Research Methodology

July 2009

A GUIDE TO SAMPLING & STATISTICAL RELIABILITY

The word “survey” is used most often to describe a method of gathering information from a sample of individuals. This “sample” is usually just a fraction of the population being studied. For example, a sample of voters is questioned in advance of an election to determine how the public perceives the candidates and the issues..…a manufacturer does a survey of the potential market before introducing a new product…..local government will commission a survey to gather information on residents’ perceptions of a service. But how can we be sure that the sample we choose represents the views of the population we’re considering?

Sampling Methodology

A variety of sampling methods are available, and it is important to choose the most appropriate method for the type of survey proposed (and equally the most appropriate survey method for the population to be studied), to help reduce possible survey error. The most commonly used sampling methods used are as follows:

> *Random sampling* is a sampling technique where you select a group of subjects (a sample) for study from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has a known, but possibly non-equal, chance of being included in the sample. By using random sampling, the likelihood of bias is reduced.
>
> *Simple random sampling* is the basic sampling technique where each individual is chosen entirely by chance and each member of the population has an equal chance of being included in the sample. Every possible sample of a given size has the same chance of selection; that is, each member of the population is equally likely to be chosen at any stage in the sampling process.

There may often be factors which divide up the population into sub-populations (groups / strata) and we may expect the measurement of interest to vary among the different sub-populations. This has to be accounted for when we select a sample from the population in order that we obtain a sample that is representative of the population. This is achieved by *stratified sampling*. A stratified sample is obtained by taking samples from each stratum or sub-group of a population.

Cluster sampling is a sampling technique where the entire population is divided into groups, or clusters, and a random sample of these clusters are selected. All observations in the selected clusters are included in the sample.

Quota sampling is a method of sampling widely used in opinion polling and market research. Interviewers are each given a quota of subjects of specified type to attempt to recruit for example, an interviewer might be told to go out and select 20 adult men and 20 adult women, 10 teenage girls and 10 teenage boys so that they could interview them about their television viewing.

Survey Error

When designing a survey, there are four basic types of survey error that need to be taken into consideration. The overall aim of reducing these errors will help to determine the size of the sample to be used, the type of survey to be carried out (e.g. telephone, postal, face-to-face), and the design of the questions to be asked.

1. Random sampling allows characteristics in a population to be estimated with precision, with larger sample sizes achieving larger degrees of accuracy. The extent to which the precision of sample survey estimates is limited by the number of people surveyed is described by the term *sampling error*.
2. *Coverage error* occurs when the list from which a sample is drawn does not include all elements of the population, making it impossible to give all elements of the population an equal or known chance of being included in the sample survey – for example the omission of people without telephones from a telephone survey.
3. *Measurement error* occurs when a respondent's answer to a survey question is inaccurate, imprecise, or cannot be compared in any useful way to other respondents' answers. This type of error results from poor question wording and questionnaire construction.
4. *Non-response error* occurs when a significant number of people in the survey sample do not respond to the questionnaire and have different characteristics from those who do respond.

Efforts must be made to reduce all four sources of survey error to acceptable levels. *Sampling* and *coverage* errors can be greatly reduced by carefully choosing an appropriate methodology for the population to be questioned (in terms of both sampling technique, and survey size), whilst *measurement* and *non-response* issues can be addressed through careful design of questions, questionnaires and implementation method.

Sample Size
In general, the larger the sample size, the more closely your sample data will match that from the overall population. However in practice, you need to work out how many responses will give you sufficient precision at an affordable cost.

There are two key elements to sample size:

1. Sample sizes are not dependent upon the size of the population. If for example, you want to estimate the percentage of people in a small town of 25,000 adults with qualifications within 3 percentage points of the actual percentage, a completed sample of abut 1,024 respondents is needed. If you want to estimate this characteristic for a city of 2 million adults with the same confidence, then around 1,067 respondents are needed. If you want to estimate this for the entire country, then again 1,067 are needed.

2. A relatively few completed questionnaires can provide surprising precision at a high level of confidence. If you can be satisfied with knowing whether an estimate from a sample survey is plus or minus (+ / -) 10% of the true population value, that could be accomplished by receiving completed questionnaires from a random sample of 100 individuals. What this means is that in a properly conducted national survey with a completed simple random sample of 100 households, in which 60% of the respondents say they own their own home, we could say with 95% confidence that between 50% and 70% of the population own their own home.

