

ASA Series
What Is a Survey?

What Is a Survey?
How to Plan a Survey
How to Collect Survey Data
Judging the Quality of a Survey
How to Conduct Pretesting
What Are Focus Groups?
More About Mail Surveys
What Is a Margin of Error?
Designing a Questionnaire
More About Telephone Surveys



Produced by
Section on Survey Research Methods
American Statistical Association
1429 Duke Street
Alexandria, VA 22314-3415 USA

ASA Series
What Is a Survey?

What is a Margin of Error?



Section on Survey Research Methods
American Statistical Association

What is a Margin of Error?

When results of surveys are reported in the media, they often include a statement like

“55 percent of respondents favor Ms. Smith in the upcoming mayoral election. There is a margin of error of 3 percentage points.”

What does a statement like this mean? This pamphlet attempts to answer this question and to provide some cautions on the use of the “*margin of error*” as the sole measure of a survey’s uncertainty.

Surveys are typically designed to provide an estimate of the true value of one or more characteristics of a population at a given time. The target of a survey might be

An estimate from a survey is unlikely to exactly equal the true population quantity of interest.

- the *average value* of a measurable quantity, such as annual 1998 income or SAT scores for a particular group.
- a *proportion*, such as the proportion of likely voters having a certain viewpoint in a mayoral election

This pamphlet, **What Is a Margin of Error?** is the eighth in the ASA series **What Is a Survey?** Here we will attempt to define the often-used journalistic phrase “Margin of Error.” Our treatment is sometimes more mathematical than in the other pamphlets in the series; still it should be accessible to most readers. Among the topics covered are how the number of observations in the sample, the type of sampling, and population size affect the margin of error. Some cautions are also provided on employing the margin of error as the sole measure of a survey’s uncertainty.

The **What is a Survey?** series is written primarily for the general public. Its overall goal is to improve survey literacy among individuals who participate in surveys or use survey results. The series is designed to promote a better understanding of what is involved in carrying out sample surveys—especially those aspects that have to be taken into account in evaluating the results of surveys.

■ the *percentage* of children under three years of age immunized for polio in 1997

An estimate from a survey is unlikely to exactly equal the true population quantity of interest for a variety of reasons. For one thing, the questions maybe badly worded. For another, some people who are supposed to be in the sample may not be at home, or even if they are, they may refuse to participate or may not tell the truth. These are sources of *“nonsampling error.”*

But the estimate will probably still differ from the true value, even if all nonsampling errors could be eliminated. This is because data in a survey are collected from only some—but not all—members of the population to make data collection cheaper or faster, usually both.

Suppose, in the mayoral election poll mentioned earlier, we sample 100 people

“Sampling error” means that results in the sample differ from a target population quantity, simply due to the “luck of the draw.”

who intend to vote and that 55 support Ms. Smith while 45 support Mr. Jones. This would seem to suggest that a majority of the town's voters, including people not sampled but who will vote in the election, would support Ms. Smith.

Of course, just by chance, a majority in a particular sample might support Ms. Smith even if the majority in the population supports Mr. Jones. Such an occurrence might arise due to *“sampling error,”* meaning that results in the sample differ from a target population quantity, simply due to the “luck of the draw”—*i.e., by which set of 100 people were chosen to be in the sample.*

Does sampling error render surveys useless? Fortunately, the answer to this question is “No.” But how should we summarize the strength of the information in a survey? That is a role for the margin of error.

Margin of Error Defined

The “margin of error” is a common summary of sampling error, referred to regularly in the media, which quantifies uncertainty about a survey result. The margin of error can be interpreted by making use of ideas from the laws of probability or the “laws of chance,” as they are sometimes called.

Surveys are often conducted by starting out with a list (known as the “sampling frame”) of all units in the population and choosing a sample. In opinion polls, this list often consists of all possible phone numbers in a certain geographic area (both listed and unlisted numbers).

In a scientific survey every unit in the population has some known positive probability of being selected for the sample, and the probability of any particular sample being chosen can be calculated. The beauty of a probability

In a scientific survey every person in the population has some known positive probability of being selected into a sample.

sample is twofold. Not only does it avoid biases that might arise if samples were selected based on the whims of the interviewer, but it also provides a basis for estimating the extent of sampling error. This latter property is what enables investigators to calculate a “margin of error.” To be precise, the laws of probability make it possible for us to calculate intervals of the form

estimate +/- margin of error.

