Pages / Home

The Minimum Sample Size in Factor Analysis

Added by Nathan Zhao, last edited by Nathan Zhao on Mar 23, 2009

Why I Studied This Issue?

I did an apprentice project studying the reasons why students withdrew from their online courses. In this project, I got a dataset that had 35 variables indicating various withdrawal reasons. I wanted to use factor analysis to reduce the 35 variables to a few categories of withdrawal reasons. However, I only have 47 casese in the dataset. Many people suggested that the number of cases was too small for performing a factor analysis. But, I really do not want to waste the time and energy I had spent and just throw away this dataset. Yes! I want to "explain the most with the least" (Henson & Roberts, 2006, p. 393).

Thus, I deceided to find out what is the minimum sample size (i.e., the minimum number of cases - some researchers called it subjects) for performing factor analysis. Here is the related information I found.

The General Recommendations

There are two categories of general recommendations in terms of minimum sample size in factor analysis. One category says that the absolute number of cases (*N*) is important, while the another says that the subject-to-variable ratio (*p*) is important. Arrindell and van der Ende (1985), Velicer and Fava (1998), and MacCallum, Widaman, Zhang and Hong (1999) have reviewed many of these recommendations.

of sample size

- Rule of 100: Gorsuch (1983) and Kline (1979, p. 40) recommanded at least 100 (MacCallum, Widaman, Zhang & Hong, 1999). No sample should be less than 100 even though the number of variables is less than 20 (Gorsuch, 1974, p. 333; in Arrindell & van der Ende, 1985, p. 166);
- Hatcher (1994) recommanded that the number of subjects should be the larger of 5 times the number of variables, or 100. Even more subjects are needed when communalities are low and/or few variables load on each factor (in David Garson, 2008).
- Rule of 150: Hutcheson and Sofroniou (1999) recommends at least 150 - 300 cases, more toward the 150 end when there are a few highly correlated variables, as would be the case when collapsing highly multicollinear variables (in David Garson, 2008).
- Rule of 200. Guilford (1954, p. 533) suggested that *N* should be at least 200 cases (in MacCallum, Widaman, Zhang & Hong, 1999, p84; in Arrindell & van der Ende, 1985; p. 166).
- Rule of 250. Cattell (1978) claimed the minimum desirable *N* to be 250 (in MacCallum, Widaman, Zhang & Hong, 1999, p84).
- Rule of 300. There should be at least 300 cases (Noru?is, 2005: 400, in David Garson, 2008).

- Significance rule. Lawley and Maxwell (1971) suggested 51 more cases than the number of variables, to support chi-square testing (in David Garson, 2008).
- Rule of 500. Comrey and Lee (1992) thought that 100 = poor, 200 = fair, 300 = good, 500 = very good, 1,000 or more = excellent They urged researchers to obtain samples of 500 or more observations whenever possible (in MacCallum, Widaman, Zhang & Hong, 1999, p84).

of subjects-to-variables (STV)ratio

- A ratio of 20:1. Hair, Anderson, Tatham, and Black (1995, in Hogarty, Hines, Kromrey, Ferron, & Mumford, 2005)
- Rule of 10. There should be at least 10 cases for each item in the instrument being used. (David Garson, 2008; Everitt, 1975; Everitt, 1975, Nunnally, 1978, p. 276, in Arrindell & van der Ende, 1985, p. 166; Kunce, Cook, & Miller, 1975, Marascuilor & Levin, 1983, in Velicer & Fava, 1998, p. 232)
- Rule of 5. The subjects-to-variables ratio should be no lower than 5 (Bryant and Yarnold, 1995, in David Garson, 2008; Gorsuch, 1983, in MacCallum, Widaman, Zhang & Hong, 1999; Everitt, 1975, in Arrindell & van der Ende, 1985; Gorsuch, 1974, in Arrindell & van der Ende, 1985, p. 166)
- A ratio of 3(:1) to 6(:1) of STV is acceptable if the lower limit of variables-to-factors ratio is 3 to 6. But, the absolute minimum sample size should not be less than 250.(Cattell, 1978, p. 508, in Arrindell & van der Ende, 1985, p. 166)
- Ratio of 2. "[T]here should be at least twice as many subjects as variables in factor-analytic investigations. This means that in any large study on this account alone, one should have to use more than the minimum 100 subjects" (Kline, 1979, p. 40).

