DOCUMENT RESUME

ED 342 809 TM 017 987

AUTHOR	Taylor, Dianne L.
TITLE	Cross-Validation in Canonical Analysis.
PUB DATE	Jan 92
NOTE	24p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Houston, TX, January 31-February 2, 1992).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	College Faculty; Correlation; *Generalizability Theory; Heuristics; Higher Education; Instructional Leadership; *Multivariate Analysis; *Research Methodology; Statistical Studies; *Validity
IDENTIFIERS	*Cross Validation; *Invariance; Research Replication; Univariate Analysis

ABSTRACT

The need for using invariance procedures to establish the external validity or generalizability of statistical results has been well documented. Invariance analysis is a tool that can be used to establish confidence in the replicability of research findings. Several approaches to invariance analysis are available that are broadly applicable across univariate and multivariate procedures. This paper explains one of these procedures, cross-validation. One form of the technique, double cross-validation, is applied in a canonical correlation analysis using a heuristic data set. A double cross-validation of the weights in a canonical correlation analysis is used to test for invariance in a study of university leadership conducted by M. L. Tucker (1990) with 105 subjects. A brief overview of both invariance testing and canonical correlation analysis is provided. Four tables present data from the analysis, and a 27-item list of references is included. An appendix contains the computer command lines used to generate the cross-validation. (Author/SLD)

Cross-Validation in Canonical Analysis

Dianne L. Taylor
Tulane University

Running Head: Cross-validation

Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX, January, 1992.

ABSTRACT

The need for employing invariance procedures to establish the external validity or generalizability of statistical results has been well documented. Invariance analysis is a tool that can be used to establish confidence in the applicability of research findings. Several approaches to invariance analysis are available which are broadly applicable across univariate and multivariate procedures. The purpose of the present paper is to explain one of these procedures, cross-validation. A form of the technique, double cross-validation, is applied in a canonical correlation analysis using a heuristic data set. A brief overview of both invariance testing and canonical correlation analysis is provided.

The need for employing invariance procedures to the establish the external validity or generalizability of statistical results has been well documented (e.g., Carver, 1978; Cooil, Winer, & Rados, 1987; Crask & Perreault, 1977; Daniel, 1989; Fish, 1986; Taylor, 1991; Thompson, 1989; Thorndike, 1978). Nevertheless, while the results of internal validity studies are commonly included in published research, one rarely sees the results of external validity analyses reported. It is possible that this circumstance arises from a mistaken belief that statistical significance tests inform the researcher regarding the likelihood that results will replicate (Thompson, 1992).

Statistical significance does not confirm the generalizability of study results, nor is statistical significance more important than result generalizability. The most important research findings are those which are both statistically significant and replicable (Tukey, 1969). Some hard sciences seem to have come to grips with this axiom. Social sciences, such as education, pay "little attention to this principle" (Huck, Cormier, & Bounds, 1974, p. 369).

Invariance analysis is a tool that can be used to establish confidence in the replicability of research findings. Several approaches to invariance analysis are available which are broadly applicable across univariate and multivariate procedures. Included among these approaches are the jackknife (Crask & Perreault, 1977; Daniel, 1989; Taylor, 1991), bootstrap (Campbell & Taylor, 1992; Diaconis & Efron, 1983), Procrustean rotation (Tucker & Taylor, 1991), the U-method (Crask & Perreault, 1977; Daniel, 1989; Prosser, 1991), and cross-validation methods (Cooil et al., 1987; Fish, 1986; Loftin, 1991).

The purpose of the present paper is to explain one of these procedures, cross-validation. A form of the technique, double cross-validation, will be applied in a canonical correlation analysis using a heuristic data set. Before turning to the invariance procedure, however, a brief overview of both invariance testing and canonical correlation analysis is provided.

A Note on Invariance Testing

Research results which are found to be invariant are relatively stable across samples; that is, they are not sample specific, and therefore, are likely to replicate in future studies. Because replicability is the "cornerstone of science" (Carver, 1978, p. 392), invariance is part of the overall result a researcher hopes to report from a study. However, sampling error may be substantial enough to lead to sample specific results.

