variables is adequately described in terms of $p-s$ linear functions of them (the discriminant functions), this implies that, within the limits of sampling error, there are s linear relations among the variables. Indeed, the test for the adequacy of $p-s$ discriminators is by implication a test for the

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

existence of s such linear relations. However, the test for the coefficients of the discriminators does not lead to a test for the coefficients of the linear relations; this is a consequence of the fact that the variables orthogonal to the hypothetical discriminators in the sample are not necessarily orthogonal to them in the population. Indeed, it does not appear possible to make exact tests or fiducial limits for the constants in a linear functional relation, since the specification of the hypothetical null variate does not lead to an explanatory variate in terms of which the variation among the popluations can be expressed.

The author (Williams 1959) has shown that important significance tests in this field may be put into the form of an analysis of covariance; the significance of the reduced variation after the elimination of an "explanatory variate" defined by the null hypothesis provides a test for the adequacy of that hypothetical explanatory variate. This has been the approach adopted in the other references mentioned at the beginning of this paper, and it will be developed explicitly in the following sections.

Incidentally it will be shown that the general test for "direction" of an explanatory variate — the test for the coefficients in the variate — is nothing more than a test of deviation of the dependent variables from regression on the explanatory variate.

2. Tests for a single explanatory variate

We suppose that we have a sample of $n+1$ observations from a $p+q$-variate population consisting of two sets of variables:

$$x_i \quad (i = 1, 2, \cdots, p)$$
$$y_j \quad (j = 1, 2, \cdots, q).$$

The discrimination problem may then be stated formally as the requirement that the regression of the x_i on the y_j may be expressed in terms of a smaller set of compound variates $\xi_i (i = 1, 2, \cdots, r)$, linear functions of the x_i; in other words, that the x_i be regarded as distributed in r rather than p dimensions, apart from sampling error. We first deal with a single explanatory variate ($r = 1$).

We use the following notation, following Williams (1955).

T matrix of sums of squares and products of the x_i

P matrix of sums of products of the x_i and the y_j

U matrix of sums of squares and products of the y_j

x an observation-vector of the x-variables

y an observation-vector of the y-variables

γ the vector of coefficients of the x_i in the explanatory variable

B matrix of sums of squares and products of regression of the x_i on the y_j.

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

Thus $\xi = \gamma' x$.

Matrices of sums of squares and products will be termed dispersion matrices.

Clearly $B = PU^{-1}P'$ and has q degrees of freedom. We write $W = T - PU^{-1}P'$, the matrix of residual dispersion, with $n - q$ degrees of freedom.

The test for the adequacy of, and the constants in, the explanatory variate ξ may be carried out as an analysis of covariance (see Williams 1955). We eliminate ξ from the total and residual dispersion of the x_i, and test the ratio of the reduced determinants. Clearly if ξ is an adequate explanatory variate, this determinantal ratio will be distributed independently of the y_j. The form of this derivation shows that it really provides a test of deviation from regression on the variate ξ. As with covariance analysis with a single dependent variable, the ratio can be broken into two factors providing tests of deviation from regression (i.e. direction) with $p - 1$ degrees of freedom, and a factor for a test of collinearity. Here we propose to derive these two tests in a different way.

The relevant sums of squares for ξ are

$$\begin{array}{ll} \text{Residual} & \gamma' W \gamma \\ \text{Total} & \gamma' T \gamma. \end{array}$$

Then the determinants of reduced dispersion of the x_i, after ξ has been eliminated, are

$$\begin{array}{ll} \text{Residual} & |W|/\gamma' W \gamma \\ \text{Total} & |T|/\gamma' T \gamma. \end{array}$$

The overall test criterion found by eliminating ξ is thus seen to be

$$\frac{|W|}{|T|} \cdot \frac{\gamma' T \gamma}{\gamma' W \gamma}. \tag{1}$$

This is a determinantal ratio with $n - 1$ total degrees of freedom, and $p - 1$ and q degrees of freedom for the (reduced) x_i and the y_j. It is denoted by $(n - 1\colon p - 1, q)$ (Williams 1959).

