CANONICAL VARIATE ANALYSIS AND RELATED TECHNIQUES

Richard B. Darlington and
Sharon L. Weinberg[1]
Cornell University
and
Herbert J. Walberg
University of Illinois

This paper discusses statistical methods for studying relations between two sets of variables, when each set contains more than one variable. The number of methods discussed is about 20 or 25, depending on how one counts. About a third of the methods are old methods criticized here, another third are old methods mentioned favorably, and the rest are new methods published here for the first time.

It has not been adequately recognized that problems involving two sets of variables arise frequently in almost every area of the behavioral sciences, as the following list of examples will attest:

1. Practical prediction problems
   (a) Test battery vs. criterion variables.
   (b) Two test batteries.
2. Theoretical problems about content relations between two sets
   (a) Stimulus vs. response variables, or independent vs. dependent variables.
   (b) Two sets of response variables—for example, interests vs. career plans (Cooley, 1967); interests vs. academic achievement (Bargmann, 1961); affective factors vs. academic achievement (Kahn, 1969); biological vs. behavioral variables.
3. Problems involving variables measured on two occasions
   (a) Time 1 vs. time 2, as in longitudinal studies.
   (b) Condition 1 vs. condition 2, as in two-condition experiments.
4. Problems involving paired subjects
   (a) Naturally paired subjects—for example, husband and wife (Hope, 1969; Massey, Frank, & Lodahl, 1968, p. 83), teacher and student (Walberg, Welch, & Rothman, 1969).
   (b) Experimentally paired subjects—for example, experimenter and subject, or two subjects.

The items in this list are not intended to be mutually exclusive or exhaustive; they are merely illustrative of the prevalence of the problems to be discussed.

Once a behavioral scientist has determined that he faces a

[1] Now at New York University.

problem involving two sets of variables, he typically seeks a source for statistical methods applicable to such problems. In our opinion, almost all existing sources on these problems have three deficiencies.

First, they fail to make clear that, for any given set of data, there are several very different hypotheses that may be tested about the relations between two sets of variables, and different statistical techniques must be used to test the different hypotheses. Instead, they usually give the impression that there is only one such hypothesis, and therefore only one statistical technique is needed—Hotelling's canonical variate analysis (CVA).

Second, most discussions of CVA are restricted almost entirely to a description of the underlying mathematical theory, computing directions, and perhaps an example of the computations. Very little is usually said about the logic of the method, so that the reader is unable to judge for himself whether the method described can actually be used to test the hypothesis of interest to him.

Third, even when CVA is appropriate, several prominent sources have recommended misleading interpretations of the statistics computed in CVA.

Perhaps because of these deficiencies, many behavioral scientists have concluded incorrectly that CVA has few or no valid and important uses in the behavioral sciences. The use originally proposed by Hotelling has been rejected by most behavioral scientists. Hotelling proposed using CVA to find the weighted average of several criterion variables forming "the most predictable criterion (1935)" for a test battery. The reaction of most behavioral scientists to this proposed use is typified by Gulliksen (1968), who wrote very sensibly, "there is no particular interest in determining a criterion simply because one is able to predict it [p. 540]."

The purpose of this paper is to remedy the deficiencies mentioned above. This paper describes several statistical techniques for studying different questions about the relations between two sets of variables, and specifies the different problems for which each technique is most appropriate. It identifies several problems for which CVA has been prominently suggested but for which other statistical techniques are more appropriate, but argues also that there are valid and important uses of CVA which are generally ignored.

A common and simple notation is used throughout the paper. The number of variables in one set X is n_x, and the number of variables in the other set Y is n_y. The correlation matrix of the (n_x+n_y) variables is R. This matrix is partitioned into the two within-set matrices R_{xx} and R_{yy}, and the between-set matrix R_{xy} and its transpose R_{yx}. All the questions discussed in this paper are questions about the nature of R.

Since several kinds of characteristics of subjects will be discussed, it will be useful to make a few arbitrary distinctions in terminology. The term *variable* will refer to the (n_x+n_y) observed variables in the X or Y sets. The terms *linear composite* and *weighted average* will be used interchangeably to refer to composites of the observed X or Y variables. The term *variate* will be used to refer only to canonical variates, which are certain weighted averages of the X and Y variables. The term *trait* will be a general term which includes all of these, plus variables not observed in a given study, and including variables for which a satisfactory operational measure may not even exist.

The paper is divided into three sections, each dealing with a different set of questions about R:

1. Questions about the number and nature of mutually independent relations between the two sets of variables.
2. Questions about the degree of overlap or redundancy between the two sets.
3. Questions about the similarity between the two within-set correlation or covariance matrices.

To illustrate these questions, suppose we have a sample of 200 elementary-school teachers, each of whom teaches a separate class. Let the two X variables be the teachers' average college grades in English and mathematics courses respectively, and let the two Y variables be measures of the mean performance on reading and math tests of the 200 classes. An artificial example of the correlation matrix of these four variables is shown in Table 1. As in any statistical method, correlation does not automatically imply causation; however, under certain conditions (for example, random assignment of subjects on the independent variables) causation can be inferred.

In this particular example, the elements in the 2×2 submatrix R_{xx} are identical to those in R_{yy}. This answers the most common question of type 3 in the above list, concerning relations between the two within-set correlation matrices; in this example the two are identical.

On the other hand, there is only a moderate relationship between the two sets of variables (question 2 in the list), since none of the elements in R_{xy} is very high.

Question 1 concerns the pattern, rather than the size, of the correlations in submatrix R_{xy}. In Table 1, the correlation of Y_2 with any X variable is exactly half that of Y_1 with the same variable. Thus

TABLE 1

An Artificial Correlation Matrix R

	X_1	X_2	Y_1	Y_2
X_1	1.0	.6	.4	.2
X_2	.6	1.0	.2	.1
Y_1	.4	.2	1.0	.6
Y_2	.2	.1	.6	1.0

the relation of Y_2 to the X variables is not independent of that of Y_1; there is only one relationship between the two sets, rather than two independent relationships.

This example has oversimplified the nature of the questions in each of the three sets. These will be developed more fully in later sections.

Since previous writers have generally failed to distinguish among these three sets of questions, CVA has been prominently suggested for all three sets. We argue that CVA is often appropriate for problems in the first set, rarely appropriate for those in the second set, and never appropriate for those in the third.

Any of the three major sections can be read and understood reasonably well by readers unfamiliar with the other two sections. However, readers of the second section will have to take our word for some of its conclusions, but the reasoning behind the conclusions will be clearer to those who have read the first section.

Canonical Variate Analysis (CVA)

What CVA Does

The questions answered by CVA have rarely been stated in the form which, in our opinion, is most meaningful and useful to behavioral scientists. This form is:

1. What is the minimum number of traits that would have to be controlled or partialled out in order to eliminate all important linear relations between sets X and Y—that is, in order to change all the elements of R_{xy} to zero or near zero? In other words, how many traits are needed to explain the relationships between sets X and Y?
2. What is the nature of those traits?