Some sampling error will be present in any study, despite care taken during the design phase to ensure that the population is randomly sampled. Further, the statistical procedures used are likely to exaggerate unusual features in data due to error. According to Thompson (1991b), "all classical analytic methods are correlational" (p. 87), and correlational procedures tend to capitalize on chance characteristics of the sample, thereby decreasing the generalizability of the findings (Thompson, 1981). The effect of this capitalization cannot be statistically minimized, but it can be detected using invariance procedures. Thus, invariance testing, not significance testing, is the researcher's protection against reporting misleading or inaccurate results which can wrongly influence a field of study.

The logic of invariance analysis is uncomplicated. A number of frequently used procedures draw subsamples from data in hand and recalculate for each subsample the statistic of interest. Results for both the total sample and the subsamples are compared empirically. If the results are similar, a basis is provided for vesting confidence in a conclusion that the study findings are generalizable. On the other hand, if the results vary substantially across resamplings, it is likely that the study findings are an artifact of the sample and will not be replicated in future research. Thompson (1991a) emphasizes the need for making these comparisons empirically, rather than relying on the way the data "look." Data that appear to be different may, through empirical comparisons, turn out to be quite similar. The reverse is also true. Subjective judgment is insufficient as a guide to making conclusions about generalizability.

One of the most commonly used invariance procedures is cross-validation. According to Cooil et al. (1987), cross-validation methods involve dividing a single sample "into two subsamples..., estimating the coefficients on one and validating on the other" (p. 272). Double cross-validation follows the same procedure, except that the coefficients from both subsamples are validated on the other. Guidelines have not been established regarding how a sample should be subdivided (Cooil et al., 1987). Some researchers opt for a 50-50 split, others prefer a 60-40, or even a 70-30 split. Cooil et al. caution that data-splitting for cross-validation purposes poses the problem of violating the sufficiency principle. This occurs

> because the estimates are not based on the entire sample. ...Underutilization of the information in the sample produces inefficient coefficient estimates and correspondingly inefficient predictions. This outcome occurs with any size sample, but small samples aggravate the problem. (Cooil et al., 1987, p. 272)

Violation of the sufficiency principal and data splitting with small samples pose serious limitations to this invariance technique. If sample size is small, it is advisable to choose some other invariance procedure, such as the jackknife, which is not so much affected by sample size. However, if sample size is adequate, Cooil et al. suggest that an uneven split of the sample compensates for the problem with the sufficiency principle.

Thompson (1984) notes that invariance testing is particularly important in canonical correlation analysis because like all parametric methods the procedure "tends...to capitalize on sampling error" (p. 41). Thorndike (1978) explains further that use of an invariance procedure "is...important for canonical analysis...because there are two sets of weights, each of which will make maximum use of sample specific covariation" (p. 180).

An overview of canonical correlation analysis will make Thorndike's caution clearer.

A Brief Overview of Canonical Correlation Analysis

Canonical correlation analysis has been called the most general linear model because it subsumes all univariate and multivariate parametric procedures as special cases (Thompson, 1991b). An advantage of using canonical correlation analysis is that this procedure does not discard variance by forcing subjects into groups, as happens with OVA procedures (e.g., ANOVA or MANOVA). Rather, all of the variability in the data is used in the canonical calculations (Thompson, 1984).

Canonical correlation analysis is a complex procedure that might be more easily understood if it is compared to a familiar procedure, such as regression. In multiple regression, the task is to explore the relationship between one dependent, or criterion, variable and multiple independent, or predictor, variables. For example, we might wish to explore the relationship between the criterion variable, student achievement, and such predictor variables as ability, family income, and class size.

