SAMPLING FOR ASSESSMENT: STRATEGIES AND TIPS

Office of Assessment of Teaching and Learning, Washington State University, December 2014 *Briana Keafer Morrison*

CONTENTS

Introduction

Sampling is used when assessing an entire group of students (a census) is overly difficult or time consuming. Sampling is particularly useful when assessment includes reviewing lengthy papers or other artifacts. A sampling strategy for surveys may also increase response rate, as over time survey fatigue is reduced. This document provides tips for sampling strategies and sampling size and examples.

Sampling Strategies

A sampling strategy is used to identify a subgroup that effectively represents the population as a whole. Below are two sampling strategies: randomized sampling and stratified sampling.

Randomized Sample

In a randomized sample, every student in the population (e.g., all seniors in your program) has an equal chance of being chosen to participate or having their paper selected for review. There are several ways to collect a random or semi-random sample. One method is to select two random numbers; the first number tells where to start in a list of students or papers and the second random number indicates how many to count before selecting a second student for the sample. For example, if you chose 32 and 8, you would start with the 32nd student on a list and count down the list including every eighth student in the sample. A second method is to use a computer program that randomly selects respondents from the pool.

For example, if you wanted to determine how satisfied students were with your degree program, it would not be sufficient to sample and survey every fifth student that came into the main department office. That technique would measure only the satisfaction of students who choose to come into the office.

Stratified Random Sample

Taking a stratified random sample involves dividing the population into sub-categories, and randomly selecting from each sub-category. A stratified random sample is taken when you want to ensure that the sample includes groups of interest (such as students from every option). To stratify a population, you first need to decide what sub-categories are of interest and in which you suspect there may be substantial differences. Next, determine the percentage of each group in the overall population, and include the same percentages of each group in the sample. For instance, if 25% of students are in option A and 75% are in option B, then the sample should include 25% students from option A and 75% from option B.

Considerations for Representation

In general, assessment data is collected locally to make local decisions. Since results do not need to be generalizable outside of a local context, the most important consideration is whether or not a sample is a representative of the entire local population. For example, does the sample include high-achieving and low-achieving students? Does the sample only include students from one degree option or does it include a representative sample from all options? If needed, sampling can be stratified to ensure inclusion of key groups.

Choosing a Sample Size

How many students need to respond to a survey for a program to be reasonably confident about the results of the survey? How many papers does a program need to collect to assess degree program learning outcomes? To answer these questions, consider the following.

Anytime you sample, the results will have some margin of error. The level of error is measured as a percentage, as is the level of confidence. The level of confidence represents how confident you feel about your error level. For example, if you have a 95% confidence interval with an error level of 10%, you are saying that if you were to conduct the same survey 100 times, the results would be within +/- 10% of the first time you ran the survey 95 times out of 100.

When deciding the sample size, consider how much sampling error can be tolerated and what confidence level is acceptable. In other words, how much precision is needed? This may change based on the types of decisions results from the assessment measure may guide. High stakes decisions, for instance, may require a lower rate of sampling error than lower stakes decisions. The following table can help you determine the level of sampling error associated with certain sample sizes. If you are using a stratified sample, use the smallest sub-category in your calculations. To calculate for other population sizes, use the following formula, or see the calculator at http://www.custominsight.com/articles/random-sample-calculator.asp.

$$Ns = \frac{(Np)(p)(1-p)}{(Np - 1)(B/C)^2 + (p)(1-p)}$$

Ns = sample size needed for desired level of precision
Np = size of population
P = proportion of population expected to choose one of the two response categories
B = acceptable amount of sampling error; .03 = +/- 3% of the true population value
C = Z statistic associated with the confidence level; 1.96 corresponds to the 95% level

Steps in Choosing a Sample Size

1. Determine the size of the target student group.
2. Decide what level of error is acceptable.
3. Decide what level of confidence is acceptable.
4. Determine the sample size.