As will be seen later, the worker has complete freedom in choosing the degree of relationship he considers "important"; he may consider any correlation at all important, or he may be interested only in fairly large relationships.

The analysis is performed in the same way regardless of whether the worker hypothesizes that the traits that must be controlled are included in the X or Y sets, or are weighted averages of variables in one or both sets, or are hypothetical traits not adequately represented by any of the variables included in the X or Y sets or by weighted averages of those variables.

In the last case, for example, suppose a theory predicts that there should be no correlation between two sets of observed variables when a unidimensional hypothetical trait called "trait anxiety" is held constant. CVA could lead a worker to reject this hypothesis even though there is not even a moderately satisfactory measure of "trait anxiety." CVA can do this by demonstrating that at least k different traits must be controlled in order to eliminate all relations between sets X and Y. In the present example, if k turns out to be greater than one, then the hypothesis that "trait anxiety" is the *only* trait which must be controlled is rejected immediately, even though further analysis may reveal that trait anxiety is *one* of the traits which must be controlled.

Hypotheses of the sort tested by CVA arise in all eight of the problem areas identified in the opening paragraph of this paper. This is illustrated by the following eight hypotheses, all of which can be tested by CVA. Each hypothesis is from a different one of the eight problem areas. The value of k shown in parentheses at the end of each example gives the number of traits which the hypothesis specifies must be controlled in order to eliminate all relationship between the two sets of variables. The examples follow:

1a. The subtests of the Wechsler IQ are not useful for predicting more than a single dimension of school grades. That is, a single weighted average of school grades can be predicted, but different combinations of Wechsler scales are not useful in predicting grades in individual courses more accurately than they can be predicted from the single weighted average. (k = 1)

1b. Test batteries X and Y have at most three dimensions in common. When those dimensions are controlled, the scales of the two batteries are independent. ($k \leqslant 3$)

2a. Several subscales measuring socioeconomic status (e.g., father's occupation, income, mother's education) have the same pattern of correlations with each subscale of an IQ test that they have with the overall IQ. (k = 1)

2b. There is no relation between the scales of an interest battery and grades in different school subjects when some measure of overall academic achievement is held constant. ($k \leqslant 1$)

3a. A battery of personality factors measured at age 10, and another battery measured at age 25, have at most three dimensions in common. ($k \leqslant 3$)

3b. There is no relation between the scales of a personality test administered in a supportive setting and the same test administered in a competitive atmosphere. (k = 0)

4a. Two dimensions of students' mean perceptions of the social environment of their classes obtained at mid-year are related to mean class achievements and attitudes in the subject-matter area measured at the end of the school year. (k = 2)

4b. After pairs of randomly-matched subjects have listened to a tape recording on race relations, and discussed it with their partners, there is no relation between the attitudes of the two partners on a variety of topics concerning race relations. (k = 0)

Of these eight examples, 4a perhaps represents the type of problem of greatest interest to educators. One of the authors and his colleagues have used CVA extensively for problems of this type: Anderson and Walberg (1968); Rothman, Welch, and Walberg (1969); Walberg (1969, 1971); Walberg and Ahlgren (1970); Walberg and Rothman (1969); Walberg, Welch, and Rothman (1969).

One reason that the uses of CVA have been so widely misunderstood is that there are several different ways of stating a hypothesis about the value of k, in addition to the one given in the opening paragraph of this section. These alternative ways may sound very different, but are in fact mathematically equivalent to the formulation stated above.

The most important of these alternative formulations concerns the properties of linear composites of the observable X and Y variables. If exactly k traits must be controlled to eliminate all relations between sets X and Y, then it is possible to find a set of k linear composites of the variables in set X that completely account for the relationship of set X to set Y. More specifically, these k composites will have the following properties:

1. The multiple regression prediction of any Y variable from the X variables is a linear function of the k composites.
2. When the k composites are held constant, there is no correlation between any variable in set X and any variable in set Y.
3. Any weighted average of the variables in set X which is uncorrelated with the k composites will be uncorrelated with all of the variables in set Y.

Since CVA is completely symmetric between sets X and Y, it will also be possible to find k linear composites of the variables in set Y with analogous properties when the roles of sets X and Y are reversed.

Notice that we have not specified that the k composites must be mutually uncorrelated, though they may always be rotated to be so. As will be noted later, CVA in fact derives the composites in such a way that they are mutually uncorrelated, though it could be argued that this is not always a centrally important property of the method.

From item 2 on the above list, it follows that if a set of k *unobserved* traits have the property hypothesized (that is, if partialling on them will remove all association between sets X and Y), then a set of *observable* linear composites of set X (or of set Y) will have the same property. This is the reason why CVA proceeds in exactly the same way regardless of whether the traits hypothesized to explain the relations between two sets are observed or unobserved.

Still another way of stating the type of hypothesis tested by CVA is that matrix R_{xy} will have rank k. This is the form of the hypothesis used in the example in Table 1; in that example, R_{xy} has rank 1. Since not all readers will know what this means, we include it here primarily to show that the hypotheses under discussion are hypotheses about the *between*-set matrix R_{xy}, and are thus completely different from the hypotheses discussed in the last section of this paper concerning the similarity of the *within*-set matrices R_{xx} and R_{yy}.[2]

Readers familiar with multiple regression, multiple discriminant analysis, and ordinary and multivariate fixed-effects analysis of variance can understand the purposes of CVA by relating it to these techniques. All these techniques are in fact special cases of CVA; all of them can be performed with a CVA computer program.[3] Consider the following array:

Characteristics of methods	Only one dependent variable	One or more dependent variables
Only nominal independent variables	Univariate analysis of variance	Multivariate analysis of variance
Nominal or interval independent variables	Multiple regression	CVA

The techniques on the right are more general than those on the left, since those on the right can consider more than one dependent variable simultaneously. Those on the bottom differ from those on the top in that, on the bottom, the independent variables are usually continuous variables rather than categorical (nominal) variables. The methods on

[2] Another reason for including the present formulation of the hypothesis is to show that the hypotheses tested by CVA are in fact identical to the hypotheses tested by Tucker's (1958) inter-battery factor analysis. The well-studied significance tests used in connection with CVA are thus preferable to the approximate significance test suggested by Tucker for testing hypotheses about the rank of R_{xy}, since the hypotheses tested by the two tests are equivalent.

The equivalence is proved as follows. Statement (1) above states that the matrix B of regression weights for predicting Y from X has rank k. $B = (X'X)^{-1} X'Y$, where X and Y are the data matrices. Since no matrix product can have a larger rank than either of the constituent matrices, the rank of B equals the smaller of the ranks of $(X'X)^{-1}$ or $(X'Y)$. The former matrix has the same rank as $X'X$, which in turn has the rank of X, which by the first rule is at least as large as the rank of $X'Y$, which has the same rank as R_{xy}. Thus B has a rank no larger than that of R_{xy}. Then applying these same rules to the equation $B(X'X) = X'Y$ proves that the rank of B cannot be smaller than that of R_{xy}, so the two ranks are equal.