In canonical correlation, one is able to explore the relationship between multiple criterion variables and multiple predictor variables. Expanding the above example, it might be of interest to use several criterion variables. In addition to student achievement, student behavior and attendance could also be examined using the same set of predictor variables (ability, income, and pupil-teacher ratio). Thus, in canonical correlation, there are multiple criterion variables as well as multiple predictor variables involved in the analysis. The use of canonical correlation when several criterion variables are being studied can lead to remarkably different results than would occur with multiple univariate analyses, each using a different criterion variable, even if the same data were used in the analyses (Fish, 1988).

Thompson (1984, 1991b) explains that canonical analysis relies on factor analytic procedures in which weights are used to create synthetic combinations of the original variables. These synthetic combinations, or functions, are comprised of composite predictor scores and composite criterion scores for each subject. According to Thompson (1991b), "the bivariate correlation between the [predictor and criterion] scores...is...the canonical correlation coefficient (Rc)" (p. 83).

There will be as many functions produced by a canonical correlation as there are variables in the smaller variable set (predictor or criterion). Determining which of the functions to interpret depends on the strength of the squared canonical correlation (Rc^2) and the test of statistical significance for each function. Statistically significant functions with a large Rc^2 should be interpreted (Thompson, 1991b).

The weights described above are not unique to canonical analysis. They are used in most statistical procedures, but are given different names. For example, in regression, they are called beta weights; in factor analysis, they are referred to as pattern coefficients; in canonical correlation, they are standardized canonical function coefficients. The weights are applied to values for the predictor and criterion variables in such a way that redundancy among the synthetic composites is eliminated. Further, the sets of weights are derived in such a way that they "maximize [the] correlation coefficient" between the predictor and criterion variables (Pedhazur, 1982, p. 722).

As with other types of multivariate analysis, there has been some controversy concerning whether the weights or the structure coefficients should be interpreted. This split in opinion has been based in part on the belief that the structure coefficients may be more stable than the weights. But Monte Carlo work by Thompson (1991a) indicated that function [weights] and structure coefficients are influenced by sampling error to roughly equal degrees. Thus, the standardized canonical function coefficients are critical not only for developing the synthetic variates on which the analysis is founded, but they are also necessary for substantive interpretations, and are of central importance in cross-validation, as will be demonstrated.

Cross-validation of the Weights

The current paper presents a double cross-validation of the weights in a canonical correlation analysis to test for invariance in a study of university leadership conducted by Tucker (1990). The sample consisted of the university chancellor, the top administrative staff, deans, department chairs, and certain faculty. One purpose of the study was to determine whether style of leadership could predict (a) satisfaction with leaders, (b) perceptions of leader effectiveness, and (c) willingness of followers to expend extra effort.

The Multifactor Leadership Questionnaire ([MLQ] Bass & Avolio, 1990) was distributed to a sample of 200 subjects; 105 usable surveys were returned. Certain subscales of the MLQ assess the extent to which leaders exhibit one of three leadership styles. One style, transactional leadership, is characterized by an exchange relationship in which leaders dispense rewards to motivate followers. A second style, transformational leadership, augments transactional skills by using charisma and other attributes to motivate followers to excel beyond expectations. Finally, laissez-faire leadership, the third style, can best be described as an absence of leadership. These three types of leadership comprised the set of predictor variables in this study. The criterion variables (satisfaction with leaders, followers' perception of leader effectiveness, and extra effort exerted by followers) were also measured through the subscales of the MLQ.

Each variable set contains three variables, therefore, the canonical analysis yielded three canonical functions (i.e., three sets of weights). These functions, the standardized canonical function coefficients, and the Rc^2 are presented in Table 1. Although both function one and two are significant, only function one has a substantial Rc^2, therefore, only function one will be used in the invariance analysis.

INSERT TABLE 1 ABOUT HERE.