Note: The sample size calculations here pertain to clean, useable data that results from assessment work. When planning, it is recommended that you include a few additional students or papers in your sample, so that you will be able to deal with incomplete data and unexpected situations (e.g., a student paper is missing pages, a rater skips a portion of the rubric, technology glitches, etc.) and take into account response rates. With surveys, you also need to consider non-response error, or the error that results when those who respond to a survey differ from those who do not, in some way that is relevant to your analysis.

Completed sample sizes needed for various population sizes, 3 levels of precision

Population size	Sample size for a 95% confidence interval		
	±20% Sampling Error	±10% Sampling Error	±5% Sampling Error
25	13	20	24
50	16	33	44
100	20	49	80
200	22	65	132
400	23	78	196
800	23	86	260
2,000	24	92	322
4,000	24	94	351
6,000	24	95	361

Figure 1: Completed sample sizes for 95% confidence interval

Glossary

Confidence interval: a range of values around a statistic that contain, with a certain probability, the true value of that statistic
Population: a collection of students to which one wants to generalize a set of findings
Sampling error: the error caused by observing sample instead of the whole population

Examples

The following examples can provide illustration of how to apply sampling strategies and decide sample size.

Surveys

Department X decides to use a senior exit survey to get a sense of student perception of their degree program. Over two semesters, the department has about 400 graduates. Concerned that students may be inundated with surveys and thus not respond to the department's survey, they decide to send the survey to a sample of graduates, indicating the importance of the survey and perhaps provide incentives to participating students.

Department X has two questions. First, what strategy will they use to sample students? Since the Department offers degrees on two campuses, they would like to be able to disaggregate survey results by campus and feel confident that the sample is representative. Consequently, they decide on a stratified random sample. Campus A has about 75% of graduates and Campus B has 25%, so they plan to take three quarters of their sample from Campus A and one quarter from Campus B. They use a computer program to randomly select students from each campus.

Second, how many students do they need to include in their sample? Since they are not aware of any high stakes decisions that may be made based on survey results, they are willing to accept possibly lower precision and decide to accept a 10% sampling error with a 95% confidence interval. Since they know that they will disaggregate by campus, they use the smallest population size of the two groups (100 seniors at Campus B) to consider the acceptable confidence interval and sampling error rates. They need to receive completed surveys from 49 students at Campus B, and because campus A has three times as many students, they will need 147 students from Campus A. Since they are concerned about response rate, Department X decides to invite 25% more students to the survey. (Note: The number by which to increase sample size depends on the context in which the survey is offered including what motivations students have to respond.) After students respond, the department also considers non-response bias by comparing the students who responded to those who did not, using available and relevant demographic information.

Student Papers

Department Y wants to use papers from their writing course to consider their communication degree learning outcome. The course has about 100 students over the course of a year, and each student completes a final paper. The Department wants to take a sample of papers that is representative; they decide that they can accept a sampling error of 10% with a 95% confidence interval. In other words, in about 19 of 20 times a random sample of papers is taken, the true population value (or average score they would have if evaluating all papers) is within 10 percentage points of the sample estimate.

Looking at the table on page 2, they see that they need to sample at least 49 papers. They decide to take an even number of 50 papers and round up to 52 (to allow for unexpected problems). To make sure it is representative, they use a computer program to randomly select 52 papers from their total number of 100 papers.

Other

As part of a 400 level course, all of Department Z's approximately 50 graduating seniors complete a substantial research project. The project is regularly graded according to course outcomes and requirements. Department Z would also like to use these projects for program assessment and have developed a rubric that evaluates students according to program-level learning outcomes.

Due to limited time and resources, Department Z knows that they cannot evaluate more than 20 projects. Given their limitations, they will sample 16 projects and accept a 20% sampling error with a 95% confidence interval. To ensure that each student project has an equal chance of being included in the sample, Department Z randomly chooses 18 projects (to allow for unexpected problems). They do this by choosing 2 random numbers between 1 and 50 (total number of students); the first number tells where to start in a list of students or papers and the second random number indicates how many to count before selecting a second student for the sample.

Reference

Dillman, D.A. (2000). *Mail and Internet Surveys: A Tailored Design Method, 2nd Ed.* New York, NY; John Wiley & Sons, Inc.