Now consider the regression function of ξ on the y_j. It is

$$X = \gamma' PU^{-1} y,$$

a linear function of the y_j. Since this regression function has been chosen to minimize the sum of squares of deviation from regression on ξ as a whole, we may analyse this sum of squares into portions for reduced regression on the x_i and residual from the x_i, regardless of the dependence of X on ξ. These two portions contain $p - 1$ and $n - p$ degrees of freedom respectively; their ratio provides the test for the coefficients in ξ.

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

We have

$$\begin{aligned} &\text{Total sum of squares of } X = \gamma' P U^{-1} P' \gamma = \gamma' B \gamma \\ &\text{Sum of products of } X,\ \xi \qquad\qquad\qquad = \gamma' B \gamma \\ &\text{Sum of squares of } \xi \qquad\qquad\qquad\quad = \gamma' T \gamma. \end{aligned}$$

Hence sum of squares of departure of X from regression on ξ

$$\begin{aligned} &= \gamma' B \gamma - (\gamma' B \gamma)^2 / \gamma' T \gamma \\ &= \gamma' B \gamma \cdot \gamma' W \gamma / \gamma' T \gamma. \end{aligned}$$

Vector of sums of products of X with the $x_i = \gamma' P U^{-1} P'$
$= \gamma' B.$

Hence, residual sum of squares of X

$$\begin{aligned} &\frac{1}{|T|} \begin{vmatrix} \gamma' B \gamma & \gamma' B \\ B \gamma & T \end{vmatrix} \\ &= \gamma' B \gamma - \gamma' B T^{-1} B \gamma \\ &= \gamma' B T^{-1} W \gamma. \end{aligned}$$

For the test of direction, the $(n - 1\colon p - 1,\ 1)$ ratio is therefore

(2) $$D = \frac{\gamma' B T^{-1} W \gamma \cdot \gamma' T \gamma}{\gamma' B \gamma \cdot \gamma' W \gamma}.$$

This ratio may be tested by means of the F test:

$$\begin{aligned} F_{(p-1,\, n-p)} &= \frac{n-p}{p-1} \frac{1-D}{D} \\ &= \frac{n-p}{p-1} \cdot \frac{\gamma' B \gamma \cdot \gamma' W \gamma - \gamma' B T^{-1} W \gamma \cdot \gamma' T \gamma}{\gamma' B T^{-1} W \gamma \cdot \gamma' T \gamma}. \end{aligned}$$

The collinearity criterion is the ratio of (1) to (2):

(3) $$C = \frac{|W|}{|T|} \cdot \frac{\gamma' B \gamma}{\gamma' B T^{-1} W \gamma}$$

which is a $(n - 2\colon p - 1,\ q - 1)$ ratio.

These results are equivalent to those given in Williams (1952a, 1955) where, however, they are expressed in terms of sample canonical correlations.

In applying these tests, we need first to test collinearity; only if there is no evidence of departure from collinearity can a valid test of direction (the discriminant function coefficients) be made.

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

3. Application where one set of variables represents group differences

The analysis of the previous section is generally applicable, but it is instructive to consider separately the situation in which one of the variables represents group differences, as for instance with a q-variate population classified into $p+1$ groups, so that there are p independent comparisons between the groups. When the hypothetical discriminator (or explanatory variate) is a function of the metrical variable, the analysis is as above; but when the discriminator is a function of the group differences, the analysis, though essentially the same, has a different appearance. A good example is given by Barnard (1935) and discussed from the present viewpoint by Williams (1959). There the explanatory variate is *time*, which is assumed to have a known value for each of the groups considered.

Suppose then that the x_i represent differences among $p+1$ groups, and that the explanatory variate ξ assigns a value to each of the groups, but that the x_i are otherwise undefined. We shall write

C dispersion matrix of the y_j between groups
V dispersion matrix of the y_j within groups
p_ξ(row) vector of sums of products of ξ with the y_j
$t_{\xi\xi}$ sum of squares of ξ.

