Credibility Intervals for Online Polling

Statistical margins of error do not apply to online non-probability polls. Ipsos has adopted Bayesian Credibility Intervals as our standard for reporting our confidence in online polling results.

Why is the classical approach to margins of error not appropriate for non-probability-based online polling?

To produce an accurate margin of error, one must know the probability of participation for each member of the survey population, or, everyone in the population must have a known chance of participating in the survey.

Except when using probability-based online panel samples, online polling among consumers or citizens does not meet this condition because of two effects: non-response bias and coverage bias. To use classical margins of error for online polling, one would have to assume that non-responders to online surveys are completely random, or that the effect of leaving out non-users is so minimal that it can be overlooked. We know that both of these are false, as we can make the following claims about online polling and those who complete online surveys:

- Not everyone has internet access;
- Non-probability-based online samples are created through an opt-in process, rather than through a random-digit-dialing or address-based sampling recruiting methodologies;

Less is known about the profiles of individuals who complete online surveys versus those who do not, or about the likelihood of an online person to complete an online survey or to participate in a survey panel.

Therefore, the probability of being included in any given opt-in online survey sample is unknown, very difficult to ascertain, or simply zero (non-internet users). Further, the nature of use of the internet is not uniform within the population, so this limits one's ability to calculate the likelihood of reaching a person through an online poll. In short, without this knowledge, a margin of error cannot be calculated.

Despite these challenges, online polling based on opt-in samples conducted scientifically has proven to yield similar results to probability sampling conducted via telephone. Put simply, online polling works! This poses a challenge, however, because it is also very important for us as researchers to provide a measure of confidence along with our data, regardless of the method of data collection. As such, Ipsos considers the methodology valid and uses of Bayesian Credibility Intervals to describe the relative robustness and uncertainty of an opt-in sample-based online survey estimate.

Introduction to Bayesian Credibility Intervals

Bayesian Credibility Intervals operate much in the same way as classical margins of error, but acknowledge the uncertainty of an estimate (in our case, the probability of any given person to complete an online survey), and incorporate external factors, such as what we know about the world, opinions, expertise, context, history and other data into its calculation to correct for the unknown. Bayesian models allow us to generalize from a sample to a population since they correct for unbalanced samples due to non-response, coverage, or other biases. One does not need to know the actual probabilities of selection, which, as described, are nearly impossible to ascertain in opt-in online polling.

For this approach to hold true, one must assume that the sample design is "conditionally ignorable", meaning that, once we control for the biases mentioned above, there is not a relationship between one's likelihood to participate in an online survey and the variables we want to measure. Ipsos is taking steps to ensure that its online samples are conditionally ignorable, such as combining multiple opt-in online panel and non-panel sources. In essence, by combining multiple sample sources, the "holes" in any one sample source can be filled by the other source.

Bayesian **Credibility Intervals** measure the **degree of certainty one has in the results based on one's experience, understanding, and knowledge of the population, tempered by the data that has been observed**. We call this the "knowledge base", and can include other published surveys, past surveys by the same pollster, historical political outcome data, a practitioner's political expertise, predictive models, and other sources. This knowledge base is continuously updated with new information.

An example of an output statement of such an approach reads as follows:

"These are findings from an Ipsos poll conducted for [CLIENT] from [DATES}. For the survey, a sample of [N=XXX] adults residing in [COUNTRY] was interviewed online. The precision of online polls is measured using a credibility interval. In this case, the poll has a credibility interval of plus or minus XX percentage points. For more information about credibility intervals, please see [insert link]. The data were weighted to the [COUNTRY] population data by [DEMOGRAPHIC VARIABLES]. Statistical margins of error do not apply to online polls. All sample surveys and polls may be subject to other sources of error, including, but not limited to coverage error and measurement error. Where figures do not sum to 100, this is due to the effects of rounding".

In summary, Ipsos is using the Bayesian approach to frame the problem of margins of errors for non-probability sampling in such a way to allow us to use the wealth of information available outside of the current survey to calibrate survey results and to provide a statistical foundation for inference even when a probabilistic sample is unavailable.

How to Calculate Bayesian Credibility Intervals

The calculation of credibility intervals assumes that Y has a binomial distribution conditioned on the parameter θ\, i.e., Y|θ~Bin(n,θ), where n is the size of our sample. In this setting, Y counts the number of "yes", or "1", observed in the sample, so that the sample mean ($\bar{y}$) is a natural estimate of the true population proportion θ. This model is often called the likelihood function, and it is a standard concept in both the Bayesian and the Classical frameworks. The Bayesian1 statistics combines both the prior distribution and the likelihood function to create a posterior distribution. The posterior distribution represents our opinion about which are the plausible values for θ adjusted after observing the sample data. In reality, the posterior distribution is one's knowledge base updated using the latest survey information. For the prior and likelihood functions specified here, the posterior distribution is also a beta distribution (π(θ/y)~β(y+a,n-y+b)), but with updated hyper-parameters.

Our credibility interval for θ is based on this posterior distribution. As mentioned above, these intervals represent our belief about which are the most plausible values for θ given our updated knowledge base. There are different ways to calculate these intervals based on $\pi(\theta/y)$. Since we want only one measure of precision for all variables in the survey, analogous to what is done within the Classical framework, we will compute the largest possible credibility interval for any observed sample. The worst-case occurs when we assume that a=1 and b=1 and $y = n/2$. Using a simple approximation of the posterior by the normal distribution, the 95% credibility interval is given by, approximately: $\bar{y} \mp \frac{1}{\sqrt{n}}$

For this poll, the Bayesian Credibility Interval was adjusted using standard weighting design effect 1+L=1.3 to account for complex weighting[2].

Examples of credibility intervals for different base sizes are below.

Sample size	Credibility intervals
2,000	2.5
1,500	2.9
1,000	3.5
750	4.1
500	5.0
350	6.0
200	7.9
100	11.2

[1] *Bayesian Data Analysis, Second Edition, Andrew Gelman, John B. Carlin, Hal S. Stern, Donald B. Rubin, Chapman & Hall/CRC | ISBN: 158488388X | 2003*

[2] *Kish, L. (1992). Weighting for unequal Pi . Journal of Official, Statistics, 8, 2, 183200.*