A COMPARISON OF FACTOR MATRICES ARISING FROM A COMMON SET OF ITEMS BUT DERIVED USING DIFFERENT CORRELATION COEFFICIENTS

Marilyn L. Bush
Regents of the University System of Georgia

In any factor analytic study the choice of an appropriate coefficient of correlation is of prime importance. The type of coefficient used is ordinarily dependent upon the type and distribution of the data under consideration because certain assumptions are necessary with the use of each. Pearson product-moment correlations are often used in factor analysis without considering the assumptions necessary for their use, e.g., both variables ought to be measured on continuous scales and regressions ought to be linear.

Since variables measured on a dichotomous scale result in point coefficients, other types of correlation coefficients should be considered as estimates of the Pearson r for use as input for factor analysis. Three other types of correlation coefficients have been studied in addition to the Pearson r in an effort to determine the effect of each on a common set of data. They are the tetrachoric r, the phi coefficient, and phi-over-phi max.

The tetrachoric r assumes both variables to be continuous, normally distributed, and linearly related even if the data have been reduced to dichotomous variables. The phi coefficient assumes the two distributions are true dichotomies and it can be applied to data that are measurable on a continuous scale if certain allowances for that fact are made. Phi-over-phi max assumes essentially the same things as phi and in addition attempts to correct for artificial restrictions on phi. The formula for the phi coefficient is identical to the formula for the Pearson correlation in the case of dichotomous variables.

The statistical technique of factor analysis is used generally in an attempt to separate common sources of variance between intercorrelated measures which have been arranged in a certain way. It is desirable to determine the smallest number of variables needed in order to account for the observed variance and to calculate the extent to which they contribute to the other measures used. Factor analysis was used in this study to determine if factors obtained by the use of different coefficients of correlation were substantially the same.

Previous research in the area of factor analysis using different types of point coefficients of correlation is not common. Comrey and Levonian (1958) in a comparison of three types of point coefficients in factor analysis of Minnesota Multiphasic Personality Inventory items make reference to problems encountered using the phi coefficient, phi-over-phi max, and tetrachoric r. After factor analyzing with phi-over phi max, they found that: (1) it seemed impossible to extract enough factors to make the loading on the last factors very small; and (2) communalities frequently exceeded 1.0 rather early in the factor extraction process. Even more extreme results occurred with the tetrachoric r. Excessively high communalities, often over 1.0, occurred too early, and late factors failed to drop satisfactorily in variance. Better results were obtained with the phi coefficient as factor loadings on late factors seemed reasonably small without obtaining excessively high communalities.

Largest loadings were reported for tetrachoric, next largest for phi-over-phi max, and smallest for phi. They further consider whether the main factors obtainable in such an analysis with different coefficients are substantially the same. The conclusion reached is that the main factors are substantially the same with phi, phi-over-phi max, and tetrachoric, but since phi-over phi max and tetrachoric frequently do lead to unreasonably high communalities, the phi coefficient is the method of choice in point correlation work where factor analysis is to follow.

Carroll (1961) in a discussion of problems of choosing correlation coefficients states that:

> No assumptions are necessary for the computation of a Pearsonian coefficient, but the interpretation of its meaning certainly depends upon the extent to which the data conform to an appropriate statistical model for making this interpretation. As the actual data depart from a fit to such a model, the limits of the correlation coefficient may contract, and the adjectival interpretations are less meaningful. The limiting case is provided when the two distributions are dichotomous and the points of dichotomy are asymmetrical between the two distributions, for here the Pearsonian coefficient (in this case, called the phi coefficient) does not, in general range between plus and minus one, as Ferguson showed some years ago.

Method

Data for this study came from ninth-graders of one Florida county participating in the Florida State-Wide Ninth Grade Testing Program during the year 1961-62.

Metropolitan Achievement Test (MAT) answer sheets from the five schools were arranged alphabetically by school. Then students within each school were arranged alphabetically. From the total group of approximately 1800 ninth graders, 311 students were selected by pulling every sixth paper.

Scoring for each of the 311 students were obtained from the MAT Reading Test, Advanced, Form EM. The responses to the 44 items were scored right or wrong, yielding a dichotomy. Scores for each of the 311 were punched on data cards and used as input for a correlation program for the IBM 709 computer. Two correlation programs were used, one for computing Pearson product-moment correlations and another for computing tetrachoric, phi-over-phi max, and phi correlations.[1]

Correlation matrices containing 44 variables were obtained for each of the four types of correlation coefficients. Each of the four types of correlation matrices was then factor analyzed using the factor analysis computer program BIMD 17 (1962). This program extracts factors by the principle axis method and performs orthogonal rotations by the varimax method. The factor analysis of the four types of correlation matrices produced, on the average, 25 factors for each type of correlation coefficient. However, the majority of notable loadings were contained in ten factors. Thus, in the final analysis only 10 factors for each type of correlation coefficient were analyzed.

