Organizational Research: Determining Appropriate Sample Size in Survey Research

James E. Bartlett, II

Joe W. Kotrlik

Chadwick C. Higgins

The determination of sample size is a common task for many organizational researchers. Inappropriate, inadequate, or excessive sample sizes continue to influence the quality and accuracy of research. This manuscript describes the procedures for determining sample size for continuous and categorical variables using Cochran's (1977) formulas. A discussion and illustration of sample size formulas, including the formula for adjusting the sample size for smaller populations, is included. A table is provided that can be used to select the sample size for a research problem based on three alpha levels and a set error rate. Procedures for determining the appropriate sample size for multiple regression and factor analysis, and common issues in sample size determination are examined. Non-respondent sampling issues are addressed.

Introduction

A common goal of survey research is to collect data representative of a population. The researcher uses information gathered from the survey to generalize findings from a drawn sample back to a population, within the limits of random error. However, when critiquing business education research, Wunsch (1986) stated that "two of the most consistent flaws included (1) disregard for sampling error when determining sample size, and (2) disregard for response and nonresponse bias" (p. 31).

Within a quantitative survey design, determining sample size and dealing with nonresponse bias is essential. "One of the real advantages of quantitative methods is their ability to use smaller groups of people to make inferences about larger groups that would be prohibitively expensive to study" (Holton & Burnett, 1997, p. 71). The question then is, how large of a sample is required to infer research findings back to a population?

Standard textbook authors and researchers offer tested methods that allow studies to take full advantage of statistical measurements, which in turn give researchers the upper hand in determining the correct sample size. Sample size is one of the four inter-related features of a study design that can influence the detection of significant differences, relationships or interactions (Peers, 1996). Generally, these survey designs try to minimize both alpha error (finding a difference that does not actually exist in the population) and beta error (failing to find a difference that actually exists in the population) (Peers, 1996).

However, improvement is needed. Researchers are learning experimental statistics from highly competent statisticians and then doing their best to apply the formulas and approaches

James E. Bartlett, II is Assistant Professor, Department of Business Education and Office Administration, Ball State University, Muncie, Indiana.

Joe W. Kotrlik is Professor, School of Vocational Education, Louisiana State University, Baton Rouge, Louisiana.

Chadwick C. Higgins is a doctoral student, School of Vocational Education, Louisiana State University, Baton Rouge, Louisiana.

they learn to their research design. A simple survey of published manuscripts reveals numerous errors and questionable approaches to sample size selection, and serves as proof that improvement is needed. Many researchers could benefit from a real-life primer on the tools needed to properly conduct research, including, but not limited to, sample size selection.

This manuscript will describe common procedures for determining sample size for simple random and systematic random samples. It will also discuss alternatives to these formulas for special situations. This manuscript is not intended to be a totally inclusive treatment of other sample size issues and techniques. Rather, this manuscript will address sample size issues that have been selected as a result of observing problems in published manuscripts.

As a part of this discussion, considerations for the appropriate use of Cochran's (1977) sample size formula for both continuous and categorical data will be presented. Krejcie and Morgan's (1970) formula for determining sample size for categorical data will be briefly discussed because it provides identical sample sizes in all cases where the researcher adjusts the t value used based on population size, which is required when the population size is 120 or less. Likewise, researchers should use caution when using any of the widely circulated sample size tables based on Krejcie and Morgan's (1970) formula, as they assume an alpha of .05 and a degree of accuracy of .05 (discussed later). Other formulas are available; however, these two formulas are used more than any others.

Foundations for Sample Size Determination

Primary Variables of Measurement

The researcher must make decisions as to which variables will be incorporated into formula calculations. For example, if the researcher plans to use a seven-point scale to measure a continuous variable, e.g., job satisfaction, and also plans to determine if the respondents differ by certain categorical variables, e.g., gender, tenured, educational level, etc., which variable(s) should be used as the basis for sample size? This is important because the use of gender as the primary variable will result in a substantially larger sample size than if one used the seven-point scale as the primary variable of measure.

Cochran (1977) addressed this issue by stating that "One method of determining sample size is to specify margins of error for the items that are regarded as most vital to the survey. An estimation of the sample size needed is first made separately for each of these important items" (p. 81). When these calculations are completed, researchers will have a range of n's, usually ranging from smaller n's for scaled, continuous variables, to larger n's for dichotomous or categorical variables.

The researcher should make sampling decisions based on these data. If the n's for the variables of interest are relatively close, the researcher can simply use the largest n as the sample size and be confident that the sample size will provide the desired results.

