



Random Sampling Error ^[1]

Assisted Self-Help ^[2] 78.7K reads

Random sampling errors are one type of experimental error that everybody should know.

Anyone who reads polls on the internet, or in newspapers, should be aware that sampling errors ^[3] could vastly influence the data and lead people to draw incorrect conclusions ^[4].

To further compound the random sampling errors, many survey ^[5] companies, newspapers and pundits are well aware of this, and deliberately manipulate polls to give favorable results.

In any experiment ^[6] where it is impossible to sample an entire population, usually due to practicality and expense, a representative sample must be used.

Of course, when you use a sample group, it can never fully match the entire population, and there will always be some likelihood of random sampling error ^[7].

Any researcher must strive to ensure that the sample is as representative as possible, and statistical ^[8] tests have inbuilt checks and balances to take this into account.

To illustrate how to ensure that your statistics ^[8] are as accurate as possible, we are going to use the example of an opinion poll. These are one of the most commonly misinterpreted representations of data, and failure to take into account the nuances of statistics can paint an incorrect picture.

Margin of Error - A False Picture

The problem is, when you see an opinion poll in a newspaper or internet site, you will usually see a margin of error ^[9], such a + or - 3%. The temptation is to think that the polls will be accurate within this figure.

For example, if a poll gives one political party (A) a 42% share of the vote, and the other (B) 39%, this opens up a number of possible results. (A) could have 45%, (B) 36%. Both could be 39% or (B) could actually be ahead, 42% versus 39%. Of course, the results could show any variation in between those extremes.
Complicated enough?

To complicate the picture further, even this random sampling error can be wildly inaccurate. Any opinion poll may give the margin of error, but this can convey a false sense of security and make people assume that the results 'must' lie within this range.

In fact, these figures could actually be completely wrong, and the numbers are only ever an estimate.

The Problem With Random Sampling Error

The problem is that these results only show the random sampling error within that specific group. They show the chances of the results in that group occurring purely by chance, exactly like the 95% confidence margin ^[10] employed by many scientific researchers.

However, this is a very narrow definition and is often misunderstood.

In an opinion poll, there is no guarantee that the sample of 1000 or 10 000 people is truly representative of the larger population as a whole.

There have been many extremely inaccurate polls conducted over the years, and they fell down due to poor design and not understanding all of the relevant factors.

For example, an opinion poll company conducting telephone polls may make the mistake of only telephoning during office hours, when most of the population is at work, skewing the data.

In addition, poorer families do not always have a fixed line telephone and use unregistered cell phones, again leaving a huge potential for inaccuracy. The margins of error would be perfectly acceptable, in these cases, but the overall findings would still be horribly wrong.

Modern polling companies are very skilled at designing polls to select samples ^[11] from many elements of the population, and via various media, so big errors rarely happen. Despite this, opinion polls must always be taken as a guide only, not an exact representation of how an election is likely to unfold.

Random Sampling Error and Experimental Design

The mistakes made by pollsters relate directly to any type of experiment ^[12] involving random sample groups.

Statistics can only work with the data ^[13] provided and, if your design is poorly thought out, will not be able to cover up these errors ^[14]. Garbage in definitely equals garbage out.

Bibliography

Husch, B. (1971). *Planning a Forest Inventory*. Rome, Italy: Food and Agriculture Organization of the United Nations

Urdan, T.C. (2005). *Statistics in Plain English*, Mahwah, NJ: Lawrence Erlbaum

Weisberg, H.F. (2005). *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: University of Chicago Press

Links

- [1] <https://staging.explorable.com/en/random-sampling-error>
- [2] <https://staging.explorable.com/en>
- [3] <https://staging.explorable.com/sampling-error>
- [4] <https://staging.explorable.com/drawing-conclusions>
- [5] <https://staging.explorable.com/survey-research-design>
- [6] <https://staging.explorable.com/experimental-research>
- [7] <http://www.pollingreport.com/sampling.htm>
- [8] <https://staging.explorable.com/statistics-tutorial>
- [9] <https://staging.explorable.com/statistics-margin-of-error>
- [10] <https://staging.explorable.com/statistics-confidence-interval>
- [11] <https://staging.explorable.com/population-sampling>
- [12] <https://staging.explorable.com/experimental-error>
- [13] <https://staging.explorable.com/statistical-data-sets>
- [14] <https://staging.explorable.com/type-i-error>