[3] This point is included for its theoretical interest. It is not recommended in practice for two reasons. First, certain computational simplifications are possible in programs written specifically for the less comprehensive methods, so that these programs will run faster than a CVA program. Second, at most computer installations the "canned" CVA programs do not print out as many auxiliary statistics as the simpler programs.

the bottom are more general, since they can perform the same analyses as those on the top through the use of dummy variables (Cohen, 1968; Walberg, 1971). Thus CVA is the most general technique in the array, encompassing all the others as special cases.

Through the use of dummy variables on the dependent side, CVA can also be used to perform multiple discriminant analyses (in which the independent variables are continuous) or simple contingency table analyses (through the use of dummy variables on the independent side as well). In the latter case, N times the sum of the squared canonical correlations (which will be defined later) equals the value of chi-square in the elementary test for independence in a two-way frequency table (Kendall & Stuart, 1961, p. 574). The parametric significance test given later may generally be used if the sample size is over 50 and the number of variables is reasonably small (Hsu & Feldt, 1969). Thus CVA is the only one of the techniques in which either set X or set Y may contain one or more variables, and the variables in either set may be continuous, categorical, or mixed.

Mechanics of CVA

This section describes the statistics computed by a typical CVA computer program. Unlike the previous section, readers familiar with previous accounts of CVA should find this section wholly familiar, and should skim or skip it. Those reading it will be struck by the lack of obvious relation between the statistics computed in CVA and the purposes of CVA described in the previous section. This gap will be bridged in the next section.

More detailed descriptions of the computational aspects of CVA can be found in Hotelling's original papers (1935, 1936); early papers by Thomson (1947), Bartlett (1948), Rao and Slater (1949), and Lubin (1950); more recent papers by Rozeboom (1965), Bock (1966), McKeon (1966), Cramer and Bock (1966), Porebski (1966b), McDonald (1968), Walberg (1971); and textbook treatments by Gulliksen (1950), Anderson (1958), Cooley and Lohnes (1962, 1971), Rozeboom (1966), Morrison (1967), Bock (1971), Tatsuoka (1971), and Van de Geer (1971). Bock (1963) and Cooley and Lohnes (1971) have published respectively computer flowcharts and programs.

The statistics computed in CVA are the *canonical correlations* and *canonical weights*, which are computed from the correlation matrix R described above. The first canonical correlation is the highest correlation that can be found between a weighted composite of X variables and a weighted composite of Y variables. Those composites are the *first canonical variates*, and the weights forming them are the *first canonical weights*, which are the elements of the *first canonical vectors*. The *second canonical correlation* is the highest correlation that can be found between X and Y weighted composites which are uncorrelated with the first canonical variates. These are the *second canonical variates*. Third, fourth, and subsequent canonical correlations and pairs of canonical variates are defined similarly. The number of nonzero canonical correlations is termed the number of *canonical relations* between sets X and Y. It cannot exceed the number of variables in the smaller set.

The most widely used significance test on the number of canonical relations is Bartlett's (1938) chi-square approximation to the distribution of Wilk's lambda. Bartlett's test of the null hypothesis that there are no more than k canonical relations[4] is:

$$(1) \quad [N - .5(n_x + n_y + 1)] \ ln \prod_{k+1}^{s} [1/(1 - r_j^2)] = \chi^2_{(n_x - k)(n_y - k)}$$

where r_j is the jth canonical correlation and s is the minimum of n_x and n_y. Schatzoff (1966) has published tables of coefficients for correcting the values of chi-square computed by this formula, but the coefficients are so close to unity for values of N above 50 that Bartlett's test can be regarded as highly accurate for these sample sizes. Furthermore, the Bartlett test is always on the conservative side. A quick computational check in the Bartlett formula is provided by the fact that, when the r_j's are .5 or less, the left side of the formula roughly equals N times the sum of the squared canonical correlations entering the formula.

The most parsimonious and definite hypothesis is that no canonical relations exist between X and Y. The next most parsimonious hypothesis is that exactly one canonical relation exists, the next that exactly two exist, and so on. The usual procedure, then, is first to test the hypothesis that no canonical relations exist. If it is rejected, the worker proceeds to test the hypothesis that exactly one exists, and so on. Successive hypotheses will always be rejected with less and less confidence, until finally some hypothesis will not be rejected. If that hypothesis states that there are exactly k relationships, then that is the most parsimonious hypothesis consistent with the data.

How to Use CVA to Determine the Number of Relations

We first introduced CVA with the statement that its major use is to determine the *number* and *nature* of the traits that must be controlled in order to eliminate all important relations between two sets of variables. From the mechanical description of CVA in the last section, it is not obvious how this can be done. The present section describes how to determine the number of traits, and the next describes how to infer their nature.

The basic rule for determining the number of traits that must be controlled is simple: it equals the number of canonical relations

[4] In this formula, the expression $\Pi[1/1 - r_j^2)]$ is the reciprocal of Wilk's lambda (λ); and $\lambda^{N/2}$ is the likelihood ratio.

between sets. For readers who want to understand why this is so, the following explanation is given:

1. Suppose there is a single trait u which, if controlled, will eliminate all relations between sets X and Y.
2. Then if x denotes any linear composite of the variables in set X, and y denotes any linear composite of the variables in set Y, we will have $r_{xy \cdot u} = 0$.
3. From the numerator of the formula for a partial correlation, (2) implies $r_{xy} = r_{ux} r_{uy}$.
4. From (3) it follows that if r_{uy} is fixed, r_{ux} is proportional to r_{xy}.
5. Thus for any fixed linear composite y, the linear composite x correlating maximally with y will also correlate maximally with u. Similarly, for any x, the y correlating maximally with it will also correlate maximally with u.
6. Thus if x and y are derived to correlate maximally with each other, both composites will be the best possible within-set predictions of u.
7. Thus any composites uncorrelated with x and y will also be uncorrelated with u; and since u provides the only link between the two sets, any such composite in set X must be uncorrelated with any such composite in set Y.

The explanation in the last paragraph applies to either a population or a sample. It is of most interest when applied to a population. That is, it is of most interest when it is used to show that the number of canonical relations in the population equals the number of traits that must be controlled in order to eliminate all correlations between sets X and Y in a population. The parallel conclusion concerning a sample is of little interest, since sampling error will generally mean that no correlations will be exactly zero in the sample even if they are zero in the population. Thus the number of canonical relations observed in a sample will generally equal the number of variables in the smaller set—that is, the smaller of n_x or n_y—even if the number of canonical relations in the population is much smaller. The minimum number of canonical relations known to exist in a population can be determined by the significance test described earlier. In the following discussion, questions about the number of canonical relations between two sets will be assumed to be questions about a population rather than a sample.

From the conclusions stated above, the basic links between the mechanics and the purposes of CVA can be drawn. If only one trait u needs to be controlled in order to eliminate all relationships between sets X and Y, then:

1. The second and subsequent canonical correlations will be zero.
2. The first canonical correlation will equal the product of the correlations of u with the first canonical variates in sets X and

Y. Though these two correlations are not known, the first population canonical correlation provides a lower bound for either one, since it equals their product and the product cannot be greater than either of the two entries when both entries are unity or less. Since the canonical correlation observed in a small sample will usually overestimate the population correlation, the sample correlation is not a lower bound.

