G3658-3

Program Development and Evaluation

Sampling

Ellen Taylor-Powell
Program Development and Evaluation Specialist

May 1998

Adapted from *Sampling in Extension*

Ellen Taylor-Powell
Texas Agricultural Extension Service
The Texas A&M University System
College Station, Texas, 1988

■ ■ ■ TABLE OF CONTENTS

A sample is a portion or a subgroup of a larger group called a population. The two standard ways to draw a sample are by probability and nonprobability sampling. If the evaluation's purpose is to generalize to the whole group, or to provide a statistical basis for saying that the sample is representative, a probability sample is appropriate. If the aim of the evaluation is to learn about individuals or cases for some purpose other than generalizing to the population, or if random selection is not feasible, then nonprobability sampling is appropriate.

Sampling is not always necessary. When a population is small, you may choose to survey all its members. Decisions about sampling depend upon population size, what you want to know, and the resources available.

The sample: who and what?

In Extension, we are usually interested in people—producers, families, youth. We can take samples of almost anything, however, such as acres, result demonstrations, livestock, lakes or 4-H records.

To identify who or what to include in a sample, start by clearly defining the population of interest. The term *population* does not necessarily refer to all the people in the town, county or state. Rather, it refers to the group or units of interest (married couples, livestock owners, watersheds, participants in an extension program, or recipients of a newsletter series) who are found in the geographic area of interest (town, county, region) during the time of interest (since 1995, during 1990-1995, and so on).

Clearly defining the population is the essential first step in selecting a sample. This process includes three parts:

- identifying the group of interest;
- naming the geographic area where the group is found; and
- indicating the time period of interest, as necessary.

Identifying the population from which to select the sample depends upon what you want to know. If the purpose is to show the impact of a financial management program on family financial well-being, *all participants in the program* would be the population of interest. If the purpose is to assess interest among the county's single parents in a prospective financial management program, the population is *all single parents in the county*.

Sampling for generalizability: probability sampling

If you want to **generalize** from the sample to the population, random sampling is necessary. A **probability sample** provides for **random selection**. This means that each unit has the same or a known chance of being selected so that it is possible to confidently make estimates about the total population based on the sample results. Random sampling increases the likelihood that the information collected is representative of the entire group.

The use of probability samples requires that each individual in the population be identifiable on a list or at a location. This so-called **sampling frame** must meet several criteria. It needs to be:

1. accurate—including only those individuals of interest;
2. complete and current—including all individuals of interest;
3. free of duplicate names; and
4. absent any pattern in the way the names are listed.

Information gained from a sample can be generalized only to the larger population from which the sample is taken. Conclusions drawn from a representative sample of teenagers in Adams County may be generalized to all teenagers in Adams County. They should not be generalized to teenagers in other counties or across the state.

Strategies

Simple random sample

In a simple random sample, each unit of the population has an equal chance of being selected. This requires a *complete* list of the total population—all participants in the program, all homeowners, all county residents, and so on. One of the following methods may be selected for drawing a random sample:

1. When the population is small, numbers or names can simply be drawn from a hat. Record the number or name that is drawn. Put the slip back into the container and continue drawing until the required sample size is obtained. If the same number is drawn again, disregard it, put it back in the container and continue.
2. For larger populations and to ensure equal selection opportunities, a random number table may be used such as those found in most statistics textbooks (see example, Appendix 1). After assigning consecutive numbers to the names on the population list, a number is randomly selected on the table (close eyes and point). Proceed either vertically or horizontally. For example, if the total population is a three-digit number (100), use the last three digits of the random table number that correspond to a number on the population list, (any number between 1 and 100) until the needed sample size is obtained. The corresponding names on the population list form the sample.

Systematic sample

This process is easier and may be substituted for random sampling with large populations. Sampling starts from some randomly selected point on the population list (again, close your eyes and point) and proceed by selecting individuals or units at intervals thereafter. The interval is determined by dividing the total number on the population list by the required sample size. For example, if a sample of 200 is desired from a population of 800 people, the interval would be 800/200 = 4. Randomly select a starting point and select every fourth name on the list thereafter. If you come to the end of the list and don't have the required sample size, go to the beginning of the list and keep counting and choosing until the required number is obtained.