Statistical Research Findings on Minimum Sample Size

Little statistical research in the fields of Education and Behaviour Science has shed light on the issue of establishing a minimum desirable level of sample size (MacCallum, Widaman, Zhang & Hong, 1999). These studies used either artificial or empirical data to investigate the minimum sample size or STV ratio that is required in order to recover the population factor structure. In this section, I will summarize the minimum sample size and STV ratio that these studies had examined.

- Barrett and Kline (1981, in MacCallum, Widaman, Zhang & Hong, 1999) used two large empirical data sets to investigate this issue. They drew sub-samples of various size from the original full samples and performed factor analysis with each sub-sample to compare the results of sub-samples with the result of full samples. They obtained good recovery:
  - from a sub-sample of N = 48 [1] for one data set that has 16 variables, which represents a STV ratio of 3.0;
  - and from a sub-sample of N = 112 for another data set that has 90 variables, which STV ratio is 1.2.

> [1] This number was reported as 50 "to be the minimum to yield a clear, recognizable factor pattern" (p. 167) in Arrindell and van der Ende's paper (1985).

- Arrindell and van der Ende (1985) used two large empirical data sets that have 1104 cases and 960 cases respectively to examine the minimum sample sizes and STV ratios that can produce stable factor structure. By drawing sub-samples from the two large data sets, the authors found that:
  - for the first data set, which had 76 variables, the minimum STV ratio (p) that required to produce clear, recognizable factor solution was 1.3 and the corresponding sample size (N) was 100;
  - for the second data set, which has 20 variables, the minimum STV ratio (p) was 3.9 and the corresponding sample size (N) was 78.

- MacCallum, Widaman, Zhang & Hong (1999) conducted a Monte Carlo Study on sample size effects. They obtained an excellent recovery (100% convergence) of population factor structure with a sample size (N) of 60 and 20 variables. However, this result was obtained only when the level of communality (over .7 in average) and overdetermination (3 loaded factors) were high (Table 1 on page 93).

- Preacher & MacCallum (2002) conducted a Monte Carlo study. Their conclusion is:
  - N had by far the largest effect on factor recovery, which exhibited a sharp drop-off below N s of 20 or so. (p.157)

The Minimum Sample Size or STV Ratio Used in Practical Studies

- Henson and Roberts (2006) reported a review of 60 exploratory factor analysis in four journals: *Educational and Psychological Measurement, Journal of Educational Psychology, Personality and Individual Differences*, and *Psychological Assessment*.
  - Minimum sample size reported: 42.
  - Minimum STV ratio reported: 3.25:1; 11.86% of reviewed studies used a ratio less than 5:1.

- Fabrigar, Wegener, MacCallum, and Strahan (1999) reported a review of articles that used EFA in two journals: *Journal of Personality and Social Psychology (JPSP)* and *Journal of Applied Psychology (JAP)*.
  - Sample size: 30 (18.9%) articles in JPSP and 8 (13.8%) in JAP were 100 or less.
  - Ratio of variable to factors: 55 (24.6%) papers in JPSP and 20 (34.4%) in JAP were 4:1 or less.

- Costello and Osborne (2005) surveyed two year's PsychINFO articles that reported principal components or exploratory factor analysis.
  -

STV ratio	% of studies	Cumulative %
2:1 or less	14.7%	14.7&
> 2:1, ? 5:1	25.8%	40.5%
> 5:1, ? 10:1	22.7%	63.2%

> 10:1, ? 20:1	15.4%	78.6%
> 20:1, ? 100:1	18.4%	97.0%
> 100:1	3.0%	100.0%

- Ford, MacCallum, and Tait (1986) examined articles published in *Journal of Applied Psychology*, *Personnel Psychology*, and *Organizational Behavior and Human Performance* during the period of 1974 - 1984.
  - RTV ratio: 27.3% of the studies were less than 5:1, 56% were less than 10:1.

