Sample Size Matters

What type of cook (researcher) are you?

Excerpted from: *Dissertation and Scholarly Research: Recipes for Success*, 2012 Edition (Simon & Goes, 2012)

Find this and many other dissertation guides and resources at www.dissertationrecipes.com

In research size matters – when it comes to sample selection.

There is no such thing as a *sample population* - A population is the entire group of individuals or items that share one or more characteristics from which data can be gathered and analyzed. A sample is a subset of the population. In the dissertation we discuss the population first and then explain how the sample was drawn from the population.

Read more: http://www.investorwords.com/3738/population.html#ixzz1JhksqO9Z

How you sample also matters – random, convenience – want to have a representative sample of the target population by whatever affordable means are available to us, we need to give serious consideration to sample size. This is a case where size matters (pun intended). Why?

Quantitative Studies

It is all about precision, tolerance for risk and cost. For samples smaller than 1000, **we always have to think about how confident we want to be that estimates are within a particular range (level of confidence and risk), and how small we want that range to be (level of precision)**. Unfortunately, they go in opposite directions. Higher levels of confidence require greater ranges (margins of error) in small sample sizes.

The Gold standard in quantitative studies is when we can be 95% confident that the true estimate for a variable is within a certain range. For example in a sample of 300 if we want to be 95% sure we have the correct solution, we may need a range of +/- 5 points. If we need to have a margin of error of 2 points within the correct answer. If we want a smaller margin of error, in an attempt to gain more precision, with the same sample size, we would either have to sacrifice certainty and may need to accept a 90% confidence or enlarge our sample size.

At the end of the day, when it comes to sample size, **you need to decide what it is more important to you, certainty or precision, and what your tolerance for risk is**, especially if your time is limited and your population is small.

A sample size and margin of error confidence calculators are found at:
http://relevantinsights.com/research-tools .

Sampling

Things to consider when conducting a sample:

Data carelessly collected may be so completely otiose that no amount of statistical torturing can salvage them - *garbage in garbage out*! There are four points to consider when collecting data.

1. Ensure the **sample size** is large enough for the required purpose.
2. When possible **do the measuring yourself.** Self-reporting tends to lead to erroneous politically correct, wishful thinking or disproportionate rounding. Distorted data.
3. How will data be collected? Mail surveys tend to get lower responses; personal interviews are time consuming and expensive but may be needed for complete data. Telephone interviews are relatively efficient and relatively inexpensive.
4. Ensure that the sample is **representative** of the population.

Sampling Methods

The following are common means of sampling:

1. **Random Sampling** (representative, proportionate) - Members of the population are selected in such a way that each has an equal chance of being selected. This needs much planning to avoid haphazard sampling. Telephone directories e.g. leave out unlisted numbers. In LA 42.5% of the telephone numbers are unlisted. Computer generated numbers are better. There is about a 20% refusal rate for telephone intervals, which could bias the study.
2. **Stratified sampling** - Members of the population are subdivided into at least two different sub populations or strata e.g. gender. Samples are then drawn from each stratum.
3. **Systematic sampling** - Similar to random sampling. Members of the population are listed in some type of roster and then every kth (e.g. 20^{th}) element is chosen.
4. **Cluster sampling** - Members of the population are divided into sections (or clusters), randomly select a few of those sections and then choose all the members for the selected sections. For example, in conducting a pre-election poll we could randomly select 30 election precincts and survey all people from those precincts. The results may need to be adjusted to correct for any disproportionate representation of groups. Used extensively by government and private research organizations.
5. **Convenience sampling** - use the results that are readily available. Sometimes this is quite good - e.g. a teacher wanting to know about left-handed students needs might use the students in her class who fit this profile; however sometimes it may be seriously biased.
6. **Snowball Sampling** -. In this method of sampling selection, one participant gives the researcher the name of another participant, who in turn provides the name of a

third, and so on. This is an especially useful technique when the researcher wants to survey or interview people with unusual characteristics who are likely to know one another.

The *size* of the survey may be decided with statistical precision. A major concern in choosing a sample size is that it should be large *enough* so that it will be representative of the population from which it comes and from which you wish to make inferences about. It ought to be large enough so that important differences can be found in subgroups such as men and women, democrats and republicans, groups receiving treatment, and control groups, etc. Two major issues to be considered when using statistical methods to choose sample size are concern with **sampling error and confidence levels**.