Such intervals are sometimes called 95 percent confidence intervals and would be expected to contain the true value of the target quantity (in the absence of non-sampling errors) at least 95 percent of the time.

Such intervals are sometimes called 95% confidence intervals and would be expected to contain the true value at least 95% of the time.

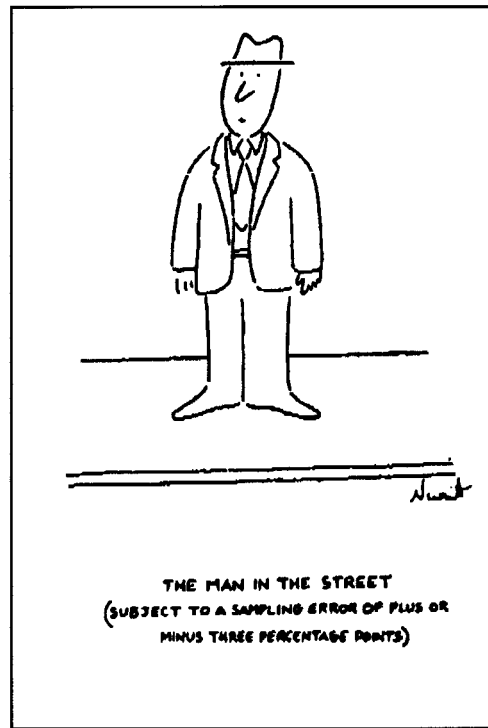
An important factor in determining the margin of error is the size of the sample. Larger samples are more likely to yield results close to the target population quantity and thus

have smaller margins of error than more modest-sized samples.

In the case of the mayoral poll in which 55 of 100 sampled individuals support Ms. Smith, the sample estimate would be that 55 percent support Ms. Smith—however, there is a margin of error of 10 percent. Therefore, a 95 percent confidence interval for the percentage supporting Ms. Smith would be

(55%-10%) to (55%+10%)

or (45 percent, 65 percent), suggesting that in the broader community the support for Ms.



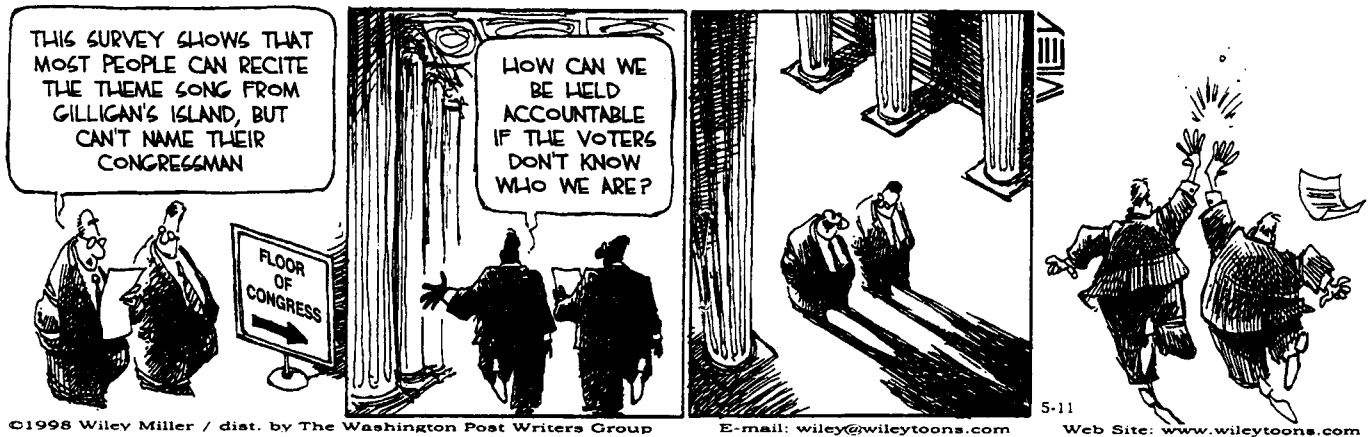
1988, The New Yorker Collection. Reprinted with permission.