Factors Related to Sample Size

Research has demonstrated that the general rule of thumb of the minimum sample size are not valid and useful (MacCallum, Widaman, Zhang, & Hong, 1999; Preacher & MacCallum, 2002). It is hard and simplicity to say whether absolute sample size is important or the STV ratio is important in factor analysis. The minimum level of *N* (sample size) was dependent on other aspects of design, such as:

- Communality of the variables
  - The communality measures the percent of variance in a given variable explained by all the factors jointly and may be interpreted as the reliability of the indicator (Gason, 2008).
  - If communalities are high, recovery of population factors in sample data is normally very good, almost regardless of sample size, level of overdetermination, or the presence of model error (MacCallum, Widaman, Preacher, and Hong, 2001, p. 636)
  - MacCallum, Widaman, Zhang, and Hong (1999) suggested communalities should all greater than .6, or the mean level of communality to be at least .7 (p. 96).
  - Item communalities are considered "high" if they are all .8 or greater - but this is unlikely to occur in real data (Costello & Osborne, 2005, p. 4).

- Degree of overdetermination of the factor (or number of factors/number of variables)
  - Overdetermination is the factor-to-variable ratio (Preacher & MacCallum, 2002).
  - Six or seven indicators per factor and a rather small number of factors is considered as high overdetermination of factors if many or all communalities are under .50 (MacCallum, Widaman, Zhang, & Hong, 1999).
  - A minimum of 3 variables per factor is critical. This confirms the theoretical results of T. W. Anderson and Rubin (1956; also see McDonald & Krane, 1977, 1979, and Rindskopf, 1984). (Velicer, & Fava, 1998, p. 243).
  - At least four measured variables for each common factor and perhaps as many as six (Fabrigar, Wegener, MacCallum, & Strahan, 1999, p. 282)
  - A factor with fewer than three itmes is generally weak and unstable (Costello & Osborne, 2005, p. 5)

- Size of loading
  - Item loading magnitude accounted for significant unique variance in the expected direction in all but one case, and in most cases was the strongest unique predictor of congruence between sample and population (Osborne, & Costello, 2004).
  - The sample-to-population pattern fit was very good for the high (.80) loading condition, moderate for the middle (.60) loading condition, and very poor (.40) for the low loading condition (Velicer & Fava, 1998).
  - 5 or more strongly loading items (.50 or better) are desirable and indicate a solid factor (Costello & Osborne, 2005, p. 5).
  - If components possess four or more variables with loadings above .60, the pattern may be interpreted whatever the sample size used . Similarly, a pattern composed of many variables per component (10 to 12) but low loadings (= .40) should be an accurate solution at all but the lowest sample sizes (N < 150). If a solution possesses components with only a few variables per component and low component loadings, the pattern should not be interpreted unless a sample size of 300 or more observations has been used. (Guadagnoli & Velicer, 1988, p. 274)

- Model fit (f)
  - It is defined in terms of the population root mean squared residual (RMSR) (Preacher & MacCallum, 2002).
  - RMSR = .00, .03, .06, respectively correspond to perfect, good, and fair model fit in the population (Preacher & MacCallum, 2002).
  - Lack of fit of the model in the population will not, on the average, influence recovery of population factors in analysis of sample data, regardless of degree of model error and regardless of sample size (MacCallum, Widaman, Preacher, & Hong, 2001, p. 611).
  - Model fit has little effect on factor recovery. It is probably very rare in practice to find factor models exhibiting simultaneously high communalities and poor fit (Preacher & MacCallum, 2002, p. 157).
  - the differences between (extraction) methods with respect to ability to reproduce the population pattern were generally minor (Velicer & Fava, 1998, p. 243)

Conclusion

- The general rule of thumb of the minimum sample size are not valid and useful.
- What I did with the data I have:
  1. Repeat the method Garson (http://www2.chass.ncsu.edu/garson/pa765/factor.htm#kmo) proposed until the KMO overall is over .60.
  2. Check the communality of each variable. Drop the variables that has the smallest communality, until the communalities of all variables are above .60.
  3. Check the mean value of all communalities to ensure that the mean value is over .07. If not, repeat step 2.
  4. Use Kaiser strategy (dropping all components with eigenvalues under 1.0) and Scree plot to determine the number of factors.
  5. Set the loading size cut-off value as .60, and drop the factors that has less than 3 variables.

Finally, with principal component analysis, I got 4 factors with 32 variables, representing a STV ratio of 1.48:1 (47/32). The overall KMO is .616, the

minimum value of all communalities is .62, the maximum value of communalities is .879, the mean value of communalities is .770 with a standard deviation of .074. There is no cross loading among the 4 factors. Two of the 4 factors each have 5 loaded variables, one has 4 loaded variables, and one has 3 loaded variables. The variable-to-factor ratio is 8 (32/4). I think this can be considered as a moderate to high degree of overdetermination.