Sampling error

Some small differences will almost always exist among samples and between them and the population from which they are drawn. One approach to measuring sample error is to report the standard error of measurement, which is computed by dividing the population standard deviation (if known) by the square root of the sample size. Minimizing sampling error helps to maximize the samples' representativeness.

Example: If the Stanford-Binet .IQ test (where standard deviation is 15) were administered to 100 participants then the standard error of the mean would be: 15/10 or 1.5

The Error of the Estimate, Confidence Intervals, and Sampling Size

The following were high temperatures reported during the month of January in San Diego County:

65	61	64	59	60	58	64	69	70	69	63
69	68	63	58	54	50	57	67	68	69	
66	74	68	66	59	56	43	54	58	61	

Let us assume that this was a pretty "typical" January and we want to use this sample to estimate the mean high temperature for all the January's in San Diego County to include in our tourism guide. Since x-bar is the best estimate of the population mean μ we have:

x-bar = 62.8 s = 6.06

The maximum error of the estimate of μ is:

$$E = z_{\partial/2}\frac{\sigma}{\sqrt{n}}$$

Thus for a 95% confidence level we can determine the "true mean" temperature of San Diego in January by first finding the value of E:

$$E = 1.96\frac{6.06}{\sqrt{31}} = 2.1$$

Thus, with a 95% confidence level, the true mean temperature in San Diego in January is between 60.7 and 64.9, or the true mean is 62.8 +/- 2.1.

Confidence Intervals

A confidence interval, or interval estimate, is a range (or an interval) of values that is likely to contain the true value of the population. It is associated with the degree of confidence, which is a measure of how certain we are that the interval contains the population parameter. The definition of the degree of confidence uses alpha (lower case) to describe the probability that corresponds to a small area under a distribution curve. The critical value separates alpha from the rest of the distribution. For a 95% confidence level when we are testing a mean value, for sample size over 30, that alpha is 1.96.

To find the confidence interval we first find the margin of error, E. For a 95% confidence level:

$$E = 1.96(\sigma / \sqrt{n})$$

Then all we need to do is Add and subtract E to the mean value we get from our sample size.
Now what does this mean? Or. What’s it all about alphie:)?

If you do a good study (random sample, pay attention to what you learnt here) and found a mean IQ of students at college X to be 115, with a standard deviation of 15, surveyed 25 random students then you can be 95% sure the true mean of this population lies between 115 +/-1.96x3 or 115+/-5.88

Confidence Levels You will need to decide how “confident” you wish, or need, to be that your sample is representative of the population. Frequently, the 95% confidence level is chosen. This means that there is a 95% change that the sample and the population will look alike and 5% chance that they will not.

Note: We can use the following formula to determine the sample size necessary to discover the “true” mean value from a population.

$n = \left[\frac{\bar{z}\sigma}{E} \right]^2$ $\bar{z}$ corresponds to a confidence level(found on a table or computer program).
Some common z values are 1.645 or 1.96, which might reflect a 95% confidence level (depending on the statistical hypothesis under investigation), and 2.33, which could reflect a 99% confidence level in a one tailed-test and 2.575 for a two-tailed test□□□□ is the standard deviation, and E is the margin of error.
Example: If we need to be 99% confident that we are within 0.25 lbs of a true mean weight of babies in an infant care facility, and s = 1.1, we would need to sample 129 babies:

$n = [2.575 (1.1)/ 0.25]^2 = 128.3689$ or 129.

Note: A formula that we can use to determine the sample size necessary to test a hypothesis involving percentages is: [these make a lot more sense later on in the course - most of these numbers come from a table.]

$n = (z/e)^2 (p) (1\text{-}p)$

where n = sample size, z= standard score corresponding to a certain confidence level. We represent the proportion of sampling error by "e," and the estimated proportion or incidence of cases by "p."

Example 1: It is estimated that 87% of North American adults have some level of mathematics anxiety. To compare and contrast math anxious and non- math anxious adults with a 90% confidence level (z = 1.96) we will need to test 44 North American adults $n = (1.96/.10)^2 (.87) (0.13) = 384.16 (0.1131) = 44$

Example 2: Suppose we wish to find the true mean IQ of the students at a school and be 99% that we are within 2 units of the true mean. Knowing the standard deviation is 15, we can determine how large our sample needs to be by performing a bit of Algebra:

$$E = z_{\partial/2} \frac{\sigma}{\sqrt{n}}$$

multiply each side by $\sqrt{n}$. This gives us:

$$\sqrt{n}E = Z_{\partial/2}\sigma$$

Solving for n we get:

$$n = \left[\frac{z_{\partial/2}}{E}\right]^2$$

Plugging in the values: we get $n = \left[\frac{2.575x15}{2.0}\right]^2$

Or $(19.3125)^2 = 373$

Thus we need at least 373 randomly selected IQ scores to be 99% confident that our sample mean is within 2.0 units of the true mean.