For X, the regression function of ξ on the y_j, we have $X = p_\xi U^{-1}y$. The analysis of variance of X is as follows:

	D.F.	Sum of squares
Deviation from regression	$p-1$	by subtraction
Residual (within groups)	$n-p$	$p_\xi U^{-1}VU^{-1}p'_\xi$
Total, reduced	$n-1$	$p_\xi U^{-1}p'_\xi - (p_\xi U^{-1}p'_\xi)^2/t_{\xi\xi}$

Thus we are led in this case also to the direction criterion

$$(n-1 : p-1, 1) = \frac{p_\xi U^{-1}VU^{-1}p'_\xi \cdot t_{\xi\xi}}{p_\xi U^{-1}p'_\xi(t_{\xi\xi} - p_\xi U^{-1}p'_\xi)} \tag{4}$$

which can be seen to be formally equivalent to (2).

$$F = \frac{n-p}{p-1} \cdot \frac{p_\xi U^{-1}CU^{-1}p'_\xi \cdot t_{\xi\xi} - (p_\xi U^{-1}p'_\xi)^2}{p_\xi U^{-1}VU^{-1}p'_\xi \cdot t_{\xi\xi}}.$$

The collinearity criterion is

$$\frac{|V|}{|U|} \cdot \frac{p_\xi U^{-1}p'_\xi}{p_\xi U^{-1}VU^{-1}p'_\xi} \tag{5}$$

which corresponds to (3).

If the matrices T and P were defined, then would

$$V = U - P'T^{-1}P;$$

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

however, it is not necessary for the determination of V that the variables x_i and the matrices T and P be defined.

4. Extension to r explanatory variates

When more than one explanatory variate is required for the interpretation of the variation among the x_i (as regards their association with the y_j), the interpretation is not so simple, nor is the case so important practically. However, the analysis is straightforward and introduces no new principle.

We consider a vector Ξ of r explanatory variates, defined in terms of the x_i by the $p \times r$ matrix Γ:

$$\Xi = \Gamma' x.$$

Then the r regression functions of Ξ on the y_j are given by the vector

$$X = \Gamma' P U^{-1} y.$$

The analysis now follows exactly the same lines as before, except that sums of squares, etc., are replaced by matrices.

The dispersion matrices of X are

Total	$\Gamma' P U^{-1} P' \Gamma$	$= \Gamma' B \Gamma$
Departure from Ξ	$\Gamma' B \Gamma (\Gamma' T \Gamma)^{-1} \Gamma' W \Gamma$	
Residual	$\Gamma' B T^{-1} W \Gamma.$	

The overal criterion is

$$\frac{|W|}{|T|} \cdot \frac{|\Gamma' T \Gamma|}{|\Gamma' W \Gamma|},$$

a $(n - r\colon p - r,\ q)$ ratio.

The direction criterion (for testing simultaneously r sets of directions in $p - r$ dimensions) is

$$D = \frac{|\Gamma' T \Gamma| \cdot |\Gamma' B T^{-1} W \Gamma|}{|\Gamma' B \Gamma| \cdot |\Gamma' W \Gamma|},$$

a $(n - r\colon p - r, r)$ ratio.

Finally the test for coplanarity is given by the $(n - 2r\colon p - r\ q - r)$ ratio

$$C = \frac{|W| \cdot |\Gamma' B \Gamma|}{|T| \cdot |\Gamma' B T^{-1} W \Gamma|}.$$

If $r \geqq q$, then clearly a test of coplanarity is not relevant, and this ratio becomes unity.

In the particular case where $r = p - 1$, these ratios can be tested by means of an analysis of variance. Such a test, of the adequacy of $p - 1$

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

explanatory variates, was wrongly described by Williams (1955) as a test of the null variate orthogonal to Ξ in the sample.

For the detailed tests we have, when $r = p - 1$, the following analysis of variance.