A potential problem with this method is the order in which names appear on the population list. Individuals may be listed in a certain order such as husbands' names preceding wives'; groupings by age; groupings by times and days of the week, or some other way. Even alphabetical listings may result in biased selection when ethnic names are alphabetically clustered, resulting in under- or overrepresentation. Systematic sampling works best if the names on the list are randomized first.

Stratified sample

In stratified sampling, the total population is divided into separate groups (strata) which differ along selected characteristics such as gender, age, size of operation, or geographical location. A random sample is drawn from each subgroup or stratum.

This method is especially appropriate when particular subgroups are known to vary or when some characteristic, such as age, is known to be related to the outcome of interest. For example, we might want to draw a sample from each age grouping in the program to ensure that each is fully represented. To do so, determine what percentage each age group makes up of the total number of participants. Use this percentage to determine the proportion of the sample that must come from each age group. Then, randomly select an equal number of participants from those enrolled in the corresponding age group. The example that follows shows the sample size for each age stratum in a nutrition program where the total number of participants is 250 and the desired sample size is 154.

Enrollment and stratified sample of a nutrition program

Age	20–30	31–40	41–50	50+	TOTAL
Number enrolled	50	112	68	20	250
% of total group	20%	45%	27%	8%	100%
Sample size	31	69	42	12	154

Sample size

When your purpose is to generalize or show representativeness, the size of the sample becomes an issue. How many individuals or units should you include so that the sample provides a fair representation? Often, a proportion of the population is used to determine the sample size. For example, you might sample 10% of all producers in the county, or 20% of all program participants. Whether you use proportions depends on the size of the total population. In a program with 200 participants, a 20% sample would produce a sample of 40 people—underrepresenting the population since there is a large chance for respondent variability. On the other hand, a 20% sample of 50,000 county inhabitants produces a sample of 10,000—an unnecessary oversampling. You do not gain much more precision than you would using a sample of 400. You can use the table in Appendix 2 as a guide to recommended sample sizes for various population sizes. Notice that the smaller the population, the larger the sample size. This is because smaller populations exhibit greater variability.

Variability, known as **sampling error,** exists within any sample. The larger the sample, the smaller the sampling error. But no sample will yield exactly the same information as if all people in the population had been included. Information collected from a sample is used to make **estimates** about the population. The intent is to produce as close an approximation of the population as possible within the constraints of time and money.

To estimate how closely the sample approximates the population, two parameters are set—the margin of error and the confidence level. The **margin of error** indicates the range of values that can result when you use a sample to estimate the population. For example, a 5% margin of error means that if 50% of the sample adopted a recommendation, you can be fairly certain that 47.5% to 52.5% of the whole population adopted it. If results show that 95% of the sample improved in parenting skills, you can feel comfortable saying that 93.5% to 97.5% of *all* participants improved in those skills.

The risk of being wrong within the margin of error is known as the **confidence level.** That is, a 95% confidence level (C = .95) means that there is a 5% chance that the results will not fall within the specified margin of error. Using the previous example, there is a 5% chance that this interpretation is not correct—that 93.5% to 97.5% of all participants did not improve their skills. We can say that we are "95% sure" that our conclusions accurately reflect the total population.

Most studies in Extension allow for a 5–10% margin of error with a confidence level of 95%. Appendix 2 gives estimates of suggested sample sizes at different margins of error. Extension practitioners recommend that when greater accuracy is needed, all individuals be included when the population is 100 or less. In this manner, the problem in sampling small populations is overcome: Because small populations exhibit greater variability, a greater proportion of the population must be sampled.

Anticipating nonresponse

Generally, a certain number of people will not respond for one reason or another. The sample size needs to be large enough to compensate for these nonresponses. A way to do this is to estimate the rate of nonresponse (30% nonresponse may be realistic) and increase the sample size accordingly. For example, if the total group of Extension participants is 150 and we set a margin of error at 10%, we need a sample size of 61 participants (Appendix 2). If we think that 70% of the individuals will respond, we actually need to begin with a sample of 87 (61/.70) to ensure that we end up with 61 people.