To compute the double cross-validation, the following steps, outlined by Fish (1986), were followed in sequence:

1) a canonical correlation was computed for the total sample;
2) the sample was randomly split into two groups at a ratio of roughly 70-30;
3) values for each of the variables in both the predictor and criterion set were converted into z-scores for both groups;
4) a canonical correlation analysis was computed for both groups to obtain the standardized canonical function coefficients;
5) four composite criterion variables and four composite predictor variables were computed for each case by multiplying the z-scores with the function coefficient for each variable as follows:
   a) composite variable CRIT11 was derived by multiplying the z-scores and weights for each criterion variable for group 1;
   b) composite variable PRED11 was derived by multiplying the z-scores and weights for each predictor variable for group 1;
   c) composite variable CRIT12 was derived by multiplying for each criterion variable the z-scores for group 1 with the weights for group 2, cross-validating the criterion variables;

d) composite variable PRED12 was derived by multiplying for each predictor variable the z-scores for group 1 with the weights for group 2, cross-validating the predictor variables;

e) in order to calculate a double cross-validation, composite variables CRIT22, PRED22, CRIT21, and PRED21 were created following the procedure just described, but adjusting to use the z-scores for group 2 and the weights for group 1;

6) four correlation coefficients were computed as follows:

a) the predictor and criterion composite variables for group 1 were correlated (Rc11);

b) the cross-validated predictor and criterion composite variables for group 1 were correlated (Rc12);

c) the predictor and criterion composite variables for group 2 were correlated (Rc22);

d) the cross-validated predictor and criterion composite variables for group 2 were correlated (Rc21);

7) each correlation coefficient was squared;

8) differences between Rc11 and Rc12, and Rc22 and Rc21 were computed.

Small differences between the squared correlation coefficients are indicative of invariant results.

The standardized canonical function coefficients used in making the calculations can be found in Table 2; z-scores for the total sample were too numerous to reproduce here. Correlations among the predictor and criterion composites computed during the invariance process are presented in Table 3. SPSSx command lines used to generate the cross-validation are contained in Appendix A.

INSERT TABLES 2 AND 3 ABOUT HERE.

Results of the double cross-validation are presented in Table 4. Each correlation coefficient in Table 3 was squared and the result of the cross-validated values were subtracted from those values not cross-validated. The resulting invariance coefficient in each case was a small .02, indicating that the results of this study are very likely to replicate in future studies.

INSERT TABLE 4 ABOUT HERE.

Summary

By using invariance procedures, researchers can gain a measure of confidence that study results are not unique to the sample employed. Although invariance results are seldom found in published research, some (Carver, 19778; Lykken, 1968; Thompson, 1989) argue that establishing generalizability is perhaps more important than evaluating statistical significance and should be a part of every research study. With regard to canonical correlation, however, Thompson (1989) counsels that use of cross-validation procedures for canonical weights is a method of establishing replicability, not stability; "function coefficients can appear to be quite different yet may yield equivalent synthetic composite variables" (p. 13).

Two issues should be kept in mind when planning an invariance analysis. One is the sensitivity of the chosen technique to small sample size. The reliability of cross-validation results is compromised by small sample size, therefore, the procedure should be avoided in such instances. The other issue has to do with interpreting invariance coefficients. Invariance testing is a relatively new field and parameters for interpreting the results are just beginning to be established. Until more work is done in this area, Thompson (1984) suggests using more than one invariance procedure, as a matter of caution and care.

References

Bass, B. M. & Avolio, B. J. (1990). Transformational leadership development: Manual for the Multifactor Leadership Questionnaire. Palo Alto, CA: Consulting Psychologists Press.

Campbell, K. T. & Taylor, D. L. (1992, January). Application of a bootstrap technique with factor analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston.

Carver, R. P. (1978). The case against statistical significance testing. Harvard Educational Review, 48, 378-399.

Cooil B., Winer, R. S., & Rados, D. L. (1987). Cross-validation for prediction. Journal of Marketing Research, 24, 271-279.

Crask, M. R. & Perreault, W. D., Jr. (1977). Validation of discriminant analysis in marketing research. Journal of Marketing Research, 14, 60-68.