	D.F.	Sum of squares
Direction	$p-1$	$1-D = 1 - \dfrac{\lvert\Gamma' T\Gamma\rvert \cdot \lvert\Gamma' BT^{-1}W\Gamma\rvert}{\lvert\Gamma' B\Gamma\rvert \cdot \lvert\Gamma' W\Gamma\rvert}$
Coplanarity	$q-p+1$	$D(1-C) = \dfrac{\lvert\Gamma' T\Gamma\rvert \cdot \lvert\Gamma' BT^{-1}W\Gamma\rvert}{\lvert\Gamma' B\Gamma\rvert \cdot \lvert\Gamma' W\Gamma\rvert} - \dfrac{\lvert W\rvert \cdot \lvert\Gamma' T\Gamma\rvert}{\lvert T\rvert \cdot \lvert\Gamma' W\Gamma\rvert}$
Residual	$n-p-q+1$	$DC = \dfrac{\lvert W\rvert \cdot \lvert\Gamma' T\Gamma\rvert}{\lvert T\rvert \cdot \lvert\Gamma' W\Gamma\rvert}$
Total	$n-p+1$	1

The direction test given in this analysis differs from that for general r in that the term for coplanarity has been eliminated from the residual. It is thus a test for 'partial direction', whereas the tests given previously are for 'simple direction'. Since if there is departure from coplanarity the test for the direction of a vector is not relevant, the simple direction test is usually appropriate.

When $p > q$ the term for coplanarity drops out of the analysis.

To show the equivalence of these results for $r = p - 1$ with those of Williams (1955) we consider the column vector β whose ith element is $(-1)^i$ (determinant of Γ with ith row omitted).

Then it may be shown that

$$\Gamma' \beta = 0.$$

Also, for any $p \times p$ matrix G,

$$\frac{\lvert\Gamma' G\Gamma\rvert}{\lvert G\rvert} = (-1)^{p-1} \begin{vmatrix} G^{-1} & \Gamma \\ \Gamma' & 0 \end{vmatrix}$$
$$= \beta' G^{-1} \beta.$$

Thus, for the direction criterion, we have

$$D = \beta' T^{-1} \beta \left(\frac{1}{\beta' B^{-1} \beta} + \frac{1}{\beta' W^{-1} \beta} \right)$$

and for the overall criterion,

$$\frac{\beta' T^{-1} \beta}{\beta' W^{-1} \beta}.$$

Hence the analysis of variance given above may be written as

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

	D.F.	Sum of squares.
Direction	$p-1$	$(\beta' T^{-1}\beta)^{-1} - (\beta' B^{-1}\beta)^{-1} - (\beta' W^{-1}\beta)^{-1}$
Coplanarity	$q-p+1$	$(\beta' B^{-1}\beta)^{-1}$
Residual	$n-p-q+1$	$(\beta' W^{-1}\beta)^{-1}$
Total	$n-p+1$	$(\beta' T^{-1}\beta)^{-1}$

which corresponds to that given by Williams (1955).

5. Distribution of determinantal ratios

A $(n : p, q)$ ratio is the ratio of two p-variate determinants of sums of squares and products, with $n-q$ and n degrees of freedom. Owing to the duality between p and q, a $(n : p, q)$ and a $(n : q, p)$ ratio have the same null distribution.

Rao (1951) has shown that, to a very close approximation, a $(n : p, q)$ ratio is distributed as the sth power of a

$$(s[n - \tfrac{1}{2}(p + q + 1)] + \tfrac{1}{2}pq + 1 : pq, 1)$$

ratio, where

$$s^2 = \frac{p^2q^2 - 4}{p^2 + q^2 - 5}.$$

This result is exact when either p or q is one or two. Thus any such ratio can be tested by means of the F test. If the ratio is denoted by R, then (approximately)

$$F = \frac{s[n - \tfrac{1}{2}(p + q + 1)] - \tfrac{1}{2}pq + 1}{pq} \cdot \frac{1 - R^{1/s}}{R^{1/s}}.$$

