Summer 2007

A : I/R
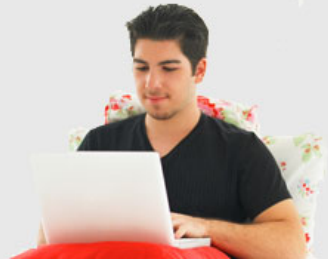
ACCURACY'S IMPACT ON RESEARCH
A Knowledge Networks Newsletter

Weight, Weight, Don't Tell Me!

By Charles DiSogra

Data junkie, left-brain brat, data-driven druid, this is just the short list of affectionate labels people around the office place on those whose decision-making is based on numbers from survey data. But let's be honest. Who would make an uninformed decision, a decision not involving some kind of data to support it? However, numbers, in and of themselves, can be a two-edged sword. When they are accurate, they will cut through the darkness and pierce the shroud of doubt. When they are inaccurate, they will obfuscate truth and be the glorious instrument of destruction.

$$s^2 = \frac{\Sigma \, (x_i - \bar{x})^2}{n - 1}$$

Good Data Gone Bad

Careful data collection alone — which some may conclude is enough to produce good data — does not guarantee that an accurate measurement will result. Good data gone bad is more often the progeny of design error and misunderstanding of the technical nuances when working with survey data. You might be tempted to attribute such tragedy to simple ignorance, but that is much too strong a word, especially since very smart people can be grossly misinformed and consequently make the wrong conclusion. History is replete with such examples. In the survey world, the 1948 presidential election newspaper headline, "Dewey Beats Truman" is the classic story of a bad conclusion from what was thought to be good survey data and the hubris that accompanies having numbers in hand. Truman won and the data junkies were wounded.

Weighting for Good Data

So how can you be certain that good survey data — the product of good instruments and good design — will deliver accurate results? Weighting the data is the answer. We use data weights to adjust each respondent's contribution to the overall results because not all survey respondents are equal. OK, time for a definition: A data weight is simply a multiplier that makes a given respondent's

contribution larger or smaller to compensate for a variety of both planned and unexpected disproportionate effects.

Planned effects are generally the result of a sample design or strategy. For example, let's say you intentionally stratify the population into two groups: pet owners and non-pet owners. And, you have a sample design that calls for equal numbers in each group. If you know that their natural proportion in the population is not really 50-50, a data weight can be applied to put them back in their correct proportion when you want to draw conclusions about the whole population using your total sample. What if among the pet owners, those who have cats are mysteriously not responding to your survey the way dog owners are? You now have a problem of non-response. This, too, can be adjusted with data weights if we know from some external source the relative proportion of cat owners among all pet owners in the population. Because all this is done after the survey is over, it is referred to as post-stratification weighting.

Knowledge Networks Gets It Right

Now you have the general two-dimensional idea. Operationally, it is much more complicated. Usually many dimensions have to be taken into account in constructing weights. When the statistics department at Knowledge Networks (KN) prepares data weights for a client's data file, there is an enormous responsibility to get it right. And we do get it right. Complex iterative computer programs are used to balance all the dimensions simultaneously, such as males and females, by multiple age groupings, race/ethnic categories, income categories, education attainment, geographic location, whether or not they have internet access, and so on. Whatever essential dimensions are available are used in these operations to adjust sample data to look like the population they are intended to represent.

Just Do It … or Not!

Do all clients, especially our non-academic and non-government clients, really want to know the rationale for their data weights, let alone how they were calculated? Not usually. Under the pressure of time, it is not an uncommon situation to be characterized by the call to "weight my data, please! But spare me the gory details." Hence, the responsibility and trust we shoulder at KN. Not all survey vendors either care to or know how to weight data accurately and thoroughly. Some may apply only post-stratification weights on just a few "obvious" dimensions and consider the job done, delivering modestly small weights as a pretense for quality data. Train your clients to expect this and bliss reigns, albeit delusional. We at KN work to educate our clients about data weights and make our procedures as transparent as possible.