3. The first canonical variates are the best within-set estimates of u.

Similar conclusions apply if the number of canonical relations is greater than one.

The smaller a canonical correlation, the less important is the statistical association it represents. Thus a worker may choose to ignore canonical correlations below a certain size, even if they are statistically significant.

Interpreting the Nature of Canonical Relations

Once a worker has determined the number of traits which must be controlled in order to remove all important relationships between two sets of variables, how does he infer the nature of those traits?

One school of thought (Dunteman & Bailey, 1967; Kahn, 1969; Ohnmacht & Olson, 1968; Thorndike, Weiss, & Dawis, 1968b; Vestre & Lorei, 1967; and others) has used canonical weights for these interpretations, while a much smaller group (notably Porebski, 1966a) has used correlations between the canonical variates and the original variables. The two sets of statistics can be very different—some variables can have high weights and low correlations, while other variables can have low weights and high correlations.

The theoretical advantages of the two types of statistic have not been adequately explicated. A detailed analysis would probably show that correlations are theoretically preferable in some situations and weights in others. However, in most cases the choice between the two is dictated by a practical rather than a theoretical consideration: sampling error. By analogy to the situation in multiple regression (Darlington, 1968, pp. 175-177), it can be inferred that the standard errors of weights are often much higher than those of correlations. This is especially true precisely in those cases when the differences between weights and correlations are greatest—when variables within a set are highly intercorrelated. Thus in small or medium-sized samples, any large divergence between the relative sizes of weights and those of correlations must throw special suspicion on the weights. In such samples, then, the worker should emphasize the correlations.

If not available from a CVA program, correlations between canonical variates and original variables can be computed readily by using a standard correlation program, since most such programs allow the user to define new variables as weighted averages of other variables.

If the original variables are not standardized, then the standardized weights computed by a CVA program must be divided by the standard deviations of the variables involved to obtain the weights to be applied to the unstandardized original variables. A useful computational check in the output of a correlation program is provided by the correlations among canonical variates—all should be zero except those between paired canonical variates, which should equal the canonical correlations.

The *relative* sizes of the correlations of a canonical variate with the variables in that set equal the relative sizes of the correlations of those same variables with the trait predicted by the canonical variate. The correlation of each variable with the trait equals the observed correlation of that variable with the canonical variate, times the correlation between the trait and the canonical variate. As described earlier, the canonical correlation provides a lower bound for this last correlation. These theorems follow by direct extension from multiple regression; their uses and derivation are discussed by Darlington (1968, pp. 170-171; 1970). From them it follows that a worker can read from the observed correlations between original variables in a set and a canonical variate in that set the relative sizes of the correlations of those variables with the unobserved trait which the canonical variate predicts.

The canonical correlation tells how accurately the nature of the trait can be inferred by these means; the higher the correlation, the more accurately the observed canonical variates represent the unobserved trait.

Devices for Facilitating the Interpretation of Canonical Variates

Since CVA derives canonical variates in a way which assures they will be uncorrelated, use of CVA in the manner described above implicitly assumes that the explanatory traits are mutually uncorrelated. This assumption does not affect the determination of the number of traits, but it does affect the determination of their nature. Assuming the traits to be mutually intercorrelated may aid in their interpretation, just as oblique rotations in factor analysis may lead to a simpler interpretation of factors.

Once canonical variates have been derived, the variates in either set may be interpreted as factors, and rotated like any other factors without regard to the other set. This is perhaps best done if the worker thinks of one set of variables as more basic than the other set; then the rotation can be done in the more basic set. This is especially useful if he thinks of the explanatory traits as actual observable weighted averages of the variables in that set, rather than as unobservable traits. A further discussion of the rotation of canonical variates has been given by Hall (1969).

A second method for facilitating the interpretation of canonical variates, called step-down analysis, was originated by Roy (1958). An excellent discussion and illustration of the method has been given by

Bock and Haggard (1968). Instead of deriving canonical variates which are then partialled out of the association between the two sets, the worker selects single variables from one set or the other which are then partialled out. Though the association between sets might not be fully explained by quite as small a number of variates as in CVA, the problem of interpreting the variates nearly vanishes since each pair of variates includes one which is simply one of the original X or Y variables.

The third method, called multivariate stepwise regression analysis, is also discussed by Bock and Haggard. The method is similar to step-down analysis, the difference being that successive variables are selected mechanically to maximize a measure of association between sets X and Y, while in step-down analysis the variables are selected theoretically or subjectively, preferably prior to the analysis.

Either of these two methods may be subject to the problem of collinearity. That is, if variables within a set are highly correlated, then controlling for any of the variables may substantially reduce both the within-set and between-set correlations among the remaining variables. Thus the order of selection of variates in multivariate stepwise regression become highly subject to chance variations, whereas in step-down analysis an unwary investigator is likely to conclude that the only important variable is whichever one he happened to choose first.

Extensions and Variations of CVA

CVA for large numbers of variables. Weinberg (1971) has described a new method for estimating canonical correlations and canonical vectors. Though the new method is inefficient for very large samples, evidence suggests that when the ratio $(n_x + n_y)/N$ is greater than about 1/10, canonical correlations and canonical weights are estimated more accurately by the new method than by standard CVA computer programs. In both methods the true canonical correlations tend to be somewhat overestimated, making cross-validation desirable.

Relations among m sets of variables. Horst (1961a, 1961b) and Carroll (1968) have discussed extensions of CVA to more than two sets of variables. Extensions of the uses of CVA discussed above can be imagined which fit these models—for example, instead of studying relations just between teachers and students, a worker could study relations among teachers, students, and principals. In general, though, as more complex statistical techniques are used, the problem of interpreting the results in meaningful terms becomes greater. Since the problem is already great enough in ordinary (two-set) CVA, the usefulness of CVA when extended to m sets remains to be seen. Also, significance tests do not yet exist for the generalized canonical correlations computed by these procedures.

CVA for fallible measures. Meredith (1964) has given a set of formulas which can be used to study the canonical relations among the "true" variables underlying a set of fallible measures. However, the same results can be obtained more simply by applying ordinary correction-for-attenuation formulas to the observed correlations among a set of measures and then applying an ordinary CVA program to the corrected correlations.

Canonical factor analysis. Rao (1955) showed that some important unsolved problems in pure factor analysis could be solved by an ingenious application of CVA. Rao's method is usually called "canonical factor analysis"; it should not be thought of as a typical application of CVA in the behavioral sciences, but rather as an important milestone in factor theory.

CVA with matched sets of variables. Another variant of CVA has its principal applications in two rather different problems: constructing batteries with maximum test-retest reliability, and studying the dimensions of similarity between matched subjects, such as twins. In both these problems, sets *X* and *Y* consist of the same variables, given either twice to the same subject or once to each of paired subjects. Modifications of CVA for this problem are discussed by DeGroot and Li (1966), Green (1950), Peel (1948), and Thomson (1940). Ordinary CVA was applied by Cooley (1967) to a similar problem, with interesting results.