Sampling for other purposes: nonprobability samples

In some instances, probability sampling may be impossible, unnecessary or even undesirable. You may be limited to only those participants or those counties that agree to be included. Or you may want more in-depth information regarding a particular program, participants or delivery method. In **nonprobability sampling**, there is no expectation that each unit has an equal chance of being included in the sample. Since the sample does not intend to represent the population, findings should not be generalized to the whole.

Strategies

Quota sample

A quota sample divides the population being studied into subgroups such as male and female, younger and older. You estimate the proportion of people in each group based on what you know about the population in general. For example, if you know that 20% of the population is headed by a single-parent family, then search for respondents until 20% of your sample is single parents. This is not the same as the stratified random sample because not every single parent is identified and has an equal opportunity of being included.

Purposeful sampling

Patton (1990) describes a number of sampling strategies that serve purposes other than representativeness or randomness (see table 1). Basic to all these is the importance of selecting "information-rich" cases from which you can learn much about issues that are important to the study. Focus on the specific rather than the general; for example, if an evaluation's purpose is to increase a program's effectiveness in reaching low-income families, you may learn more by conducting an in-depth query of the few poor families in the program than by gathering standardized information from a random sample of all participants.

Sample size

When your purpose is to examine selected cases in greater depth, the previous recommendations for sample size do not apply.

Often, in purposeful sampling the sample size is very small—possibly even just one case study (n=1). For example, you may wish to conduct a single case study of a low-income participant and the difference a program made in his or her life. Or, your purpose may require an in-depth analysis of successful community collaborations, highlighting the factors affecting success. Another option is to ask knowledgeable people to identify such collaborations and then select a certain number to include in your sample. If the purpose of the evaluation is to document diversity or variation, a larger number of cases may be necessary to capture the variety. What is the recommended sample size? There are no rules. It depends upon what you want to know, what will be useful, what will be credible, and what can be accomplished within the time and resources you have available.

When sampling for the purpose of generalizability, the sample size is set in advance. With purposeful sampling, on the other hand, the sample size may change as the study progresses. For instance, based on early investigations, you may identify other "information-rich" cases to include. Or, you may terminate data collection when no new information is forthcoming from the new sampled units.

More than likely, the factors determining sample size will be time and resources. Select the individual cases carefully to choose those that are most likely to provide the information you are seeking. Finally, explain and justify your sampling procedures and decisions so that information users and decisionmakers understand your logic. As a professional, you are obliged to describe the strengths and weaknesses of the sampling procedures that are relevant for understanding the findings. Be careful not to generalize but to focus on the intention and strengths of a purposeful sample.

Table 1. Purposeful sampling strategies

Extreme or deviant case sampling	Learning from highly unusual cases, such as outstanding successes/notable failures; top of the class/dropouts; crises.
Intensity sampling	Information-rich cases that manifest the phenomenon intensely, but not extremely, such as good students/poor students; above average/below average.
Maximum variation sampling	Purposefully pick cases to illustrate a wide range of variation on dimensions of interest. Identifies important common patterns that cut across variations.
Homogeneous sampling	Pick cases that are alike. Focuses, reduces variation, simplifies analysis, faclitates group interviewing.
Typical case sampling	Illustrates or highlights what is typical, normal, average.
Stratified purposeful sampling	Illustrates characteristics of particular subgroups of interest; facilitates comparisons.
Critical case sampling	Permits *logical* generalization and maximum application of information to other cases because if it's true of this one case, it's likely to be true of all other cases.
Snowball or chain sampling	Relies on people identifying other people or cases to investigate. As they identify new names, the snowball gets bigger. A few key names may be mentioned repeatedly indicating their special importance. Particularly useful when there is no population list or you want to draw a sample based on recommendations.
Criterion sampling	Picking all cases that meet some criterion, such as all children abused in a treatment facility.
Theory-based or operational construct sampling	Finding manifestations of a theoretical construct of interest so as to elaborate and examine the construct.
Confirming and disconfirming cases	Elaborating and deepening initial analysis, seeking exceptions, testing variation.
Opportunistic sampling	Following new leads during fieldwork, taking advantage of the unexpected, requires flexibility.
Random purposeful sampling (small sample size)	Uses random selection to select limited number of cases from a larger purposeful sample; adds credibility (not for generalizations or representativeness).
Politically important or sensitive case sampling	Attracts attention to the study (or avoids attracting undesired attention by eliminating politically sensitive cases from the sample).
Convenience sampling	Takes individuals who are available or cases as they occur. Saves time, money and effort. Poorest rationale; lowest credibility.
Combination or mixed purposeful sampling	Combination of above approaches, meets multiple interests and needs.