Exercises:

3. A sample of 35 skulls is obtained for Egyptian males who lived around 1859 B.C. The maximum breath of each skull is measured with the result that x = 134.5 and s = 3.48. Using these sample results, construct a 95% confidence interval for the population mean μ.
4. A psychologist had developed a test for spatial perception and she wants to estimate the mean score of male pilots. How many people must she test if she wants the sample mean to be in error by no more than 2.0 points with 95% confidence? An earlier study suggest that σ = 21.2

Answer: 1. 133.3 mm < μ, 135.7mm
2. 432

A quick guide:

Linda Suskie in her book Questionnaire Survey Research What Works, provides a guide to determine how many people you can survey based on sampling error alone, when non standardized instruments are used. A sampling error of 5%, would mean that you can be 95% sure that you have the correct statistics. The table below is from her book.

Random Sample Size	Sample Error
196	7%
264	6%
364	5%
1,067	4%
2,401	2%
9,604	1%

In most studies 5% sampling error is acceptable. Below are the sample sizes you need from a given population. These numbers assume a 100% response rate

Population Size	Sample Size
10,000	370
5,000	357
2,000	322
1,000	278
500	217
250	155
100	80

Gay (1996, p.125) has suggested the following guidelines for selecting a sample size
For small population (N<100), there is little point in sampling. Survey the entire population. If the population size is around 500, 50% of the population should be sampled. If the population size is 1,500, 20% should be sampled. Beyond a certain point (at approximately N=5,000), the population size is almost irrelevant, and a sample size of 400 will be adequate. Thus, the larger the population the smaller the *percentage*, one needs to get a representative sample.

Researchers should, thus, consider the following when determining a sample size for a quantitative study:

Statistical validity. The larger the sample, the more accurate the results. (http:www.bourgetresearch.com/resources_sample.htm)

Characteristics of the sample. Larger samples needed for heterogeneous populations; smaller samples needed for homogeneous populations (Leedy & Ormond, 2001, p. 221).

Cost of the study. Minimum number of participants needed to produce valid results. (http://www.oandp.org/jpo/library/1995_04_137.asp)

The ability to generalize results from the study to the larger population.

Knowledge of the behavior of the data (variance and standard deviation).

Methods available for determining sample size (http://www.oandp.org/jpo/library/1995_04_137.asp)

Statistical power needed. Larger samples yield greater the statistical power. In experimental research, power analysis is used to determine sample size (requires calculations involving statistical significance, desired power, and the effect size)

Confidence level desired (reflects accuracy of sample, Babbie, 2001)

Purpose of the study. Merriam (1998) stated that "selecting the sample is dependent upon the research problem" (p. 67).

Availability of the sample. Convenience samples used when researcher is limited to available groups (Creswell, 2003, p. 164)

Availability of sampling frames (a sampling frame is a list from which a probability sample is selected, e.g. telephone list, Babbie, 2001).

Most importantly, the sample size must represent the characteristics or " behavior" of the larger population.

Check out: http://www.surveysystem.com/ssformu.htm

Zoomerang offers a free service to compute sample size: http://panel.zoomerang.com/ZS_Sample_Size_Calculations.html?gclid=CMrq5O-k8qcCFQEHbAodASdNZw

Power Analysis

Power analysis is the process of determining the sample size for a research study. It is the probability of detecting an effect when it exists. In most cases, power analysis involves a number of simplifying assumptions, in order to make the problem tractable, and running the analyses numerous times with different variations to cover all of the contingencies.

G*Power is a software program capable of determining the sample size needed for tests of two independent proportions as well as for tests of means. To begin, the program should be set to the z family of tests, to a test of proportions, and to perform the 'A Priori' power analysis necessary to identify sample size.

For more on G-power see: http://www.ats.ucla.edu/stat/gpower/indeppropor1.htm

http://www.jeremymiles.co.uk/misc/power/

There are 3 types of power analysis: a priori, post hoc, and compromise. Compromise power analysis is very complex is rarely used, and slightly controversial.

A Priori Power Analysis

Ideally, power analysis is carried out a priori, that is, during the design stage of the study. A study can conclude that a null hypothesis should be rejected or not.