It is only when population size decreases to a few thousand or less that the number of completed questionnaires needed for a given level of precision declines significantly, as shown in Table 1. This table illustrates the four factors that must be taken into consideration in determining how large a sample size is needed in order to make population estimates:

- How much sampling error can be tolerated (i.e. how precise the estimates need to be).
- The population size from which the sample is to be drawn.
- How varied the population is with respect to the characteristic. This table assumes a yes/no question with maximum variation – i.e. a 50/50 split in responses.
- The amount of confidence you want to have in the estimates made (generally 95% confidence is used).

Table 1. Completed sample sizes needed for various population sizes at three levels of precision

Population Size	Sample size for the 95% confidence level		
	+/- 10% sampling error	+/- 5% sampling error	+/- 3% sampling error
100	49	80	92
200	65	132	169
400	78	196	291
600	83	234	384
800	86	260	458
1,000	88	278	517
2,000	92	322	696
4,000	94	351	843
6,000	95	361	906
8,000	95	367	942
10,000	95	370	965
20,000	96	377	1,013
40,000	96	381	1,040
100,000	96	383	1,056
1,000,000	96	384	1,066
1,000,000,000	96	384	1,067

If you want to analyse survey data for sub-groups (such as by age, ethnicity or gender), you also need to take this into account before starting a survey – as you need to consider the reliability of the data for each individual group. If you're happy to work to +/- 10%, you will need at least 100 responses from each group to ensure a degree of accuracy – so you could begin by deciding how you want to analyse the results and build up a sample size from there. For example if you wanted to look at the differences between men and women – you will need at least 100 men and 100 women in your sample. If you have a range of different age groups or ethnic backgrounds, then again you will need to ensure you have at least 100 people in each category.

Statistical Reliability

Having looked at how big a sample size should ideally be, we can now look at the issue of reliability – that is, having carried out a survey and received a certain number of responses, how confident can we be in the results.

As we know, the residents who take part in a survey are only a sample of the total "population" of residents in an area, so we cannot be certain that the figures obtained are exactly those that would have been reached were everyone had responded (the "true" values). We can, however, predict the variation between the sample results and the "true" values from knowledge of the size of the samples on which the results to each question is based, and the number of times a particular answer is given. As with sample size, the confidence with which we can make this prediction is usually chosen to be 95% - that is, the chances are 95 in 100 that the "true" value will fall within a specified range. The following illustrates the predicted ranges for different sample sizes and percentage results at the "95% confidence interval":

Table 2. Confidence intervals for various sample sizes

Size of sample on which survey result is based	Approximate sampling tolerances applicable to percentages at or near these levels		
	10% or 90% + / -	30% or 70% + / -	50% + / -
100 responses	6	9	10
200 responses	4	6	7
500 responses	3	4	4
1,000 responses	2	3	3
1,500 responses	1	2	3

For example, with a sample size of 1,000 where 30% give a particular answer, the chances are, 19 in 20 that the "true" value (i.e. the one which would have been obtained if the whole population had been interviewed) will fall within the range of +/- 3 percentage points from the survey result (i.e. between 27% and 33%).

When results are compared between separate groups within a sample (e.g. males versus females), different results may be obtained. The difference may be "real," or it may occur by chance (because not everyone in the population has been interviewed).

To test if the difference is a real one - i.e. if it is "statistically significant" - we again have to know the size of the samples, the percentage giving a certain answer and the degree of confidence chosen. If we once again assume a "95% confidence interval", the differences between the results of two separate groups must be greater than the values given in the following table:

Table 3. Confidence intervals when comparing results for sample groups

Size of sample on which survey result is based	Differences required for significance at or near these percentage levels		
	10% or 90% + / -	30% or 70% + / -	50% + / -
100 vs. 100	8	13	14
200 vs. 200	6	9	10
500 vs. 500	4	6	6
500 vs. 1,000	3	5	5

It is important to note that, strictly speaking, the above confidence interval calculations relate only to samples that have been selected using strict probability sampling methods. However, in practice it is reasonable to assume that these calculations provide a good indication of the confidence intervals relating to most random sampling approaches used.

Christine Wright, Principal Research Officer
Research Sandwell
July 2009