Patton, 1990: 182–183 (Adapted with permission by author. Reprinted with permission of Sage Publications.)

Summary

Sampling is the selection of a smaller number of units from among a larger group or population. To generalize from the sample to the population, probability or random sampling is needed to ensure that the sample is representative. If generalizations are not desired or necessary, nonprobability sampling is appropriate and can yield very useful information if cases are selected thoughtfully.

Many people are concerned about how many cases to include in their samples. There is no simple answer. Determining the sample size depends upon:

1. The purpose of the study
2. Expectations of information users, and
3. The resources available.

References

Abbey-Livingston, Diane and Davis S. Abbey. 1982. *Enjoying Research? A 'How-to' Manual on Needs Assessment*. Ontario: Ministry of Tourism and Recreation.

Cochran, W.G. 1963. *Sampling Techniques*, 2nd Ed. N.Y.: John Wiley and Sons, Inc.

Fink, Arlene. 1995. *How to Sample in Surveys*. Thousand Oaks, Calif: Sage Publications.

Ilvento, Tom W., Paul D. Warner, Richard C. Maurer. 1986. *Sampling Issues for Evaluations in the Cooperative Extension Service*. Kentucky Cooperative Extension Service, Lexington, KY.

Isaac, Stephen and William B. Michael. 1981. *Handbook in Research and Evaluation*, 2nd Ed. San Diego, EdITS Publishers.

Kerlinger, Fred N. 1973. *Foundations of Behavioral Research*, 2nd Ed. N.Y.: Holt, Rinehart, and Winston, Inc.

Patton, Michael Q. 1990. *Qualitative Evaluation and Research Methods*. 2nd ed. Newbury Park, Calif: Sage Publications.

Rohs, F. Richard. 1985. *Sampling Techniques for Extension Program Planners and Evaluators.* The Cooperative Extension Service, University of Georgia, Athens, Georgia.

Sabrosky, Laurel K. 1959. "Sampling" in *Evaluation in Extension,* pg. 37-45. ES-USDA, Washington, D.C.

Smith, M.F. 1983. *Sampling Considerations in Evaluating Cooperative Extension Programs*. Florida Cooperative Extension Service, Gainesville, Florida.