Three-Part Harmony

Understanding what goes into the KN weights requires understanding three critical pieces of information. First is how we sample the U.S. population to construct and maintain KnowledgePanel^{SM}. Second is how the sample for a particular study is arrived at. And third is adjusting the actual respondents to each study to correctly look like the population they are designed to represent. Although it can get more complicated, these are essentially the three parts to the multidimensional weighting harmony. Most people only think of that third part of weighting — making respondents look like the population of interest after the data collection is over. The study population could be anything from grocery shoppers, sports enthusiasts, beer drinkers, persons with a particular health condition, to the nation's general adult population. Whatever and whomever is being studied, only final weights incorporating all three parts deliver the accurate picture with all segments properly adjusted and in their correct relative proportion to each other.

The Base Part

If you think of the U.S. population as a vast and deep ocean, KnowledgePanel^SM is about a bathtub-sized sample of ocean water. However, that sample is constructed from a multitude of thimble-sized samples drawn from across the length, breadth and depth of that larger body of water. In a perfect world, every part of the ocean gets sampled, and all those thimbles show up filled to the brim. But, alas, perfection is elusive. Despite the fact that we use the very best random-digit dial (RDD) telephone survey methods to recruit our panel, only about 85% of U.S. households have landline phones for us to call. It is also impossible to make contact with all the households we repeatedly dial in the RDD sample. There is also disproportionate success in recruiting participants from across all geographic, race/ethnic, and socio-economic groups that make up our diverse national population. To address all these disparities, we statistically re-adjust the dynamically changing panel every three weeks to reflect an accurate representative picture of that drifting ocean of a national population. We do this using monthly revised data released from the U.S. Census Bureau plus our own accounting for natural changes in the panel's composition due to attrition and newly recruited members from national RDD samples. This intensive adjustment effort produces that first or "base" weight attached to every panel member.

The Middle Part

The next step is driven by how your sample is drawn from KnowledgePanel^SM in order to meet the needs of your study. This is like extracting a glass of water from the bathtub — or a bucket, if it is either a very large sample or one requiring us to screen many participants to find a small and very rare group. The vagaries of sampling, especially involving small samples, demand yet another set of adjustments to keep this sample in representative balance for the population of interest that may or may not be of the general-population variety. This can be where complications set in, but I will spare you that torture.

The After Part

Finally, since not everyone selected for your sample actually completes your questionnaire, the final sample requires that third stage adjustment that brings it all home and moves your data cases into their true representational glory. This is yet another complex weighting procedure. This simultaneously and correspondingly adjusts every case across multiple dimensions so that each respondent's contribution to the results is proportionately correct relative to their size in your study population. This final and quintessential weight is a multiplication factor incorporating all the weights from the previous parts. Additionally, it is not unusual for studies to require many uniquely calculated weights so that specific segments of the sample can be independently analyzed.

If the Shoe Fits

My grandfather, an Italian immigrant, was a shoemaker. He had a multitude of colorful metaphors that sprang from his humble trade when he spoke about life and the world. In that tradition, let me say that evaluating the weights in your data file is like evaluating a pair of shoes. Style and materials aside, it all comes down to the right size and comfort. Driven by the sample design, the number of segments studied, the sample sizes within segments, the response patterns, etc., the range of weights for a study can be wide. It is not an unusual practice for good statisticians to conservatively trim the few outlier weights to be no less than the lower 1.0 percentile and no more than the upper 99.0 percentile of the distribution of weights. This may still leave weights ranging from a small fraction to somewhat larger numbers up in the teens. Yet, these are the weights, and they are a product of the characteristics of the study, sample and response. Depending on the analyses to be

applied, these weights, whatever their size, are to be used to get the most accurate answer. When we deliver weights to our clients, they must pass our comfort test that the size and range is reasonable, not too unusual, and correctly developed. We know they are accurate, and we tell our clients to use them with confidence. To use the survey data without these weights would be like wearing a shoe without a sole. At first glance, it may look good, but it will fail its intended purpose.

Now don't say I didn't tell you!

Dr. Charles DiSogra is Knowledge Networks' Chief Statistician and heads the statistics unit at the Menlo Park office.

He brings over 20 years experience in survey research, sample design, data analysis, and administration. Dr. DiSogra has a masters degree in public health and a doctorate in nutritional epidemiology with an emphasis in biostatistics and policy analysis from the University of California at Berkeley. He can be reached at cdisogra@knowledgenetworks.com.