Measures of Similarity Between Two Sets

Questions about the "similarity" or "overlap" or "redundancy" between two sets of variables have some connection to the hypotheses tested by CVA, since the hypothesis of zero similarity is the hypothesis that there are no canonical relations. However, the connection between the two types of question ends there.

Questions about the similarity of two sets of variables can arise in any of the eight problem areas identified in the opening paragraph of this paper. For example:

1a. How valid is test battery *X* for predicting a set of criterion variables *Y*?
1b. How redundant are test batteries *X* and *Y*? Should one be dropped?
2a. How much of the variance in a set of dependent variables *Y* is determined by a set of school achievement variables *X*?
2b. How much overlap is there between a set of personality variables *X* and a set of school achievement variables *Y*?
3a. How much overlap is there between a set of measures of intellectual functioning in fifth grade and a set in seventh grade?
3b. To what degree is a set of personality measures taken under stress related to the same set of measures taken in a non-stress situation?
4a. To what degree are the personality characteristics of wives predictable from those of their husbands?

4b. To what degree are the responses of experimental subjects predictable from the responses of the experimenters with which they were paired?

Previous Measures of Similarity

Previous writers have suggested four different measures of the "similarity" between two sets of variables, all based on the sizes of the canonical correlations between the sets. This subsection criticizes all four. The next suggests several new methods of analysis for answering questions about the similarity between two sets of variables. No single measure of similarity applies to all situations.

Lohnes and Marshall (1965), Rentz, Fears, and White (1968), and Thorndike, Weiss, and Dawis (1968a) have suggested basing measures of similarity between sets of variables on the highest few canonical correlations between the sets. This may be misleading, since a few variables in each set that correlate very highly with variables in the other set could then lead a worker to conclude that the two sets are highly "redundant," even though the vast majority of the variables in either set may be wholly uncorrelated with any variables in the other set.

A measure which Rozeboom (1965) called the "correlation between two sets of variables," and which was interpreted as a measure of "redundancy" by Stewart and Love (1968), has the same property; it can be very high, even unity, even though the vast majority of the variables in either set may correlate zero with all variables in the other set.

Stewart and Love proposed measuring similarity by the mean of the squared multiple correlations for predicting the variables in one set from those in the other. As they point out, this measure is asymmetric between sets. Miller and Farr (1971) have shown that it can be decomposed into additive components contributed by mutally orthogonal factors. The Stewart and Love measure might be a useful measure of similarity in some situations. However, use of this or any other single measure of the "similarity" of two sets may often be misleading. For example, if a test battery "overlaps strongly" with a set of criterion variables, it may still be true that the one criterion variable that is least well predicted from the battery happens to be the most important one. Or, if one test battery Y is highly "redundant" with another battery X because every Y variable can be accurately (though not perfectly) predicted from battery X, it may still happen that some composite, say $Y_1 - Y_2$, is very useful in predicting some external criterion variable which is predicted very poorly from X. For example, suppose that set Y contains two variables and they correlate very highly and equally with X_1. Then the Stewart-Love measure would be very high, even though $Y_1 - Y_2$ (where Y_1 and Y_2 are standardized) would correlate exactly zero with X_1.

If variables in one set (say set Y) are mutually uncorrelated, the sum of squared canonical correlations can be termed "the amount of variance in set Y accounted for by set X," since it equals the sum of the n_y squared multiple correlations predicting variables in set Y from those in set X. If variables in both sets are uncorrelated within sets, then the sum of squared canonical correlations can be termed "the amount of variance in either set accounted for by the other." In this case it also equals the sum of the $n_x \times n_y$ squared simple correlations between X and Y. The sum of squared canonical correlations is thus a useful measure of similarity between two sets of factors which are uncorrelated within sets, or which could be rotated to be so. This measure was proposed by Wrigley and Neuhaus (1955) for measuring similarity between two sets of factors. However, in the more general case, in which variables in both sets are correlated within sets, these interpretations do not apply. In an extreme case, it could happen that all but one of the canonical correlations are perfect, even though none of the variables in either set is well predicted from the other set.[5] The sum of squared canonical correlations should thus not be used routinely for practical problems like assessing the redundancy between two test batteries.

Some New Methods for Measuring Similarity

A far more conservative, yet perhaps more useful, measure of the "redundancy" between two test batteries would be the *smallest* canonical correlation between the two. The last canonical variate in the smaller set is the weighted composite in that set *least* predictable from the larger set, and the smallest canonical correlation is the multiple correlation between that composite and the larger set. Only in the unlikely case in which this figure is high can it be said that every possible weighted average in the smaller set can be predicted accurately from the larger set. (If R is nonsingular and the sets contain different numbers of variables, then there is always some weighted composite in the larger set completely unpredictable from the smaller set.)

One drawback in the measure just proposed is that the multiple correlations for predicting variables in one set may be low due to unreliability of the variables in that set. An alternate method of assessing redundancy would be to compute the variances and correlations of the portions of the variables in one set orthogonal or residual to the other set.[6] If these variances (corrected for shrinkage, by Formula 11, Darlington, 1968, p. 172) roughly equal the amounts of

---

[5] This could happen if the first principal component of set Y correlates zero with all X variables, yet accounts for almost all the variance in set Y. Yet all subsequent principal components of set Y could in theory be perfectly predictable from set X, making all canonical correlations perfect except one.

[6] The variance-covariance matrix of the orthogonal components of X is $C_{xx} - C_{xy}C_{yy}^{-1}C_{yx}$.

unreliable variance estimated to be in the original variables, then the second set predicts all the reliable variance in the first set. Unlike the previous analysis, this procedure makes sense in either direction, since the smaller set may account for all reliable variance in the larger.

If the reliabilities of variables in the first set are unknown, then the correlations between the orthogonal components may provide useful information, since any high correlation inplies that the two residual variables involved cannot be measuring pure error. If all the residual correlations are low, this corroborates (but does not prove) the hypothesis that the residuals measure only measurement error and that the first battery is redundant with the second.

Sometimes questions about the "similarity" of two sets are answered most fruitfully by inspection of simple correlations between specific variables in the two sets. For example, O'Hara and Tiedeman (1959) used CVA to study the relation between a set of self-report measures and a set of objective-test measures of interests, aptitudes, values, and social class. They found that the first canonical correlation between self-report and objective-test measures was higher for a sample of older students than for younger students. From this they concluded that self-report becomes more accurate (that is, more similar to the results of objective tests) with increasing age. However, from their paper it is impossible to determine whether the increased correlations were between the self-reported and objective measures of the same variables, or whether the increase might be between unpaired variables, for example, between an objective-test measure of social class and a self-report measure of aptitudes. A comparison of individual correlations would seem to be more applicable to this problem than CVA.