Given the three factors alpha, sample size and effect size, a fourth variable can be calculated, called beta. Where alpha is the probability of a type I error (i.e. rejection of a correct null hypothesis) beta is the probability of a type II error (acceptance of a false null hypothesis)

Post Hoc Analysis

Whereas a priori analysis is done before a study has been carried out, post-hoc analysis is done after a study has been carried out to help to explain the results if a study which did not find any significant effects

Alternate Theories for Different Studies

Multiple Regression

The determination of an adequate sample size for a correlational study can be calculated based on variables The formula (Sample Size = 104 + m) recommended by Tabachnick and Fidell (2001).

B. Gather an Adequate Sample
   - --When the number of events is small, the observed R^2 tends to increase. As the number of predictor variables approaches the number of events in the sample, the R^2 also nears 1.0 even if the associations actually are nonexistent. Hence, researchers using multiple regression correlation with small samples may mislead themselves into thinking that they have identified substantial effects when, in fact, their large R coefficients are largely artifacts.
   - --One guideline advises researchers to have 104 events plus the number of independent variables if **they wish to test regression coefficients** (Tabachnick & Fidell, 2001, p. 117).
   - --Another popular rule of thumb is that a sample must include ***at least*** **15 events per predictor variable** (J. P. Stevens, 2002, p. 143).
   - --The size of the population multiple correlation coefficient (ρ^2) makes a difference. One study (C. Park & Dudycha, 1974) found that the sample size required to keep R^2 from deviating from R^2 corrected for shrinkage increases as

the population ρ^2 to be detected decreases. For instance, with four predictor variables, when the population ρ^2 is .50, the required sample is 66 (which equals 16.5 events per predictor variable). Similarly, when the population ρ^2 is .25, the required sample is 93 (which equals 23.25 events per predictor variable).

Qualitative Studies

Although there is no total agreement on the sample size for a qualitative study, there are several guidelines that can be followed.

1. Creswell (2002) recommends that 3-5 participants be used for case study research, along with other types of data.
2. With respect to phenomenological studies, sample size recommendations range from 6 (Morse & Chung, 2003) to 10 (Creswell).
3. For grounded theory research, sample size guidelines have ranged from 15-20 participants (Creswell) to 20-30 participants (Morse & Chung, 2003).
4. With regard to ethnographic Morse and Chung (2003) suggests that 30-50 interviews need to be conducted.
5. The completion of data collection and the resulting sample size may be the result of data saturation. After enough data have been collected to determine themes or categories, the researcher may decide that if the next few participants' experiences are captured by the existing themes or categories, the phenomenon of study is saturated or complete. This means that the researcher's construct represents the phenomenon of study, and no further data collection is necessary

References

Babbie, E. (2001). The practice of social research. Australia: Wadsworth Thomson Learning.

Bourget Research Group. (n.d.). Resources: Sample size considerations. Retrieved April 5, 2004, from http:www.bourgetresearch.com/resources_sample.htm

Creswell, J..W. (2003). Research design: Qualitative, quantitative, and mixed methods approaches. (2nd ed.). Thousand Oaks, CA: Sage Publications.

Gay, L. R. (1996). Educational Research: Competencies for Analysis and Application. Beverly Hill, CA: Sage Publications.

Leedy and Ormond (2001). Practical research. Upper Saddle River, NJ:Merrill Prentice Hall.

Lunsford and Lunsford. (n.d.). Research forum. Retrieved April 5, 2004, from the American Academy of orthotists and prosthetists: from http://www.oandp.org/jpo/library/1995_04_137.asp

Merriam, S. (1998). Qualitative research and case study applications in education. San Francisco: Jossey-Bass Publishers.

Morse, J. M. & Chung, S. E. (2003). Toward holism: The significance of methodological pluralism. *International Journal of Qualitative Methods, 2* (3). Article 2. Retrieved [INSERT DATE] from http://www.ualberta.ca/~iiqm/backissues/ 2_3final/ html/morsechung.html

Simon, M. (2006) Dissertation and Scholarly Research: Recipes for Success. Dubuque, IA: Kendall-Hunt publishing.

Suskie, Linda (1996) Questionnaire Survey Research: What works 2nd edition. Washington, D.C.: Assn for International Research. ISBN: 1882393058

http://www.tufts.edu/~gdallal/sizenotes.htm

Tabachnick, B. G. and Fidell, L. S. (2001). *Using Multivariate Statistics* (4th Ed.). Boston: Allyn and Bacon.