Statistical Confidence in the GA Survey

Everyone that depends on the results of a survey is concerned about response rates and the level of confidence they can place in the survey results. The GA Survey is designed to maximize our statistical confidence. However, this is one part of a survey project that does require fundamental understanding of some statistics. For laypersons it is easiest, and most intuitive, to consider the attention given to the four factors that determine our statistical confidence in the GA Survey:

Size of the population. The population of interest is all general aviation aircraft registered with the FAA that are based in the United States and potentially operating during a calendar year. The 2012 survey population included 281,804 aircraft. Each year, a sample is drawn from the population and the survey is administered to the sample. The 2012 survey sampled 30.3% of the general aviation aircraft (85,403 aircraft). By any best practice research measure, either absolute number or as a percent of the population, the number of aircraft invited into the survey is large. The percentage of the sample that responded to the survey invitation was 43.8%. That percentage is the response rate. Increasing response rates improve the data quality and the accuracy of the survey.

Segmentation analysis desired. We typically analyze the data along some demographic segmentation, for example, aircraft type, region, primary and actual use, or age of aircraft. Each segment in essence is another population. Because critical policy decisions will be focused on the analysis of a segment, then statistical confidence must be focused on the segment, not the population. For this reason we sample 100% of all aircraft authorized to operate on-demand FAR Part 135, turbine-powered aircraft, rotorcraft, aircraft in Alaska, special light-sport aircraft, and any aircraft manufactured in the last five years.

Degree of variance in responses from the population. This factor can be the most difficult to understand. If the respondents' responses tend to be tightly clustered, then we don't need to sample as many aircraft in a segment to get the same confidence as we would if the responses range widely. Imagine you polled your office colleagues, and the first five people gave the same answer. Would you continue polling? Probably not. What if you got five different responses? You'd probably keep polling. Why? More variability requires larger samples. Because we conduct the survey and analyze the data each year, we know a lot about the variance we expect to see in the data and this enables us to select sample sizes that optimize the statistical confidence in our results for each segment of interest.

Tolerance for error. How accurate do you need the results to be? If you're going to impose a new regulation or make multi-million dollar business decisions, then you probably have less tolerance for error. The standard error (SE) is a measure of the variation among the estimates derived from all possible samples. For convenience, the standard error is often expressed as a percent of the estimate and is called a relative standard error (RSE). A sample estimate and an estimate of its SE or RSE can be used to construct interval estimates that have a prescribed confidence that the interval includes the average of the estimates derived from all possible samples with a known probability (see Appendix A of the published report for a detailed example of how to do this.) As a rule of thumb, RSE's below 20 are considered reasonable and by this rule the GA Survey has very low sampling error—RSE's range from 0.5 percent to 5 percent for estimates by aircraft type. Sampling error gives us some idea of the precision of our statistical estimate. A low sampling error means that we had relatively less variability or range in the sampling distribution. The standard error is also related to the sample size. The greater the sample size, the smaller the standard error. Why? Because the greater the sample size, the closer the sample is to the actual population itself. If you take a sample that consists of the entire population you actually have no sampling error because you don't have a sample, you have the entire population. In that case, the mean you calculate is the parameter.

Optimizing Data Quality with Best Practice Research Methods

Thus far, we have only been discussing only sampling error in relation to the statistical confidence we have in the GA Survey results. But all surveys and censuses are also subject to nonsampling errors. Nonsampling errors are of two kinds – *random and non-random*. **Random nonsampling errors** arise because of the varying interpretation of questions by respondents and varying actions of coders and other survey processers. Some randomness is also introduced when respondents estimate their answers (e.g., proportion of hours flown in different flight conditions, average fuel burn rate). Random nonresponse errors usually, but not always, result in an understatement of sampling errors and thus an overstatement of the precision of the survey estimates.

Nonrandom sampling errors result from total nonresponse (no usuable data are obtained for a sampled aircraft), partial or item nonresponse (only a portion of a response may be usable), inability or unwillingness on the part of respondents to provide correct information, difficulty interpreting questions, and coverage problems. To compensate for suspected nonrandom errors, adjustment of the sample estimates are often made. For example, adjustments are frequently made for nonresponse by weighting survey responses. (See Appendix A of published report for more details on weighting of the GA Survey.) Adjustments made for item nonresponse are referred to as imputations. Imputation for item non-response is made by substituting for a missing item the response to that item of a respondent having demographic or ancillary characteristics that are "similar" to those of the nonrespondent.

Nonsampling errors are minimized through strict quality controls on the collection process and careful review of all reported data. In addition, the GA Survey incorporates the following best practice survey research methods as defined by the American Association of Public Opinion Research (AAPOR), the preeminent association of survey research firms and professionals:

Have specific goals for the survey.

Consider alternative methods of data collection.

Select samples that well represent the population to be studied.

Use designs that balance costs with errors.

Take great care in matching question wording to the concepts being measured and the population studied.

Pretest questionnaires and procedures.

Check quality at each stage of data collection and analysis.

Maximize cooperation or response rates.

Use appropriate statistical analytic and reporting techniques.

Develop and fulfill pledges of confidentiality given to respondents.

Disclose all methods of the survey to allow for evaluation and replication

We strongly encourage users of the GA Survey to read Appendix A of the published reports for a more comprehensive discussion of the methodology of the GA Survey.