> More commonly, there is a sufficient variation among the n's so that we are reluctant to choose the largest, either from budgetary considerations or because this will give an over-all standard of precision substantially higher than originally contemplated. In this event, the desired standard of precision may be relaxed for certain of the items, in order to permit the use of a smaller value of n (Cochran, 1977, p. 81).

The researcher may also decide to use this information in deciding whether to keep all of the variables identified in the study. "In some cases, the n's are so discordant that certain of them must be dropped from the inquiry; . . ." (Cochran, 1977, p. 81).

Error Estimation

Cochran's (1977) formula uses two key factors: (1) the risk the researcher is willing to accept in the study, commonly called the margin of error, or the error the researcher is willing to accept, and (2) the alpha level, the level of acceptable risk the researcher is willing to accept that the true margin of error exceeds the acceptable margin of error; i.e., the probability that differences revealed by statistical analyses really do not exist; also known as Type I error. Another type of error will not be addressed further here, namely, Type II error, also known as beta error. Type II error occurs when statistical procedures result in a judgment of no significant differences when these differences do indeed exist.

Alpha Level. The alpha level used in determining sample size in most educational research studies is either .05 or .01 (Ary, Jacobs, & Razavieh, 1996). In Cochran's formula, the alpha level is incorporated into the formula by utilizing the t-value for the alpha level selected (e.g., t-value for alpha level of .05 is 1.96 for sample sizes above 120). Researchers should ensure they use the correct t- value when their research involves smaller populations, e.g., t-value for alpha of .05 and a population of 60 is 2.00. In general, an alpha level of .05 is acceptable for most research. An alpha level of .10 or lower may be used if the researcher is more interested in identifying marginal relationships, differences or other statistical phenomena as a precursor to further studies. An alpha level of .01 may be used in those cases where decisions based on the research are critical and errors may cause substantial financial or personal harm, e.g., major programmatic changes.

Acceptable Margin of Error. The general rule relative to acceptable margins of error in educational and social research is as follows: For categorical data, 5% margin of error is acceptable, and, for continuous data, 3% margin of error is acceptable (Krejcie & Morgan, 1970). For example, a 3% margin of error would result in the researcher being confident that the true mean of a seven point scale is within $\pm$.21 (.03 times seven points on the scale) of the mean calculated from the research sample. For a dichotomous variable, a 5% margin of error would result in the researcher being confident that the proportion of respondents who were male was within $\pm$5% of the proportion calculated from the research sample. Researchers may increase these values when a higher margin of error is acceptable or may decrease these values when a higher degree of precision is needed.

Variance Estimation

A critical component of sample size formulas is the estimation of variance in the primary variables of interest in the study. The researcher does not have direct control over variance and must incorporate variance estimates into research design. Cochran (1977) listed four ways of estimating population variances for sample size determinations: (1) take the sample in two steps, and use the results of the first step to determine how many additional responses are needed to attain an appropriate sample size based on the variance observed in the first step data; (2) use pilot study results; (3) use data from previous studies of the same or a similar population; or (4) estimate or guess the structure of the population assisted by some logical mathematical results. The first three ways are logical and produce valid estimates of variance; therefore, they do not need to be discussed further. However, in many educational and social research studies, it is not feasible to use any of the first three ways and the researcher must estimate variance using the fourth method.

A researcher typically needs to estimate the variance of scaled and categorical variables. To estimate the variance of a scaled variable, one must determine the inclusive range of the scale, and then divide by the number of standard deviations that would include all possible values in the range, and then square this number. For example, if a researcher used a seven-point scale and given that six standard deviations (three to each side of the mean) would capture 98% of all responses, the calculations would be as follows:

$$\underline{S} = \frac{7 \text{ (number of points on the scale)}}{6 \text{ (number of standard deviations)}}$$

When estimating the variance of a dichotomous (proportional) variable such as gender, Krejcie and Morgan (1970) recommended that researchers should use .50 as an estimate of the population proportion. This proportion will result in the maximization of variance, which will also produce the maximum sample size. This proportion can be used to estimate variance in the population. For example, squaring .50 will result in a population variance estimate of .25 for a dichotomous variable.

Basic Sample Size Determination

Continuous Data

Before proceeding with sample size calculations, assuming continuous data, the researcher should determine if a categorical variable will play a primary role in data analysis. If so, the categorical sample size formulas should be used. If this is not the case, the sample size formulas for continuous data described in this section are appropriate.

