DOCUMENT RESUME

ED 065 552 TM 001 697

AUTHOR	Sachdeva, Darshan
TITLE	Multivariate Analog of Hays Omega-Squared.
PUB DATE	Apr 72
NOTE	5p.; Paper presented at the Annual Meeting of the American Educational Research Association (56th, Chicago, Illinois, April 3-7, 1972)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Analysis of Variance; *Mathematical Concepts; *Statistical Analysis
IDENTIFIERS	*Hays (W L)

ABSTRACT

The multivariate analog of Hays omega-squared for estimating the strength of the relationship in the multivariate analysis of variance has been proposed in this paper. The multivariate omega-squared is obtained through the use of Wilks' lambda test criterion. Application of multivariate omega-squared to a numerical example has been provided so as to help understand the mechanics of the formulas necessary for the computation of multivariate omega-squared. (Author/DB)

21.12

Multivariate Analog of Hays ω^2

Darshan Sachdeva

University of California, Berkeley **

ABSTRACT

The multivariate analog of Hays ω^2 for estimating the strength of the relationship in the multivariate analysis of variance has been proposed in this paper. The multivariate ω^2 (denoted by ω^2_{mult}) is obtained through the use of Wilks's Λ test criterion. Application of ω^2_{mult} to a numerical example has been provided so as to help understand the mechanics of the formulas necessary for the computation of ω^2_{mult}.

** Paper presented at the 56th Annual Meeting of the American Educational Research Association, Chicago, Ill. April 3 - 7, 1972.

Multivariate Analog of Hays ω^2

Darshan Sachdeva

University of California, Berkeley

It has periodically been suggested to the behavioral researchers that test statistics such as t-ratios or F-ratios serve to indicate the statistical significance of observed results and these statistics do not describe their practical importance. The assessment of the practical significance of the experimental effects, given that they have been found to be statistically significant, has received attention from time to time and Hays (1963) suggested the following statistic for this purpose:

$$\hat{\omega}^2 = \frac{SSB - (k-1)MSW}{SST + MSW} = \frac{SSB - \frac{k-1}{N-k}SSW}{SST + \frac{1}{N-k}SSW} \qquad (1)$$

where SSB is the between-group sum of squares, MSW is the within-group mean square, and SST is the total sum of squares for k groups in one-way analysis of variance.

Clearly, $\hat{\omega}^2$ as given by (1) is applicable to univariate analysis of variance. The univariate analysis of variance refers to experiments in which one or more independent variables may be employed but only one dependent variable. On the other hand, multivariate analysis of variance is a simple extension of the univariate analysis of variance where analysis is being carried out using two or more dependent variables simultaneously. This paper examines the notion of ω^2 in case of multivariate analysis of variance.

The multivariate extension of (1) is obtained by replacing each sum of squares by the determinant of the corresponding matrix of sums of squares and sums of cross-products. For one-way multivariate analysis of variance for k groups involving p dependent variables, the elements of matrix W for the within-group sums of squares and sums of cross-products are given by

$$SS(W_i, W_j) = \sum_{s=1}^{k} \sum_{t=1}^{n_s} (X_{ist} - \bar{X}_{is})(X_{jst} - \bar{X}_{js}), \qquad \begin{array}{l} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p. \end{array}$$

It may be noted that when $i = j$, the above expression reduces to

$$SSW_i = \sum_{s=1}^{k} \sum_{t=1}^{n_s} (X_{ist} - \overline{X}_{is})^2 ,$$

which is the within-group sum of squares for the ith dependent variable as obtained in the univariate analysis of variance. In like manner, the research worker can define the matrix B for the between-group sums of squares and sums of cross-products. The matrix T for the total sums of squares and sums of cross-products is obtained in a similar fashion. Now the multivariate extension of (1) is given by

$$\hat{\omega}^2_{mult} = \frac{|B| - \frac{k-1}{N-k}|W|}{|T| + \frac{1}{N-k}|W|} \tag{2}$$

where N is the total number of subjects in k groups, and $|B|$, $|W|$, $|T|$ are the determinants of matrices B, W, and T. After some algebraic simplifications (2) reduces to

$$\hat{\omega}^2_{mult} = 1 - \frac{N \cdot \Lambda}{N - k + \Lambda} \tag{3}$$

where $\Lambda = |W| / |T|$ is known as Wilks's lambda test criterion (Anderson, 1958; Morrison, 1967) and $|B|$ has been replaced by ($|T| - |W|$).

The $\hat{\omega}^2_{mult}$ as defined by (3) can also be estimated by F-ratio using the fact (Rao, 1965) that

$$F = \frac{1 - \Lambda^{1/S}}{\Lambda^{1/S}} u \tag{4}$$

where

$$S = \sqrt{\left[p^2 (df_h)^2 - 4\right] \div \left[p^2 + (df_h)^2 - 5\right]}$$

and

$$u = \frac{S(2df_h + df_e - p - 1) - p(df_h) + 2}{2p(df_h)}$$

and where p is the number of dependent variables, df_h and df_e are the degrees of freedom for the hypothesis and for within-group sum of squares respectively.

By using (4) in (3), it is seen that

$$\hat{\omega}^2_{mult} = 1 - \frac{N u^S}{(N-k)(F + u)^S + u^S} \tag{5}$$

An example is presented illustrating the use of formulas (3) and (5). The computed F-value used in formula (5) was obtained through the application of a computer program written by Finn (1968).

An Example:

In an analysis of language development among four groups of elementary school children who were measured on five variables it was found through the application of one-way multivariate analysis of variance for testing the null hypothesis of equality of mean vectors that $F = 6.6355$ with $\nu_1 = 15$ and $\nu_2 = 513.8655$. The multivariate null hypothesis was rejected as the p-value associated with the above F-value was less than $p = 0.0001$. The total number of subjects used in four groups was $N = 194$. It was also found that $\Lambda = |W| / |T| = 471333.0832/761723.7792 = 0.6188$, where $T = B + W$ and B and W have the same definitions accorded to them earlier in this narration.

The estimate of ω^2_{mult} using either formula (4) or formula (5) was found to be $\omega^2_{mult} = 0.37$. Therefore, nearly 37% of the total group variability is accounted for by the five language variables. In addition, it is seen that the estimates of Hay's ω^2 for each of the five variables in this case were: $\omega_1^2 = 0.011$; $\omega_2^2 = 0.100$; $\omega_3^2 = 0.029$; $\omega_4^2 = 0.322$; $\omega_5^2 = 0.015$ respectively. Since the dependent variables are correlated, each of the ω_i^2 represents the variation in groups that is accounted for purely by the ith variable only plus the variation that is explained jointly in some unknown fashion by the remaining dependent variables. Thus, $\omega^2_{mult} = 0.37$ represents the proportion of the variance that is explained by the five language variables when they are subjected to analysis simultaneously. Furthermore, it may be inferred that there is roughly 37% of discriminatory power in this set of dependent variables for the present data used in this example.

REFERENCES

Anderson, T. W. (1958) An introduction to multivariate statistical analysis. New York: John Wiley and Sons.

Finn, J. D. (1968) Multivariance-univariate and multivariate analysis of variance, covariance, and regression: a fortran IV program. Department of Educational Psychology, State University of New York at Buffalo.

Hays, W. L. (1963) Statistics for psychologists. New York: Holt, Rinehart and Winsto

Morrison. D. F. (1967) Multivariate statistical methods. New York: McGraw-Hill.

Rao. C. R. (1965) Linear statistical inference and its applications. New York: John Wiley and Sons.