Appendix 1. Table of random numbers

59391	58030	52098	82718	87024	82848	04190	96574	90464	29065
99567	76364	77204	04615	27062	96621	43918	01896	83991	51141
10363	97518	51400	25670	98342	61891	27101	37855	06235	33316
86859	19558	64432	16706	99612	59798	32803	67708	15297	28612
11258	24591	36863	55368	31721	94335	34936	02566	80972	08188
95068	88628	35911	14530	33020	80428	39936	31855	34334	64865
54463	47237	73800	91017	36239	71824	83671	39892	60518	37092
16874	62677	57412	13215	31389	62233	80827	73917	82802	84420
92494	63157	76593	91316	03505	72389	96363	52887	01087	66091
15669	56689	35682	40844	53256	81872	35213	09840	34471	74441
99116	75486	84989	23476	52967	67104	39495	39100	17217	74073
15696	10703	65178	90637	63110	17622	53988	71087	84148	11670
97720	15369	51269	69620	03388	13699	33423	67453	43269	56720
11666	13841	71681	98000	35979	39719	81899	07449	47985	46967
71628	73130	78783	75691	41632	09847	61547	18707	85489	69944
40501	51089	99943	91843	41995	88931	73631	69361	05375	15417
22518	55576	98215	82068	10798	86211	36584	67466	69373	40054
75112	30485	62173	02132	14878	92879	22281	16783	86352	00077
80327	02671	98191	84342	90813	49268	95441	15496	20168	09271
60251	45548	02146	05597	48228	81366	34598	72856	66762	17002
57430	82270	10421	00540	43648	75888	66049	21511	47676	33444
73528	39559	34434	88596	54086	71693	43132	14414	79949	85193
25991	65959	70769	64721	86413	33475	42740	06175	82758	66248
78388	16638	09134	59980	63806	48472	39318	35434	24057	74739
12477	09965	96657	57994	59439	76330	24596	77515	09577	91871
83266	32883	42451	15579	38155	29793	40914	65990	16255	17777
76970	80876	10237	39515	79152	74798	39357	09054	73579	92359
37074	65198	44785	68624	98336	84481	97610	78735	46703	98265
83712	06514	30101	78295	54656	85417	43189	60048	72781	72606
20287	56862	69727	94443	64936	08366	27227	05158	50326	59566
74261	32592	86538	27041	65172	85532	07571	80609	39285	65340
64081	49863	08478	96001	18888	14810	70545	89755	59064	07210
05617	75818	47750	67814	29575	10526	66192	44464	27058	40467
26793	74951	95466	74307	13330	42664	85515	20632	05497	33625
65988	72850	48737	54719	52056	01596	03845	35067	03134	70322
27366	42271	44300	73399	21105	03280	73457	43093	05192	48657
56760	10909	98147	34736	33863	95256	12731	66598	50771	83665
72880	43338	93643	58904	59543	23943	11231	83268	65938	81581
77888	38100	03062	58103	47961	83841	25878	23746	55903	44115
28440	07819	21580	51459	47971	29882	13990	29226	23608	15873
63525	94441	77033	12147	51054	49955	58312	76923	96071	05813
47606	93410	16359	89033	89696	47231	64498	31776	05383	39902
52269	45030	96279	14709	52372	87832	02735	50803	72744	88208
16738	60159	07425	62369	07515	82721	37875	71153	21315	00132
59348	11695	45751	15865	74739	05572	32688	20271	65128	14551
12900	71775	29845	60774	94924	21810	38636	33717	67598	82521
75086	23537	49939	33595	13484	97588	28617	17979	70749	35234
99495	51434	29181	09993	38190	42553	68922	52125	91077	40197
26075	31671	45386	36583	93459	48599	52022	41330	60651	91321
13636	93596	23377	51133	95126	61496	42474	45141	46660	42338

Appendix 2. Recommended sample sizes for two different precision levels

Population Size	Sample size +5%	10%	Population size	Sample size 5%	10%
10	10		275	163	74
15	14		300	172	76
20	19		325	180	77
25	24		350	187	78
30	28		375	194	80
35	32		400	201	81
40	36		425	207	82
45	40		450	212	82
50	44		475	218	83
55	48		500	222	83
60	52		1000	286	91
65	56		2000	333	95
70	59		3000	353	97
75	63		4000	364	98
80	66		5000	370	98
85	70		6000	375	98
90	73		7000	378	99
95	76		8000	381	99
100	81	51	9000	383	99
125	96	56	10,000	385	99
150	110	61	15,000	390	99
175	122	64	20,000	392	100
200	134	67	25,000	394	100
225	144	70	50,000	397	100
250	154	72	100,000	398	100

Source: Isaac and Michael, 1981; Smith, M. F., 1983

Author: Ellen Taylor-Powell is a program development and evaluation specialist for Cooperative Extension, University of Wisconsin–Extension.

You can obtain copies of this publication from Cooperative Extension Publications, Room 170, 630 W. Mifflin Street, Madison, WI 53703, or by phoning (608)262-3346.