Smith could plausibly range from 45 percent to 65 percent. This implies—because of the small sample size—considerable uncertainty about whether a majority of townspeople actually support Ms. Smith.

Instead, if there had been a survey of 1,000 people, 550 of whom support Ms. Smith, the sample estimate would again be 55 percent, but now the margin of error for Ms. Smith's support would only be about 3 percent. A 95 percent confidence interval for the proportion supporting Ms. Smith would thus be

(55%-3%) to (55%+3%)

or (52 percent, 58 percent), which provides much greater assurance that a majority of the town's voters support Ms. Smith.



1988, Washington Post Writer's Group. Reprinted with permission.

What Affects the Margin of Error?

Three things that seem to affect the margin of error are sample size, the type of sampling done, and the size of the population.

Sample Size—As noted earlier, the size of a sample is a crucial actor affecting the margin of error. In sampling, to try an estimate a population proportion—such as in telephone polls—a sample of 100 will produce a margin of error of no more than about 10 percent, a sample of error of 500 will produce a margin of error of no more than about 4.5

percent, and a sample of size 1,000 will produce a margin of error of no more than about 3 percent. This illustrates that there are diminishing returns when trying to reduce the margin of error by increasing the sample size. *For example*, to reduce the margin of error to 1.5% would require a sample size of well over 4,000.

Telephone surveys often rely on the technique of "random digit dialing."

Probability Sampling Designs—The survey researcher also has control over the design of the sample, which can affect the margin of error. Three common types are simple random sampling, random digit dialing, and stratified sampling.

- **A simple random sampling** design is one in which every sample of a given size is equally likely to be chosen. In this case, individuals might be selected into such a sample based on a randomizing device that gives each individual a chance of selection. Computers are often used to simulate a random stream of numbers to support his effort.
- Telephone surveys that attempt to reach not only people with listed phone numbers but also people with unlisted numbers often rely on the technique of **random digit dialing**.
- **Stratified sampling designs** involve defining groups, or strata, based on characteristics known for everyone in the population, and then taking independent samples within each stratum. Such a design offers flexibility, and, depending on the nature of the strata, they can also improve the precision of estimates of target quantities (or equivalently, reduce their margins of error).

Stratified samples are especially advantageous ... when the target is to estimate differences in viewpoints between different groups.

Of the three types of probability sampling, stratified samples are especially advantageous when the target of the survey is not necessarily to estimate the proportion of an entire population with a particular viewpoint but instead is to estimate differences in viewpoints between

different groups. *For example*, if there was a desire to compare attitudes between individuals of Inuit (Alaskan native) origin versus other Americans on their opinion about drilling for oil on federal land, it would not make sense to take a simple random sample of all Americans to answer this question because very few Inuit would likely fall into such a sample. Instead, one might prefer to take a stratified sample in which Alaskan Native Americans compose one half of the sample and non-Inuit compose the other one half.

Sometimes samples are drawn in clusters in which only a few counties or cities are sampled or only the interviewer visits a few blocks. This tends to increase the margin of error and should be taken into account by whoever calculates sampling error.

Size of Population—Perhaps surprising to some, one factor that generally has little influence on the margin of error is the size of the population. That is, a **sample size of 100 in a population of 10,000 will have almost the same margin of error as a sample size of 100 in a population of 10 million.**

Interpreting the Margin of Error

In practice, nonsampling errors occur that can make the margin of error reported for a poll smaller than it should be if it reflected all sources of uncertainty. *For example*, some respondents to the mayoral survey may not have been eligible to vote but may have answered anyway, while others may have misled the interviewer about their preferences.

Why isn't the margin of error adjusted to reflect both sampling and nonsampling uncertainties? The answer is that, unlike sampling error, the extent of nonsampling error cannot usually be assessed from the sample itself, even if the sample is a probability sample.

Some things that help assess nonsampling uncertainties, when available, include the percentage of respondents who answer "don't know" or "undecided." Be wary when these quantities are not given. Almost

Nonsampling errors occur that can make the margin of error reported for a poll smaller than it should be to reflect all sources of uncertainty.

always there are people who have not made up their mind. How these cases are handled can make a big difference. Simply splitting them in proportion to the views of those who gave an opinion can be misleading in some settings.