Factor Methods of Measuring Similarity

One type of similarity measure which may be useful for some purposes is the degree to which a single set of factors will fit the two sets of variables. Contrary to Burt (1948), CVA is not useful for deriving such a measure. For an extreme example of this point, imagine that there are k orthogonal factors in each set, in declining order of importance, and that each of the *first* $k/2$ factors in each set correlates perfectly with one of the *last* $k/2$ factors in the other set. Then all canonical correlations would be perfect even though all of the first $k/2$ factors in set X would correlate zero with all of the first $k/2$ factors in set Y, so that a single set of factors would not fit the two sets well.

If the factors in question are principal components (PCs), then the most direct method for studying the degree to which a single set of PCs will fit both sets of variables is to perform a PC analysis on the combined set of X and Y variables and perform separate PC analyses on the individual X and Y sets. If (and only if) the first PCs in X and Y are identical, then the amount of variance (measured by the eigenvalue) accounted for by the first PC of the combined sets will equal the sum of the amounts of variance accounted for by the first PCs of the individual sets. The same is true for the second, third, and remaining PCs. If the two sets of PCs are not identical, then the amount of variance accounted for by the PCs of the combined sets will always be less than the sum of the amounts of variance accounted for by the PCs derived from the individual sets. Thus, the difference between these two amounts serves as a useful measure of the degree to which a single set of PCs will fit the two sets of variables. This difference can be computed separately for the first, second, third, and successive PCs. Alternatively, the worker can compute the ratio rather than the difference between the two amounts of variance, enabling him to make statements such as, "The first m PCs of the combined sets account for 95% as much variance as the first m PCs of the individual sets."

When the factors in question are hypothetical constructs, as in pure factor analysis, the general logic of the solution is the same as in the last paragraph. The procedure is altered, however, to conform to the different definition of goodness of fit used in this case—size of residual correlations rather than amount of variance accounted for. Analogous to the procedure described in the last paragraph, the procedure is to perform a factor analysis on the combined sets of variables, plus two separate factor analyses within the individual sets. As before, this procedure may be performed for any specified number of factors m, or separately for each possible number.

Equality of Two Correlation or Covariance Matrices

The hypotheses to be discussed in this section have been confused with those of the previous section, notably by Horst (1961b, p. 129), though they are in fact very different. Again drawing illustrations from the list in the opening paragraph, we can list some sample hypotheses, though more selectively this time:

1a. Are the intercorrelations among the tests of a test battery the same as those among the criterion variables the test is supposed to predict?
1b. Are the intercorrelations among the tests in battery X the same as among the tests in battery Y?
3a. Are the intercorrelations among a set of measures the same at age 8 as at age 12?
3b. Does an experimental manipulation have an effect not only on the means of several dependent variables, but also on their standard deviations and intercorrelations?
4a. Are the intercorrelations among a set of traits measured on husbands the same as among their wives?

The hypothesis of equal correlation or covariance matrices differs from the hypotheses of the last section in that the *matrices*, rather than the *variables*, are hypothesized to be similar. For example, the two within-set matrices R_{xx} and R_{yy} may be identical even though the between-set correlations in R_{xy} are all zero. The correlation matrices are then similar though the variables are not. The present problem is defined whether the two sets of variables are measured in the same sample or in two independent samples, whereas the problems of previous sections have meaning only in terms of a single sample.

A test for the equality of two covariance matrices C_{xx} and C_{yy} (which implies equality of both variances and correlations) has been described by Anderson (1958, pp. 256-259) for the case in which the two sets of variables are measured in separate, unmatched samples. When there is a single sample of subjects, as in repeated-measures designs, a simple test of the same hypothesis exists if R_{xy} can be assumed to be symmetric. This assumption implies, for example, that if a battery of tests is administered twice, in years X and Y, then the correlation between, say, test 1 in year X and test 2 in year Y equals the correlation between test 1 in year Y and test 2 in year X. If this assumption holds, then it can be shown that the hypothesis $C_{xx} = C_{yy}$ is equivalent to the hypothesis that all the differences between matched X and Y variables correlate zero with all the sums. That is, the variables $X_1 - Y_1$, $X_2 - Y_2$, etc., all correlate zero with all the variables $X_1 + Y_1$, $X_2 + Y_2$, etc. This hypothesis is tested by a CVA with the sums in one set and the differences in the other.

References

Anderson, G. J., & Walberg, H. J. Classroom climate and group learning. *International Journal of Educational Sciences*, 1968, **2**, 175-180.

Anderson, T. W. *Introduction to multivariate statistical analysis.* New York: Wiley, 1958.

Bargmann, R. E., Multivariate statistical analysis in psychology and education. *Bulletin de l'Institut International de Statistique*, 1961, **38**, 79-86.

Bartlett, M. S. Further aspects of the theory of multiple regression. *Proceedings of the Cambridge Philosophical Society*, 1938, **34**, 33-40.

Bartlett, M. S. Internal and external factor analysis. *British Journal of Psychology*, Statistical Section, 1948, **1**, 73-81.

Bock, R. D. Programming univariate and multivariate analysis of variance. *Technometrics*, 1963, **5**, 95-117.

Bock, R. D. Contributions of multivariate experimental designs to educational research. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology*. Chicago: Rand McNally, 1966. Pp. 820-840.

Bock, R. D. *Multivariate statistical methods in behavioral research.* New York: McGraw-Hill, 1971.

Bock, R. D., & Haggard, E. A. The use of multivariate analysis in behavioral research. In Dean K. Whitla (Ed.), *Handbook of measurement and assessment in the behavioral sciences.* Reading, Mass.: Addison-Wesley, 1968. Pp. 100-142.

Burt, C. Factor analysis and canonical correlations. *British Journal of Psychology*, Statistical Section, 1948, **1**, 95-106.

Carroll, J. D. Generalization of canonical correlation analysis to three or more sets of variables. *Proceedings of the American Psychological Association*, 1968, **76**, 227-228.

Cohen, J. Multiple regression as a general data-analytic system. *Psychological Bulletin*, 1968, **70**, 426-443.

Cooley, W. W. Interactions among interests, abilities and career plans. *Journal of Applied Psychology*, 1967, **51** (Monogr. Suppl. 640).

Cooley, W. W., & Lohnes, P. R. *Multivariate procedures for the behavioral sciences.* New York: Wiley, 1962.

Cooley, W. W., & Lohnes, P. R. *Multivariate data analysis.* New York: Wiley, 1971.

Cramer, E. M., & Bock, R. D. Multivariate analysis. *Review of Educational Research,* 1966, **36**, 604-617.

Darlington, R. B. Multiple regression in psychological research and practice. *Psychological Bulletin,* 1968, **69**, 161-182.

Darlington, R. B. Some techniques for maximizing a test's validity when the criterion variable is unobserved. *Journal of Educational Measurement,* 1970, **7**, 1-14.

DeGroot, M. H., & Li, C. C. Correlations between similar sets of measurements. *Biometrics,* 1966, **22**, 781-790.