The rotated factor matrices were compared visually for similarities of pattern and relative loadings. Then to further match the factors, the four orthogonally rotated factor matrices, each consisting of 10 factors, were used as input for a transformation analysis program which computed indices of similarity between the four sets of 10 factors. The program used the methods of Ahmavaara (1954) and Kaiser (1960) and was compiled by King (1960).

If indices for a majority of the factors were .80 or above, regardless of position in the vector, the two types

---

[1]In this program prepared by King and Martin (1962), the tetrachoric computation was a polynomial solution, not a table look-up or Cosine-Pi approximation.

of matrices were considered to have produced the same pattern of factors. If indices were .50 to .79, this combination of matrices was considered as having produced factors nearly the same. However, if indices were .49 or less, it was judged that the factors produced by the different types of correlation methods were not the same nor nearly the same. These limits were arbitrarily established using commonly accepted values for the degree of relationship shown by a correlation coefficient. This was done since the indices range from -1.0 to +1.0 as does the correlation coefficient.

Results

No attempt was made to name factors psychologically nor to interpret them by subject content. The factor analytic patterns of the four factor matrices (Pearson, tetrachoric, phi-over-phi max, and phi), together with the indices of similarity were used as the basis for this analysis.

The results of factor matching showed a fairly high degree of similarity in the factor patterns resulting from the use of the four different types of correlation coefficients. The highest agreement was found between factor matrices for Pearson and phi. This was expected due to the equivalency of the formulae in the dichotomous case. Nine of the ten factors had indices of similarity of .97 or better. The average of the remaining 91 table entries (the 9 of .97+ excluded) was .013.

The next highest relationship was found between factor matrices for Pearson and phi-over-phi max. Three of the 10 factors had indices of similarity of .80 or better with 7 of the 10 registering at least .50. Here, the average of the 93 remaining table entries was .009.

Summarizing for the other pairs: tetrachoric with phi-over-phi max showed 3 indices of .80+, 6 of 10 at least .50, and an average of .018 for the remaining 94 table entries; tetrachoric with phi, and phi-over-phi max with phi, each had 2 indices of .80+, 7 of 10 at least .50, and averages of -.018 and .008 respectively for the remaining 93 table entries; Pearson with tetrachoric: 2 indices of .80+, with 5 of 10 at least .50, and an average of .028 for the remaining 95 table entries.

From this analysis, Pearson and phi produced the same pattern of factors for these data. The remaining combinations were less clear in relationship. Furthermore, the Pearson-tetrachoric combination produced the smallest number of factors with indices of similarity of .50 or better (5 of 10). Under these circumstances, it is hard to say that all of the methods produce factors which could be considered the same, but they agree highly on enough factors to be termed nearly the same. No standard statistical tests are available for determining the significance of a given magnitude of similarity.

Discussion and Summary

The purpose of this study was to determine if factors obtained by the use of Pearson product-moment correlations, tetrachoric correlations, phi-over-phi max correlations, and phi coefficients were substantially the same for a common set of items.

The reader is reminded that these results are based on one sample of empirical data. Additional empirical evidence on other populations is required for greater generalization of the results. Considerable evidence could be obtained by "Monte-Carlo" computer methods which involve the generation of a large amount of artificial data. Rigorous proof of equivalency or non-equivalency of factor structure would need to be done mathematically.

The evidence from this sample shows that factor patterns resulting from the use of Pearson and from phi were definitely the same. All of the combinations produced at least 5 factors of .50 or above with the average being between 6 and 7 out of a possible 10. Thus, the use of any of the four types of correlation coefficients should yeild fairly comparable results in a similar situation.

References

1. Ahmavaara, Y. The mathematical theory of factorial invariance under selection. _Psychometrika_, 19 (1954), 27-38.

2. Carroll, J. B. The nature of the data, or how to choose a correlation coefficient. _Psychometrika_, 26 (1961), 347-371.

3. Comrey, A. L. and Levonian, E. A comparison of three point coefficients in factor analysis of MMPI items. _Educational and Psychological Measurement_, 19 (1958), 739-754.

4. Division of Biostatistics, Department of Preventive Medicine and Public Health. "Factor Analysis and Rotate for Monitor and Non-Monitor Operation-BIMD 17." U.C.L.A., 1962.

5. Kaiser, H. F. Varimax solution for primary mental abilities. _Psychometrika_, 25 (1960), 153-158.

6. King, F. J. Program for transformation analysis. Unpublished report, Department of Educational Research and Testing, Florida State University, 1960.

7. King, F. J. and Martin, D. Correlation of dichotomized variables. Unpublished report, Department of Educational Research and Testing, Florida State University, 1962.

8. _Metropolitan Achievement Tests_. New York: Harcourt Brace and World, Incorporated, 1961.