6. Large-sample analysis

When $n+1$, the sample size, is large (or the population covariance matrix is known) the analysis takes a particularly simple form. We shall assume that $B = o(n)$, and write

$$\operatorname*{Lt}_{n\to\infty} \frac{T}{n} = \operatorname*{Lt}_{n\to\infty} \frac{W}{n} = A$$

for the covariance matrix of the x_i. Then in the limiting case we have for the determinantal ratios considered earlier

$$\frac{|W|}{|T|} = |WT^{-1}| = |I - n^{-1}BA^{-1}| \to 1 - n^{-1} \operatorname{tr}(BA^{-1})$$

$$\frac{|\Gamma' T\Gamma|}{|\Gamma' W\Gamma|} \to 1 + n^{-1} \operatorname{tr}(\Gamma' B\Gamma(\Gamma' A\Gamma)^{-1})$$

$$\frac{|\Gamma' BT^{-1}W\Gamma|}{|\Gamma' B\Gamma|} \to 1 - n^{-1} \operatorname{tr}(\Gamma' BA^{-1}B\Gamma(\Gamma' B\Gamma)^{-1})$$

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

The traces are distributed as χ^2, with degrees of freedom pq, rq, and $p+q-r$ respectively.

Hence we have the analysis of χ^2

	D.F.	χ^2
Explanatory variables	rq	$\operatorname{tr}(\Gamma' B\Gamma(\Gamma' A\Gamma)^{-1})$
Direction	$r(p-r)$	$\operatorname{tr}(\Gamma' BA^{-1}B\Gamma(\Gamma' B\Gamma)^{-1} - \Gamma' B\Gamma(\Gamma' A\Gamma)^{-1})$
Coplanarity	$(p-r)(q-r)$	$\operatorname{tr}(BA^{-1} - \Gamma' BA^{-1}B\Gamma(\Gamma' B\Gamma)^{-1})$
Total	pq	$\operatorname{tr}(BA^{-1})$

This analysis reduces when $r = 1$ to the form

	D.F.	χ^2
Explanatory variable	q	$\dfrac{\gamma' B\gamma}{\gamma' A\gamma}$
Direction	$p-1$	$\dfrac{\gamma' BA^{-1}B\gamma}{\gamma' B\gamma} - \dfrac{\gamma' B\gamma}{\gamma' A\gamma}$
Collinearity	$(p-1)(q-1)$	$\operatorname{tr}(BA^{-1}) - \dfrac{\gamma' BA^{-1}B\gamma}{\gamma' B\gamma}$
Total	pq	$\operatorname{tr}(BA^{-1})$

We see that in this limiting case the two tests required are included as part of the same analysis. In other examples, the population covariance matrix may be known apart from a constant factor. Then an independent estimate of the variance is required; the analysis of χ^2 will be replaced by an analysis of variance, and the significance will be tested by the F-test. Examples have been given by Williams (1952b: interpretation of interactions; 1959: concurrent and proportional regressions).

References

Barnard, M. M. (1935) — The secular variations of skull characters in four series of Egyptian skulls. Ann. Eugen., Lond., 6, 352—371.

Bartlett, M. S. (1951) — The goodness of fit of a single hypothetical discriminant function in the case of several groups. Ann. Eugen., Lond., 16, 199—214.

Bartlett, M. S. (1957) — A note on tests of significance for linear functional relationships. Biometrika 44, 268—269.

Rao, C. R. (1951) — An asymptotic expansion of the distribution of Wilks' Λ criterion. Bull. Int. Stat. Inst., 33, 177—180.

Williams, E. J. (1952a) — Some exact tests in multivariate analysis. Biometrika 39, 17—31.

Williams, E. J. (1952b) — The interpretation of interactions in factorial experiments. Biometrika 39, 65—81.

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press

Williams, E. J. (1952c) — Use of scores for the analysis of association in contingency tables. Biometrika 39, 274—289.
Williams, E. J. (1953) — Tests of significance for concurrent regression lines. Biometrika 40, 297—305.
Williams, E. J. (1955) — Significance tests for discriminant functions and linear functional relationships. Biometrika 42, 360—381.
Williams, E. J. (1959) — "Regression Analysis". Wiley, New York.

Division of Mathematical Statistics
C.S.I.R.O. Canberra.

https://doi.org/10.1017/S1446788700026719 Published online by Cambridge University Press