It is important to learn if the survey results are actually from a probability sample at all. Many media surveys are based on what are called quota samples, and, although margins of error are reported from them, they do not strictly apply.

Overall, nonresponse in surveys has been growing in recent years and is increasingly a

consideration in the interpretation of reported results. Media stories typically do not provide the response rate, even though these can be well under 50 percent. When the results are important to you, always try to learn what the non-response rate is and what has been done about it.

For more information, see “Judging the Quality of a Survey”—an earlier brochure in the *What Is a Survey?* series.

Keep Your Eye on What is Being Estimated

It is common for political polls to quote a margin of error of plus or minus 3 percent. It might happen, however, that in two separate polls between Jones and Smith in the same week one might have Jones ahead by 2 percent in one poll while the other poll might have Jones ahead by 10 percent. How can this be?

A misleading feature of most current media stories on political polls is that they report the margin of error associated with the proportion favoring one candidate, not the margin of error of the **lead** of one candidate over another. To illustrate the problem, suppose one poll finds that Mr. Jones has 45 percent support, Ms. Smith has 41 percent support, 14 percent are undecided, and there is a 3 percent margin of error for each category.

If we note that Mr. Jones might have anywhere from 42 percent to 48 percent support in the voting population and Ms. Smith might have anywhere from 38 percent to 44 percent

Overall nonresponse in surveys has been growing in recent years and is increasingly a consideration in the uncertainty of reported results.

support, then it would not be terribly surprising for another poll to report anything from a 10-point lead for Mr. Jones (such as 48 percent to 38 percent) to a 2-point lead for Ms. Smith (such as 44 percent to 42 percent).

In more technical terms, a law of probability dictates that the difference between two uncertain proportions (*e.g.*, the lead of one candidate over another in a political poll in which both are estimated) has more uncertainty associated with it than either proportion alone. Accordingly, the margin of error associated with the lead of one candidate over another should be larger than the margin of error associated with a single proportion, which is what media reports typically mention (**thus the need to keep your eye on what's being estimated!**).

Until media organizations get their reporting practices in line with actual variation in results across political polls, a rule of thumb is to multiply the currently reported margin of error by 1.7 to obtain a more accurate estimate of the margin of error for the lead of one candidate over another. Thus, a reported 3 percent margin of error becomes about 5 percent and a reported 4 percent margin of error becomes about 7 percent when the size of the lead is being considered.

It is common for political polls to quote a margin of error of plus or minus 3%.

There is a lot more to be said about the use of the term “margin of error.” Surprisingly, there is even some controversy about its meaning. For those interested in reading more about this controversy, a Sunday, June 14, 1998, “Unconventional Wisdom” column by Richard Morin in *The Washington Post* may be a good start.

Where Can I Get More Information?

In addition to the pamphlets in this series, ASA also makes other brochures available upon request:

■ ***Ethical Guidelines for Statistical Practice***

■ ***Surveys and Privacy***, produced by the ASA Committee on Privacy and Confidentiality.

For the above brochures or other pamphlets in the *What Is a Survey?* series, contact:

Section on Survey Research Methods

American Statistical Association

1429 Duke Street

Alexandria, VA 22314-3415 USA

(703) 684-1221/fax: (703) 684-2037

Email: asainfo@amstat.org

Web site: [http://www.amstat.org/
sections/srms/](http://www.amstat.org/sections/srms/)

Besides the ASA, there are many other associations that are concerned with the proper collection and use of survey data:

■ **The American Association for Public Opinion Research** (AAPOR) offers a number of publications—perhaps the most relevant of these is the one entitled ***Best Practices for Survey and Public Opinion Research Survey Practices AAPOR Condemns***. To obtain copies, call (313) 764-1555 or visit their Web site at <http://www.aapor.org>.

■ **The National Council on Public Polls** publishes another useful pamphlet, ***Twenty Questions a Journalist Should Ask About Poll Results***. To obtain a copy, call (800) 239-0909.

■ **The Research Industry Coalition, Inc.**, publishes a brochure, ***Integrity and Good Practice in Marketing and Opinion Research***. To obtain a copy, call (516) 928-6803.

■ **The Council of American Survey Research Organizations** publishes a pamphlet, ***Surveys and You***. To obtain a copy, call (516) 928-6954, or visit their Web site at <http://www.casro.org>.