Daniel, L. G. (1989, January). Use of the jackknife statistic to establish the external validity of discriminant analysis results. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston. (ERIC Document Reproduction Service No. ED 305 382)

Diaconis, P. & Efron, B. (1983). Computer-intensive methods in statistics. Scientific American, 248(5), 116-130.

Fish, L. (1986, November). The importance of invariance procedures as against tests of statistical significance. Paper presented at the annual meeting of the Mid-South Educational Research Association, Memphis.

Fish, L. (1988). Why multivariate methods are usually vital. Measurement and Evaluation in Counseling and Development, 21, 130-137.

Huck, S. W., Cormier, W. H. & Bounds, W. G., Jr. (1974). Reading statistics and research. New York: Harper & Row.

Loftin, L. B. (1991, January). The importance of invariance procedures: Application of cross-validation in discriminant analysis. Paper presented at the annual meting of the Southwest Educational Research Association, San Antonio.

Lykken, D. T. (1968). Statistical significance in psychological research. Psychological Bulletin, 70, 151-159.

Pedhazur, El J. (1982). Multiple regression in behavioral research. 2nd ed. New York: Holt, Rinehart and Winston.

Prosser, B. (1991, January). Use of the U-Method to establish the external validity of discriminant analysis results. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio.

Taylor, D. (1991, January). Evaluating the sample specificity of discriminant analysis results using the jackknife statistic. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio.

Thompson, B. (1981). Utility of invariance coefficients. Perceptual and Motor Skills, 52, 708-710.

Thompson, B. (1984). Canonical correlation analysis: Uses and interpretation. Newbury Park, CA: Sage.

Thompson, B. (1988, April). Canonical correlation analysis: An explanation with comments on correct practice. Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. ED 295 957)

Thompson, B. (1989). Statistical significance, result importance, and result generalizability: Three noteworthy but somewhat different issues. Measurement and Evaluation in Counseling and Development, 22, 2-6.

Thompson, B. (1989, March). Invariance of multivariate results: A Monte Carol study of canonical coefficients. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Thompson, B. (1991a). Invariance of multivariate results. Journal of Experimental Education, 59, 367-382.

Thompson, B. (1991b). A primer on the logic and use of canonical correlation analysis. Measurement and Evaluation in Counseling and Development, 24, 80-95.

Thompson, B. (1992, April). Bootstrap and other approaches to evaluating statistical significance. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Thorndike, R. M. (1978). Correlational procedures for research. New York: Gardner Press.

Tucker, M. L. (1990). Higher education leadership: Transformational leadership as a predictor of satisfaction, effectiveness, and extra effort. Unpublished doctoral dissertation, University of New Orleans, New Orleans.

Tucker, M. L. & Taylor, D. L. (1991, January). Applying Procrustean rotation to evaluate the generalizability of research results. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio.

Tukey, J. W. (1969). Analyzing data: Sanctification or detective work? American Psychologist, 24, 83-91.

Table 1

Standardized Canonical Function Coefficients

	Function I**	Function II*	Function III
Transformational	-.98	.91	.61
Transactional	-.01	1.33	.19
Laissez-faire	.03	.29	1.15
Extra Effort	-.49	.75	1.26
Effectiveness	-.50	.47	-1.52
Satisfaction	-.10	-1.45	.33
Rc^2	.87	.11	.01

*$p < .001$
**$p < .01$

Table 2

Standardized Canonical Function Coefficients
for Function 1 for both Groups

	Function I	
	Grp 1	Grp 2
Transformational	1.005	-.975
Transactional	-.035	-.055
Laissez-faire	-.036	-.039
Extra Effort	.523	-.365
Effectiveness	.506	-.558
Satisfaction	.063	-.155

Table 3

Correlations between the Predictor and Criterion Composites

PRED11 x CRIT11 = .94
PRED12 x CRIT12 = .93

PRED22 x CRIT22 = .94
PRED21 x CRIT21 = .93

Table 4

Invariance Coefficients

Squared Correlation		Squared Invariance Correlation		Invariance Coefficient
$Rc11^2$	(minus)	$Rc12^2$	(equals)	Inv. Coef.
.88	-	.86	=	.02
$Rc22^2$	(minus)	$Rc21^2$	(equals)	Inv. Coef.
.88	-	.86	=	.02