Assume that a researcher has set the alpha level a priori at .05, plans to use a seven point scale, has set the level of acceptable error at 3%, and has estimated the standard deviation of the scale as 1.167. Cochran's sample size formula for continuous data and an example of its use is presented here along with the explanations as to how these decisions were made.

$$\underline{n}_0 = \frac{(\underline{t})^2 * (\underline{s})^2}{(\underline{d})^2} = \frac{(1.96)^2(1.167)^2}{(7*.03)^2} = 118$$

Where $\underline{t}$ = value for selected alpha level of .025 in each tail = 1.96
(the alpha level of .05 indicates the level of risk the researcher is willing to take that true margin of error may exceed the acceptable margin of error.)

Where $\underline{s}$ = estimate of standard deviation in the population = 1.167.
(estimate of variance deviation for 7 point scale calculated by using 7 [inclusive range of scale] divided by 6 [number of standard deviations that include almost all (approximately 98%) of the possible values in the range]).

Where $\underline{d}$ = acceptable margin of error for mean being estimated = .21.
(number of points on primary scale * acceptable margin of error; points on primary scale = 7; acceptable margin of error = .03 [error researcher is willing to except]).

Therefore, for a population of 1,679, the required sample size is 118. However, since this sample size exceeds 5% of the population (1,679*.05=84), Cochran's (1977) correction formula should be used to calculate the final sample size. These calculations are as follows:

$$\underline{n} = \frac{\underline{n}_0}{(1 + \underline{n}_0 / \text{Population})} = \frac{(118)}{(1 + 118/1679)} = 111$$

Where population size = 1,679.
Where $\underline{n}_0$ = required return sample size according to Cochran's formula = 118.
Where $\underline{n}_1$ = required return sample size because sample > 5% of population.

These procedures result in the minimum returned sample size. If a researcher has a captive audience, this sample size may be attained easily. However, since many educational and social research studies often use data collection methods such as surveys and other voluntary participation methods, the response rates are typically well below 100%. Salkind (1997) recommended oversampling when he stated that "If you are mailing out surveys or questionnaires, ... count on increasing your sample size by 40%-50% to account for lost mail and uncooperative subjects" (p. 107). Fink (1995) stated that "Oversampling can add costs to the survey but is often necessary" (p. 36). Cochran (1977) stated that "A second consequence is, of course, that the variances of estimates are increased because the sample actually obtained is smaller than the target sample. This factor can be allowed for, at least approximately, in selecting the size of the sample" (p. 396). However, many researchers criticize the use of over-sampling to ensure that this minimum sample size is achieved and suggestions on how to secure the minimal sample size are scarce.

If the researcher decides to use oversampling, four methods may be used to determine the anticipated response rate: (1) take the sample in two steps, and use the results of the first step to estimate how many additional responses may be expected from the second step; (2) use pilot study results; (3) use responses rates from previous studies of the same or a similar population; or (4) estimate the response rate. The first three ways are logical and will produce valid estimates of response rates; therefore, they do not need to be discussed further. Estimating response rates is not an exact science. A researcher may be able to consult other researchers or review the research literature in similar fields to determine the response rates that have been achieved with similar and, if necessary, dissimilar populations.

Therefore, in this example, it was anticipated that a response rate of 65% would be achieved based on prior research experience. Given a required minimum sample size (corrected) of 111, the following calculations were used to determine the drawn sample size required to produce the minimum sample size:

Where anticipated return rate = 65%.
Where $\underline{n}_2$ = sample size adjusted for response rate.
Where minimum sample size (corrected) = 111.
Therefore, $\underline{n}_2$ = 111/.65 = 171.

Categorical Data

The sample size formulas and procedures used for categorical data are very similar, but some variations do exist. Assume a researcher has set the alpha level a priori at .05, plans to use a proportional variable, has set the level of acceptable error at 5%, and has estimated the standard deviation of the scale as .5. Cochran's sample size formula for categorical data and an example of its use is presented here along with explanations as to how these decisions were made.

$$\underline{n}_0 = \frac{(\underline{t})^2 * (p)(q)}{(\underline{d})^2}$$

$$\underline{n}_0 = \frac{(1.96)^2(.5)(.5)}{(.05)^2} = 384$$

Where $\underline{t}$ = value for selected alpha level of .025 in each tail = 1.96.
(the alpha level of .05 indicates the level of risk the researcher is willing to take that true margin of error may exceed the acceptable margin of error).
Where (p)(q) = estimate of variance = .25.
(maximum possible proportion (.5) * 1- maximum possible proportion (.5) produces maximum possible sample size).
Where $\underline{d}$ = acceptable margin of error for proportion being estimated = .05
(error researcher is willing to except).

