



Sampling A Diverse Audience

What steps do I take to draw a *stratified* sample of a target audience in my county to do a needs assessment or gather baseline data? Also, why is a *stratified* sample so important?

There are two steps in the sampling process. The first is to compile a comprehensive and accurate list of the potential target audience: producers in the swine industry for a potential bio security program or coaches in youth sports for a potential program on teen diet and health for example. This list is called the *sampling frame* or *universe*. The second step is to draw a *stratified random sample* from that list. By far, the more challenging and often neglected task is the first.

Step One: A Sampling Frame: Your goal in constructing a county list for a needs assessment is to make sure the list is comprehensive. For example, for a potential bio security program it should include every swine producer with 10 hogs or more, and for a potential teen nutrition program it should include all coaches employed as volunteers in public, private and club sports who work with teens 13-18 years of age. The list should include all producers or coaches whether they are currently cooperators or not. (If the list has other swine or sports-related people, you must be able to extract them from the names of the swine producers and coaches).

Why does the list need to be comprehensive?

- Although the impact of problems like poor bio-security or teen dieting have been established through research, the extent to which behavior or lack of knowledge on the part of the target audience contributes to that problem is not known, and, what you want to find out. You want this information so you know which behaviors or knowledge to emphasize in videos, slides, printed materials and educational meetings. Thus, it is important to include a complete list of the target audience (producers or coaches) in the sampling frame in order to find out *the widest distribution of management practices and knowledge that relate to the problem*.
- Conducting a needs assessment or collecting baseline data to find the distribution of the practices and knowledge will have an immediate and direct bearing on the content of the program.
- There are also long-term benefits of collecting such data. Having an up-to-date list of swine producers and coaches in a county means that a county agent will have a complete list of the principal target audience for marketing and evaluating the educational program.
- Sometimes a seminal list exists at the state level but you might have to add to it. For swine producers: first, you can obtain one for your county by calling the Bureau of Animal Industries at 717/783-5301. Next, compare this list with those of co-ops in the area. Check with farm realtors for farmers who have recently sold their business and for the names of the new owners. Check with other appropriate sources.
- At other times, such as compiling a list of coaches, you may have to start at the county level. Begin with local schools or chapters of professional organizations.
- The list *does not* have to be in any particular *order*.

Continue down for additional information.



Sampling A Diverse Audience

Step Two: Drawing a Sample: Since it is impossible to interview every member of the target audience in the county, methods have been established to select a fraction to accurately represent the whole. That fraction is known as the “sample.”

What is important from a scientific point of view is to make certain everyone in the *sampling frame* gets a known chance of being selected. That means two things: no conscious selection of any particular member of the target audience should be made, but on the other hand, the selection of members of the target audience should not be “hit or miss” either.

Sometimes you need to think about sampling from *several* subgroups in a target audience because different subgroups of the large audience represent different management practices of knowledge or attitudes. Sampling from subgroups is called Stratified Sampling. Before you can draw a sample, you need to know the proportion of each sub-group to the total number in the target audience in your county.

For example in the swine program, if you have two groups, 50 independent swine producers and 25 contract producers, then find the fractional proportion for each subgroup:

$$\text{Independents} = \frac{50}{75}$$

$$\text{Contract} = \frac{25}{75}$$

(It may be easier to keep track of these fractions if you don't reduce them.)

These fractional proportions will be used below. If you have three groups, apply the same principle to obtain three fractions.

There are four easy steps in drawing a sample:

1. You want to decide the size of the overall sample first. Start with the number of interviews you plan to do in your county (for example, 15) and as a rule of thumb, add 5-7 people for outright rejection and another 5-8 people to cover a variety of other reasons such as producers who withdraw at some point in the process. Thus, the sample here stands at 30. It is better to have more names in your sample than too few because you don't want to have to draw another sample.
2. Once you know the overall sample size you want, the next step is to decide what proportion of that number should be in each subsample of independents and contracts. To find out how many producers to have in each subsample, take the overall sample (30) and multiply it by the fractional part for each group (see above).

$$\frac{50}{75} \times 30 = 20 \text{ in subsample of Independents}$$

$$\frac{25}{75} \times 30 = 10 \text{ in subsample of Contract Producers}$$

Continue down for additional information.



Sampling A Diverse Audience

3. To select the names from each list, do two things for each group: first,

For Independents: Divide the number of independents in the subsample into the total number of independents in the sampling frame.

$$\frac{20 \text{ subsample independents}}{50 \text{ sampling frame independents}} = 2.5$$

**Round down to 2.
Be conservative.**

The answer tells you to select every 2nd name in your list of independent producers.

For Contracts: Divide the number of contract producers in the subsample into the total number of contracts in the sampling frame.

$$\frac{10 \text{ subsample contracts}}{50 \text{ sampling frame contracts}} = 5.0$$

The answer tells you to select every 5th name in your list of contract producers.

4. ***Don't begin to select names from the top of the list. You want to start in a random place.***

To find out where on each list to start, put numbers in a container from 1 to the answer in Step 3 above. For the first group the numbers would be #1 and #2. For the second group, the numbers would be #1, #2, #3, #4, and #5. Pull out a number. This is the place on the list you *start* to select the names in your subsample. The next names on the list will be every nth (Step 3) name after your first selection (2 in the first group and 5 in the second). Repeat process for each subgroup.

When you finish selecting names, you should have two subsamples. The number in each subsample will be somewhat larger than the number of interviews you need to complete in each subsample. To set up dates for the interviews, you can begin with any name in the sample since they were all randomly chosen.

A stratified random sample is designed to investigate the widest distribution of certain behavior, attitudes or knowledge. This type of sample is valuable to program planers when there is reason to believe that the behavior, attitudes or knowledge in subgroups of the target audience are different.

Nancy Ellen Kiernan, Ph.D., Program Evaluator, nekiernan@psu.edu

The reference citation for this Tipsheet is: Kiernan, Nancy Ellen (2001)
Sampling A Diverse Audience: Tipsheet #58, University Park, PA: Penn State Cooperative Extension.
Available at: <http://www.extension.psu.edu/evaluation/pdf/TS58.pdf>

This Web site is copyrighted by The Pennsylvania State University. The information may be used for educational purposes but not sold for profit.