Appendix

```
TITLE "CANONICAL CORRELATION ANALYSIS FOR SERA 1992"
SET WIDTH=80
FILE HANDLE DT/NAME='D4.DAT'
DATA LIST  FILE=D4.DAT /ITEM1 TO ITEM80 1-80 LEADER 81-88(4)
 RELATE 89-90

*****************************************************************
A series of COMPUTE statements here created the variables
EEFFT, EFFEC, SAT, TFSCOR, TASCOR, and LF that are used in the
canonical correlation.
*****************************************************************

SUBTITLE 'CANONICAL CORRELATIONS ANALYSIS OF ENTIRE DATA SET'
MANOVA  EEFFT EFFEC SAT WITH TFSCOR TASCOR LF
    /PRINT=SIGNIF(EIGEN DIMENR) CELLFINO (MEANS)
     DISCRIM(STAN,COR,ALPHA(1.00))
    /DESIGN

*****************************************************************
A canonical correlation analysis has been run on the total data
set.  The data set will now be split into two groups to perform a
cross-validation using the variable RELATE.
*****************************************************************

DO IF (RELATE2 EQ 1 OR RELATE2 EQ 2 OR RELATE2 EQ 3 OR RELATE2 EQ
 4 OR RELATE2 EQ 5 OR RELATE2 EQ 6)
COMPUTE A=1
ELSE
COMPUTE A=2
END IF

SORT CASES BY A
SPLIT FILE BY A
DESCRIPTIVES VARIABLES=EEFFT EFFEC SAT TFSCOR TASCOR LF
 /SAVE
LIST VARIABLES=ZEEFFT ZEFFEC ZSAT ZTFSCOR ZTASCOR ZLF
 /FORMAT=NUMBERED/CASES=500

MANOVA EEFFT EFFEC SAT WITH TFSCOR TASCOR LF
    /PRINT=SIGNIF(EIGEN DIMENR) CELLFINO (MEANS)
     DISCRIM(STAN,COR,ALPHA(1.00))
    /DESIGN
SPLIT FILE OFF

DO IF A=1
COMPUTE CRIT11=(ZEEFFT*.523) + (ZEFFEC*.506) + (ZSAT*.063)
COMPUTE PRED11=(ZTFSCOR*1.005) + (ZTASCOR*-.035) + (ZLF*-.036)
END IF
LIST VARIABLES=CRIT11 PRED11
CORRELATION VARIABLES=CRIT11 WITH PRED11
```

```
DO IF A=1
COMPUTE CRIT12=(ZEEFFT*-.365) + (ZEFFEC*-.558) + (ZSAT*-.155)
COMPUTE PRED12=(ZTFSCOR*-.975) + (ZTASCOR*-.055) + (ZLF*-.039)
END IF
LIST VARIABLES=CRIT12 PRED12
CORRELATION VARIABLES=CRIT12 WITH PRED12

DO IF A=2
COMPUTE CRIT22=(ZEEFFT*-.365) + (ZEFFEC*-.558) + (ZSAT*-.155)
COMPUTE PRED22=(ZTFSCOR*-.975) + (ZTASCOR*-.055) + (ZLF*-.039)
END IF
LIST VARIABLES=CRIT22 PRED22
CORRELATION VARIABLES=CRIT22 WITH PRED22

DO IF A=2
COMPUTE CRIT21=(ZEEFFT*.523) + (ZEFFEC*.506) + (ZSAT*.063)
COMPUTE PRED21=(ZTFSCOR*1.005) + (ZTASCOR*-.035) + (ZLF*-.036)
END IF
LIST VARIABLES=CRIT21 PRED21
CORRELATION VARIABLES=CRIT21 WITH PRED21
```