Dunteman, G. H., & Bailey, J. P. A canonical correlational analysis of the Strong Vocational Interest Blank and the Minnesota Multiphasic Personality Inventory for a female college population. *Educational and Psychological Measurement,* 1967, **27**, 631-642.

Green, B. F., Jr. A note on the calculation of weights for maximum battery reliability. *Psychometrika,* 1950, **15**, 57-61.

Gulliksen, H. *Theory of mental tests.* New York: Wiley, 1950.

Gulliksen, H. Methods for determining the equivalence of measures. *Psychological Bulletin,* 1968, **70**, 534-544.

Hall, C. E. Rotation of canonical variates in multivariate analysis of variance. *Journal of Experimental Education,* 1969, **38**, 31-38.

Harman, H. H. *Modern factor analysis.* (2nd ed.) Chicago: Univ. of Chicago, 1967.

Heck, D. L. Charts of some upper percentage points of the distribution of the largest characteristic root. *The Annals of Mathematical Statistics,* 1960, **31**, 625-642.

Hope, K. The study of hostility in the temperament of spouses: Definitions and methods. *British Journal of Mathematical and Statistical Psychology,* 1969, **22**, 67-95.

Horst, P. Generalized canonical correlations and their application to experimental data. *Journal of Clinical Psychology,* 1961, **17**, 331-347. (a)

Horst, P. Relations among *m* sets of measures. *Psychometrika,* 1961, **26**, 129-149. (b)

Hotelling, H. The most predictable criterion. *Journal of Educational Psychology,* 1935, **26**, 139-142.

Hotelling, H. Relations between two sets of variates. *Biometrika,* 1936, **28**, 321-377.

Hotelling, H. Multivariate quality control. In C. Eisenhart *et al.* (Eds.), *Selected techniques of statistical analysis.* New York: McGraw-Hill, 1947.

Hsu, T., & Feldt, L. S. The effect of limitations on the number of criterion score values on the significance level of the F-test. *American Educational Research Journal,* 1969, **6**, 515-528.

Kahn, S. B. Affective correlates of academic achievement. *Journal of Educational Psychology,* 1969, **60**, 216-221.

Kendall, M. G., & Stuart, A. *The advanced theory of statistics.* Vol. 2. London: Griffin, 1961.

Lawley, D. N. A generalization of Fisher's z-test. *Biometrika,* 1938, **30**, 180-187.

Lohnes, P. R., & Marshall, T. O. Redundancy in student records. *American Educational Research Journal,* 1965, **2**, 19-23.

Lubin, A. Linear and non-linear discriminating functions. *British Journal of Psychology,* Statistical Section, 1950, **3**, 90-103.

Massey, W. F., Frank, R. E., & Lodahl, T. *Purchasing behavior and personal attributes.* Philadelphia: Univ. of Pennsylvania, 1968.

McDonald, R. P. A unified treatment of the weighting problem. *Psychometrika,* 1968, **38**, 351-381.

McKeon, J. J. Canonical analysis: some relations between canonical correlation, factor analysis, discriminant function analysis, and scaling theory. *Psychometric Monographs,* 1966, No. 13.

Meredith, W. Canonical correlations with fallible data. *Psychometrika*, 1964, 29, 55-65.

Miller, J. K., & Farr, S. D. Bimultivariate redundancy: A comprehensive measure of interbattery relationship. *Multivariate Behavioral Research*, 1971, 6, 313-324.

Morrison, D. F. *Multivariate statistical methods.* New York: McGraw-Hill, 1967.

Mukherjee, B. N. Application of canonical correlational analysis to learning data. *Psychological Bulletin*, 1966, 66, 9-21.

O'Hara, R. P., & Tiedeman, D. Vocational self-concept in adolescence. *Journal of Counseling Psychology*, 1959, 6, 292-301.

Ohnmacht, F. W., & Olson, A. V. Canonical analysis of reading readiness measures and the Frostig developmental test of visual perception. *Educational and Psychological Measurement*, 1968, 28, 479-484.

Peel, E. A. Prediction of a complex criterion and battery reliability. *British Journal of Psychology*, Statistical Section, 1948, 1, 84-94.

Pillai, K. C. S. Some new test criteria in multivariate analysis. *Annals of Mathematical Statistics*, 1955, 26, 117-121.

Pillai, K. C. S. On the distribution of the largest characteristic root of a matrix in multivariate analysis. *Biometrika*, 1965, 52, 405-414.

Pillai, K. C. S. Upper percentage points of the largest root of a matrix in multivariate analysis. *Biometrika*, 1967, 54, 189-194.

Porebski, O. R. Discriminatory and canonical analysis of technical college data. *British Journal of Mathematical and Statistical Psychology*, 1966, 19, 215-236. (a)

Porebski, O. R. Interrelated nature of multivariate statistics in discriminatory analysis. *British Journal of Mathematical and Statistical Psychology*, 1966, 19, 197-214. (b)

Rao, C. R. Estimation and tests of significance in factor analysis. *Psychometrika*, 1955, 20, 93-111.

Rao, C. R., & Slater, P. Multivariate analysis applied to differences between neurotic groups. *British Journal of Psychology*, Statistical Section, 1949, 2, 17-29.

Rentz, R. R., Fears, E. B., & White, W. F. Personality correlates of group structure—a canonical correlational analysis. *Journal of Psychology*, 1968, 70, 163-167.

Rothman, A. I., Welch, W. W., & Walberg, H. J. Physics teacher characteristics and student learning. *Journal of research in science teaching*, 1969, 6, 59-63.

Roy, J. Step-down procedures in multivariate analysis. *Annals of Mathematical Statistics*, 1958, 29, 1177-1187.

Rozeboom, W. W. Linear correlations between sets of variables. *Psychometrika*, 1965, 30, 57-71.

Rozeboom, W. W. *Foundations of the theory of prediction.* Homewood, Illinois: Dorsey, 1966.

Schatzoff, M. Exact distributions of Wilk's likelihood ratio criterion. *Biometrika*, 1966, 53, 347-358.

Stewart, D., & Love, W. A general canonical correlation index. *Psychological Bulletin*, 1968, 70, 160-163.

Tatsuoka, M. M. *Multivariate analysis: Techniques for educational and psychological research.* New York: Wiley, 1971.

Thomson, G. H. Weighting for battery reliability and prediction. *British Journal of Psychology*, 1940, 30, 357-366.

Thomson, G. H. The maximum correlation of two weighted batteries. *British Journal of Psychology*, Statistical Section, 1947, 1, 27-34.

Thorndike, R. M., & Weiss, D. J. Stability of canonical components. *Proceedings of the American Psychological Association*, 1970, 78, 107-108.

Thorndike, R. M., Weiss, D. J., & Dawis, R. V. Canonical correlation of vocational interests and vocational needs. *Journal of Counseling Psychology*, 1968, 15, 101-106. (a)

Thorndike, R. M., Weiss, D. J., & Dawis, R. V. Multivariate relationships between a measure of vocational interests and a measure of vocational needs. *Journal of Applied Psychology*, 1968, **52**, 491-496. (b)

Tucker, L. A method for synthesis of factor analysis studies. Personnel Research Section Report, No. 984. Washington, D.C.: Dept. of the Army, 1951.