> "As long as communalities are high, the number of expected factors is relatively small, and model error is low (a condition which often goes hand-in-hand with high communalities), researchers and reviewers should not be overly concerned about small sample sizes." (Preacher & MacCallum, 2002, p. 160)

> "Strong data" in factor analysis means uniformly high communalities without cross loadings, plus several variables loading strongly on each factor. (Costello and Osborne, 2005, p. 4)

References

- Anderson, T. W., & Rubin, H. (1956). Statistical inference in factor analysis. In J. Neyman (Ed.), *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (pp. 111-150). Berkeley: University of California Press.

- Arrindell, W. A., & van der Ende. J. (1985). An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. *Applied Psychological Measurement, 9*, 165 - 178.

- Barrett, P. T., & Kline. P. (1981). The observation to variable ratio in factor analysis. *Personality Study in Group Behavior, 1*, 23-33.

- Bryant, F. B., & Yarnold, P. R. (1995). Principal components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & R R. Yarnold (Eds.), *Reading and understanding multivariale statistics* (pp. 99-136). Washington, DC: American Psychological Association.

- Cattell, R. B. (1978). *The Scientific Use of Factor Analysis*. New York: Plenum

- Comrey, A. L., & Lee, H. B. (1992). *A first Course in Factor Analysis*. Hillsdale, NJ: Erlbaum.

- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10*(7). Retrieved July 3, 2008 from http://pareonline.net/pdf/v10n7a.pdf.

- Everitt, 1:1. S. (1975). Multivariate analysis: The need for data, and other problems. *British Journal of Psychiatry*. 126, 2S7-240.

- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology, 39*, 291-314.

- Garson, D. G. (2008). *Factor Analysis: Statnotes*. Retrieved March 22, 2008, from North Carolina State University Public Administration Program, http://www2.chass.ncsu.edu/garson/pa765/factor.htm.

- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale,NJ: Erlbaum.

- Guadagnoli, E., & velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological bulletin, 103*, 265-275.

- Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw-Hill.

- Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black,W. C. (1995). Multivariate data analysis (4th ed.). Saddle River, NJ: Prentice Hall.

- Hatcher, L. (1994). *A Step-by-Step Approach to Using the SAS® System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute, Inc.

- Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement, 65*, 202-226.

- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement, 66*, 393-416.

- Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Thousand Oaks, CA: Sage Publications.

- Kline, P. (1979). *Psychometrics and psychology*. London: Acaderric Press.

- Kunce, J. T., Cook, W. D., & Miller, D. E. (1975). Random variables and correlational overkill. *Educational and Psychological Measurement, 35,* 529-534.

- Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method*. London: Butterworth and Co.

- McDonald, R. P., & Krane, W. R. (1977). A note on local identifiability and degrees of freedom in the asymptotic likelihood ratio test. *British Journal ofMathematical and Statistical Psychology, 30,* 198-203.

- McDonald, R. P., & Krane, W. R (1979). A Monte Carlo study of local identifiability and degrees of freedom in the asymptotic likelihood ratio test. *British Journal of Mathematical and Statistical Psychology, 32,* 121-132.

- Marascuilo, 1.. A., & Levin, J. R (1983). *Multivariate statistics in the social sciences*. Monterey, CA: Brooks/Cole.

- MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research, 36*, 611-637.

- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99.

- Noru?is, M. J. (2005). *SPSS 13.0 Statistical Procedures Companion*. Chicago: SPSS, Inc.

- Nunnally, J. C. (1978). *Psychometric theory* (2nd Ed.). New York: McGraw-Hill.

- Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research & Evaluation, 9*(11). Retrieved July 1, 2008 from http://PAREonline.net/getvn.asp?v=9&n=11.

- Preacher, K. J., & MacCallum, R. C. (2002). Exploratory Factor Analysis in Behavior Genetics Research: Factor Recovery with Small Sample Sizes. *Behavior Genetics, 32*, 153-161.

- Rindskopf, D. (1984). Structural equation models: Empirical identification, Heywood cases, and related problems. *Sociological Methods and Research, 13,* 109-119.

- Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3*, 231-251.

No labels