

# TRY THIS SAMPLE ON FOR SIZE

by Dale Atkinson

Have you ever dreamed of working in the Sample Design Section? Probably not! However, you may have wondered how sample sizes are determined. Well, you're in luck, because the topic on our talk show today is one even Oprah Winfrey hasn't thought of – sample size estimation. Our special guest is that famous sampler, Mr. Sam Puller. Let's get right to the phones and hear from our listeners.

*Caller 1:* The sample size for one of our State surveys has just been increased from 250 to 1,000. Why was such a large increase necessary?

*Mr. P:* The biggest factor in determining sample size is the level of precision desired in the survey indications. Precision is measured by something called the coefficient of variation - CV, for short. The smaller the CV, the better the precision. In your case, it was determined that a lower CV is needed. Quadrupling the sample size will reduce the CV by half. If the CV had been 20 percent with a sample size of 250, it will probably be about 10 percent with a sample size of 1,000. If your sample size were increased even more - for example, quadrupled to 4,000 - then the CV would be reduced to 5 percent. Sample size is determined by the required precision of the survey indications.

*Caller 2:* The sample size for our labor survey is 400, but the sample size for our December Crops/Stocks Survey is 2,000. Why the big difference?

*Mr. P:* Estimation objectives differ from survey to survey. For the labor survey, sample sizes are determined with the objective of producing regional estimates, not State estimates. Thus, when the sample size for your State is lumped in with the other States in the region, fairly precise regional estimates can be obtained. The December Crops/Stocks Survey sample sizes are designed to provide precise State estimates.

*Caller 3:* Well, our labor survey has a sample size of 500, and our Agricultural Resource Management Survey (ARMS) has a sample size of 1,000. We do not do State estimates for either survey in our State, only regional estimates. So why is there a difference in these survey sample sizes?

*Mr. P:* Different types of data demonstrate differing amounts of variability. The labor sample sizes are designed to provide precise estimates of number of workers and wage rates. These data tend to be less variable than total farm expenditures - the primary item of interest in the ARMS. The more variability in an item, the larger the sample size needed to estimate it well.

*Caller 4:* Our State has a sample size of 1,500 for the December Crops/Stocks Survey. Our neighbor State to the north also has a sample size of 1,500 although they have twice as many farms. Why is that?

*Mr. P:* Sample size is usually not affected by population size (i.e., number of farms). Thus, if a sample size of 1,500 is needed to obtain a certain CV, it doesn't usually matter if there are 50,000 or 100,000 farms in the State - the sample size still needs to be 1,500.

It is a common misconception to think that if you have a smaller population size, your sample size will be less. NASS has larger sample sizes in some States because they produce a larger share of the U.S. production, and NASS wants a lower CV in those States. Of course, States which are more agricultural tend to have more farms, but it is the lower CV that is usually affecting the sample size - not the number of farms. Sampling experts take population sizes into account only when more than 10 percent of the population will be in the sample, or the State sample size must be allocated to a set of strata.

*Caller 5:* The list sample size for our December Crops/Stocks Survey was increased, and although the crop acreage CV's decreased somewhat, we did not see any change in our grain stocks CV's. Why was this?

*Mr. P:* For multiple frame surveys, the area frame component, or nonoverlap (NOL) domain, must be taken into account – both the CV for NOL and the percentage of the indication that comes from the NOL. The stocks estimates in your State probably have a large NOL component, so that simply increasing the list sample size did not result in a noticeable improvement in the multiple frame CV. Rather than increasing the list sample size, your State probably needs to improve the list coverage for stocks (capacity) so that less of the total estimate comes from the NOL domain.

*Caller 6:* Does stratification have any effect on sample sizes?

*Mr. P:* Stratification is a tool that gives us some control over data variability. Stratified samples usually produce estimates with much better precision (in other words, lower CV's) than would be realized if the sample size were used in an unstratified design. The quality of the control data used to construct the strata is an important point in effective use of stratification. High quality, current control data on the sampling frame will result in better stratification and lower CV's.

*Caller 7:* We need to obtain a 5 percent CV on our hog indication, but we cannot afford the sample size which is required to get a 5 percent CV.

*Mr. P:* Resources do have some effect on sample size. For your situation, you must decide if the sample size that you can afford will result in a level of precision that will make the survey worth doing. Survey organizations often find themselves using the sample sizes they can afford rather than the sizes they desire. However, a survey organization would be better off not doing a survey if the affordable sample size will produce such large CV's that the survey indications are not usable.

Well, this has been very interesting. I hate to end the discussion, but I have to run. So, until next time, happy sampling!