Therefore, for a population of 1,679, the required sample size is 384. However, since this sample size exceeds 5% of the population (1,679*.05=84), Cochran's (1977) correction formula should be used to calculate the final sample size. These calculations are as follows:

$$\underline{n}_1 = \frac{\underline{n}_0}{(1 + \underline{n}_0 / \text{Population})}$$

$$\underline{n}_1 = \frac{(384)}{(1 + 384/1679)} = 313$$

Where population size = 1,679
Where $\underline{n}_0$ = required return sample size according to Cochran's formula= 384
Where $\underline{n}_1$ = required return sample size because sample > 5% of population

These procedures result in a minimum returned sample size of 313. Using the same oversampling procedures as cited in the continuous data example, and again assuming a response rate of 65%, a minimum drawn sample size of 482 should be used. These calculations were based on the following:

Where anticipated return rate = 65%.
Where $\underline{n}_2$ = sample size adjusted for response rate.
Where minimum sample size (corrected) = 313.
Therefore, $\underline{n}_2$ = 313/.65 = 482.

Sample Size Determination Table

Table 1 presents sample size values that will be appropriate for many common sampling problems. The table includes sample sizes for both continuous and categorical data assuming alpha levels of .10, .05, or .01. The margins of error used in the table were .03 for continuous data and .05 for categorical data. Researchers may use this table if the margin of error shown is appropriate for their study; however, the appropriate sample size must be calculated if these error rates are not appropriate.

Other Sample Size Determination Considerations

Regression Analysis. Situations exist where the procedures described in the previous paragraphs will not satisfy the needs of a study and two examples will be addressed here. One situation is when the researcher wishes to use multiple regression analysis in a study. To use multiple regression analysis, the ratio of observations to independent variables should not fall below five. If this minimum is not followed, there is a risk for overfitting, ". . . making the results too specific to the sample, thus lacking generalizability" (Hair, Anderson, Tatham, & Black, 1995, p. 105). A more conservative ratio, of ten observations for each independent variable was reported optimal by Miller and Kunce (1973) and Halinski and Feldt (1970).

These ratios are especially critical in using regression analyses with continuous data because sample sizes for continuous data are typically much smaller than sample sizes for categorical data. Therefore, there is a possibility that the random sample will not be sufficient if multiple variables are used in the regression analysis. For example, in the continuous data illustration, a population of 1,679 was utilized and it was determined that a minimum returned sample size of 111 was required. The sample size for a population of 1,679 in the categorical data example was 313. Table 2, developed based on the recommendations cited in the previous paragraph, uses both the five to one and ten to one ratios.

Table 1: Table for Determining Minimum Returned Sample Size for a Given Population Size for Continuous and Categorical Data

	Sample size					
	Continuous data (margin of error=.03)			Categorical data (margin of error=.05)		
Population size	alpha=.10 t=1.65	alpha=.05 t=1.96	alpha=.01 t=2.58	p=.50 t=1.65	p=.50 t=1.96	p=.50 t=2.58
100	46	55	68	74	80	87
200	59	75	102	116	132	154
300	65	85	123	143	169	207
400	69	92	137	162	196	250
500	72	96	147	176	218	286
600	73	100	155	187	235	316
700	75	102	161	196	249	341
800	76	104	166	203	260	363
900	76	105	170	209	270	382
1,000	77	106	173	213	278	399
1,500	79	110	183	230	306	461
2,000	83	112	189	239	323	499
4,000	83	119	198	254	351	570
6,000	83	119	209	259	362	598
8,000	83	119	209	262	367	613
10,000	83	119	209	264	370	623

NOTE: The margins of error used in the table were .03 for continuous data and .05 for categorical data. Researchers may use this table if the margin of error shown is appropriate for their study; however, the appropriate sample size must be calculated if these error rates are not appropriate. Table developed by Bartlett, Kotrlik, & Higgins.

Table 2: Minimum Number of Regressors Allowed for Sampling Example

Sample size for:	Maximum number of regressors if ratio is: 5 to 1	10 to 1
Continuous data: n = 111	22	11
Categorical data: n = 313	62	31

As shown in Table 2, if the researcher uses the optimal ratio of ten to one with continuous data, the number of regressors (independent variables) in the multiple regression model would be limited to 11. Larger numbers of regressors could be used with the other situations shown. It should be noted that if a variable such as ethnicity is incorporated into the categorical example, this variable must be dummy coded, which will result in multiple variables utilized in the model rather than a single variable. One variable for each ethnic group, e.g., White, Black, Hispanic, Asian, American Indian would each be coded as 1=yes and 2=no in the regression model, which would result in five variables rather than one in the regression model.