Tucker, L. An inter-battery method of factor analysis. *Psychometrika*, 1958, **23**, 111-136.

Van de Geer, J. P. *Introduction to multivariate analysis for the social* sciences. San Francisco: W. H. Freeman, 1971.

Vestre, N. R., & Lorei, T. W. Relationships between social history factors and psychiatric symptoms. *Journal of Abnormal Psychology*, 1967, **72**, 247-250.

Walberg, H. J. Social environment as a mediator of classroom learning. *Journal of Educational Psychology*, 1969, **60**, 443-448.

Walberg, H. J. Generalized regression models in educational research. *American Educational Research Journal*, 1971, **8**, 71-91.

Walberg, H. J., & Ahlgren, A. Predictors of the social environment of learning. *American Educational Research Journal,* 1970, **7**, 153-167.

Walberg, H. J., & Rothman, A. Teacher achievement and student learning. *Science Education*, 1969, **53**, 253-257.

Walberg, H. J., Welch, W. W., & Rothman, A. I. Teacher heterosexuality and student learning. *Psychology in the Schools,* 1969, **6**, 258-266.

Wienberg, Sharon. Large variable canonical anslysis: A proposed variant of canonical variate anslysis. Unpublished doctoral dissertation, Cornell University, 1971.

Wrigley, C., & Neuhaus, J. E. The matching of two sets of factors. *American Psychologist*, 1955, **10**, 418-419.

LINKED CITATIONS

You have printed the following article:

Canonical Variate Analysis and Related Techniques
Richard B. Darlington; Sharon L. Weinberg; Herbert J. Walberg
Review of Educational Research, Vol. 43, No. 4. (Autumn, 1973), pp. 433-454.
Stable URL:
http://links.jstor.org/sici?sici=0034-6543%28197323%2943%3A4%3C433%3ACVAART%3E2.0.CO%3B2-Y

References

Programming Univariate and Multivariate Analysis of Variance
R. Darrell Bock
Technometrics, Vol. 5, No. 1. (Feb., 1963), pp. 95-117.
Stable URL:
http://links.jstor.org/sici?sici=0040-1706%28196302%295%3A1%3C95%3APUAMAO%3E2.0.CO%3B2-3

Multivariate Analysis
Elliot M. Cramer; R. Darrell Bock
Review of Educational Research, Vol. 36, No. 5, Methodology of Educational Research. (Dec., 1966), pp. 604-617.
Stable URL:
http://links.jstor.org/sici?sici=0034-6543%28196612%2936%3A5%3C604%3AMA%3E2.0.CO%3B2-K

Some Techniques for Maximizing a Test's Validity When the Criterion Variable Is Unobserved
Richard B. Darlington
Journal of Educational Measurement, Vol. 7, No. 1. (Spring, 1970), pp. 1-14.
Stable URL:
http://links.jstor.org/sici?sici=0022-0655%28197021%297%3A1%3C1%3ASTFMAT%3E2.0.CO%3B2-D

LINKED CITATIONS

Correlations between Similar Sets of Measurements
M. H. DeGroot; C. C. Li
Biometrics, Vol. 22, No. 4. (Dec., 1966), pp. 781-790.
Stable URL:
http://links.jstor.org/sici?sici=0006-341X%28196612%2922%3A4%3C781%3ACBSSOM%3E2.0.CO%3B2-J

Charts of Some Upper Percentage Points of the Distribution of the Largest Characteristic Root
D. L. Heck
The Annals of Mathematical Statistics, Vol. 31, No. 3. (Sep., 1960), pp. 625-642.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28196009%2931%3A3%3C625%3ACOSUPP%3E2.0.CO%3B2-Z

Relations Between Two Sets of Variates
Harold Hotelling
Biometrika, Vol. 28, No. 3/4. (Dec., 1936), pp. 321-377.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28193612%2928%3A3%2F4%3C321%3ARBTSOV%3E2.0.CO%3B2-F

The Effect of Limitations on the Number of Criterion Score Values on the Significance Level of the F-Test
Tse-Chi Hsu; Leonard S. Feldt
American Educational Research Journal, Vol. 6, No. 4. (Nov., 1969), pp. 515-527.
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28196911%296%3A4%3C515%3ATEOLOT%3E2.0.CO%3B2-3

A Generalization of Fisher's z Test
D. N. Lawley
Biometrika, Vol. 30, No. 1/2. (Jun., 1938), pp. 180-187.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28193806%2930%3A1%2F2%3C180%3AAGOFT%3E2.0.CO%3B2-7

Redundancy in Student Records
Paul R. Lohnes; Thomas O. Marshall
American Educational Research Journal, Vol. 2, No. 1. (Jan., 1965), pp. 19-23.
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28196501%292%3A1%3C19%3ARISR%3E2.0.CO%3B2-K

LINKED CITATIONS

Some New Test Criteria in Multivariate Analysis
K. C. S. Pillai
The Annals of Mathematical Statistics, Vol. 26, No. 1. (Mar., 1955), pp. 117-121.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195503%2926%3A1%3C117%3ASNTCIM%3E2.0.CO%3B2-R

On the Distribution of the Largest Characteristic Root of a Matrix in Multivariate Analysis
K. C. Sreedharan Pillai
Biometrika, Vol. 52, No. 3/4. (Dec., 1965), pp. 405-414.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196512%2952%3A3%2F4%3C405%3AOTDOTL%3E2.0.CO%3B2-F

Upper Percentage Points of the Largest Root of a Matrix in Multivariate Analysis
K. C. Sreedharan Pillai
Biometrika, Vol. 54, No. 1/2. (Jun., 1967), pp. 189-194.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196706%2954%3A1%2F2%3C189%3AUPPOTL%3E2.0.CO%3B2-6

Step-Down Procedure in Multivariate Analysis
J. Roy
The Annals of Mathematical Statistics, Vol. 29, No. 4. (Dec., 1958), pp. 1177-1187.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195812%2929%3A4%3C1177%3ASPIMA%3E2.0.CO%3B2-M

Exact Distributions of Wilks's Likelihood Ratio Criterion
Martin Schatzoff
Biometrika, Vol. 53, No. 3/4. (Dec., 1966), pp. 347-358.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28196612%2953%3A3%2F4%3C347%3AEDOWLR%3E2.0.CO%3B2-B

Generalized Regression Models in Educational Research
Herbert J. Walberg
American Educational Research Journal, Vol. 8, No. 1. (Jan., 1971), pp. 71-91.
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28197101%298%3A1%3C71%3AGRMIER%3E2.0.CO%3B2-%23

LINKED CITATIONS

Predictors of the Social Environment of Learning
Herbert J. Walberg; Andrew Ahlgren
American Educational Research Journal, Vol. 7, No. 2. (Mar., 1970), pp. 153-167.
Stable URL:
http://links.jstor.org/sici?sici=0002-8312%28197003%297%3A2%3C153%3APOTSEO%3E2.0.CO%3B2-F