In the continuous data example, if a researcher planned to use 14 variables in a multiple regression analysis and wished to use the optimal ratio of ten to one, the returned sample size must be increased from 111 to 140. This sample size of 140 would be calculated from taking the number of independent variables to be entered in the regression (fourteen) and multiplying them by the number of the ratio (ten). Caution should be used when making this decision because raising the sample size above the level indicated by the sample size formula will increase the probability of Type I error.

Factor Analysis. If the researcher plans to use factor analysis in a study, the same ratio considerations discussed under multiple regression should be used, with one additional criteria, namely, that factor analysis should not be done with less than 100 observations. It should be noted that an increase in sample size will decrease the level at which an item loading on a factor is significant. For example, assuming an alpha level of .05, a factor would have to load at a level of .75 or higher to be significant in a sample size of 50, while a factor would only have to load at a level of .30 to be significant in a sample size of 350 (Hair et al., 1995).

Sampling non-respondents. Donald (1967), Hagbert (1968), Johnson (1959), and Miller and Smith (1983) recommend that the researcher take a random sample of 10-20% of non-respondents to use in non-respondent follow-up analyses. If non-respondents are treated as a potentially different population, it does not appear that this recommendation is valid or adequate. Rather, the researcher could consider using Cochran's formula to determine an adequate sample of non-respondents for the non-respondent follow-up response analyses.

Budget, time and other constraints. Often, the researcher is faced with various constraints that may force them to use inadequate sample sizes because of practical versus statistical reasons. These constraints may include budget, time, personnel, and other resource limitations. In these cases, researchers should report both the appropriate sample sizes along with the sample sizes actually used in the study, the reasons for using inadequate sample sizes, and a discussion of the effect the inadequate sample sizes may have on the results of the study. The researcher should exercise caution when making programmatic recommendations based on research conducted with inadequate sample sizes.

Final Thoughts

Although it is not unusual for researchers to have different opinions as to how sample size should be calculated, the procedures used in this process should always be reported, allowing the reader to make his or her own judgments as to whether they accept the researcher's assumptions and procedures. In general, a researcher could use the standard factors identified in this paper in the sample size determination process.

Another issue is that many studies conducted with entire population census data could and probably should have used samples instead. Many of the studies based on population census data achieve low response rates. Using an adequate sample along with high quality data collection efforts will result in more reliable, valid, and generalizable results; it could also result in other resource savings.

The bottom line is simple: research studies take substantial time and effort on the part of researchers. This paper was designed as a tool that a researcher could use in planning and conducting quality research. When selecting an appropriate sample size for a study is relatively easy, why wouldn't a researcher want to do it right?

References

Ary, D., Jacobs, L. C., & Razavieh, A. (1996). *Introduction to research in education*. Fort Worth, TX: Harcourt Brace College Publishers.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: John Wiley & Sons.

Donald, M. N. (1967). Implications of non-response for the interpretation of mail questionnaire data. *Public Opinion Quarterly, 24*(1), 99-114.

Fink, A. (1995). *The survey handbook*. Thousand Oaks, CA: Sage Publications.

Hagbert, E. C. (1968). Validity of questionnaire data: Reported and observed attendance in an adult education program. *Public Opinion Quarterly, 25*, 453-456.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1995). *Multivariate data analysis* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Halinski, R. S. & Feldt, L. S. (1970). The selection of variables in multiple regression analyses. *Journal of Educational Measurement, 7*(3), 151-158.

Holton, E. H., & Burnett, M. B. (1997). Qualitative research methods. In R. A. Swanson, & E. F. Holton (Eds.), *Human resource development research handbook: Linking research and practice*. San Francisco: Berrett-Koehler Publishers.

Johnson, P. O. (1959). Development of the sample survey as a scientific methodology. *Journal of Experiential Education, 27*, 167-176.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement, 30*, 607-610.

Miller, D. E., & Kunce, J. T. (1973). Prediction and statistical overkill revisited. *Measurement and Evaluation in Guidance, 6*(3), 157-163.

Miller, L. E., & Smith, K. L. (1983). Handling nonresponse issues. *Journal of Extension, 21*, 45-50.

Peers, I. (1996). *Statistical analysis for education and psychology researchers*. Bristol, PA: Falmer Press.

Salkind, N. J. (1997). *Exploring research* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wunsch, D. (1986). Survey research: Determining sample size and representative response. *Business Education Forum, 40*(5), 31-34.