

Estimating the Validity of a Poll

As we lead up to elections we usually have some group poll the electorate and say that the results are accurate to $\pm 4\%$, 19 times out of 20. At other times, when marketing or someone else, does a poll, how accurate is the result? How many people should you poll (this is called the sample space) if you want $\pm 5\%$ accuracy 99 times out of 100? Remember, you only get 100% accuracy if you poll every person in the population.

Well it just so happens, there are a couple of easy formulas that answer these questions. The formula is: Divide 98% by the square root of the number of people polled, and you get the 95% margin of error of the survey. A 95% level of margin of error means that you are sure 19 times out of 20 that you are a certain percentage of the correct result if you polled all of the population. So the formula could be written as:

$$M = \frac{98\%}{\sqrt{N}}, \text{ where "M" = Margin of error, and "N" = Number of people polled. If you want more accuracy,}$$

and would like to get to a 99% margin of error (99 times out of 100), then the formula now changes to:

$$M = \frac{129\%}{\sqrt{N}}$$

I got this formula from a really neat book: *Struck by Lightning* by Jeffrey Rosenthal. Here are some examples of using this formula.

(1) A poll had a sample size of 1 100 people. How accurate are the results? Using the formulas:

$$M = \frac{98\%}{\sqrt{1100}} = 2.95\%. \text{ The results of the poll are within 2.95\% of the actual answer 19 times out of 20.}$$

$$M = \frac{129\%}{\sqrt{1100}} = 3.89\%. \text{ The results of the poll are within 3.89\% of the actual answer 95 times out of 100.}$$

(2) I need to accurate to 2%, 19 times out of 20 (95% margin of error). How many people should I poll?

$$2\% = \frac{98\%}{\sqrt{N}}, \text{ so multiply both sides by } \sqrt{N}$$

$$2\% \sqrt{N} = 98\%, \text{ now divide both sides by 2\%}$$

$$\sqrt{N} = \frac{98\%}{2\%} = 49, \text{ now square both sides}$$

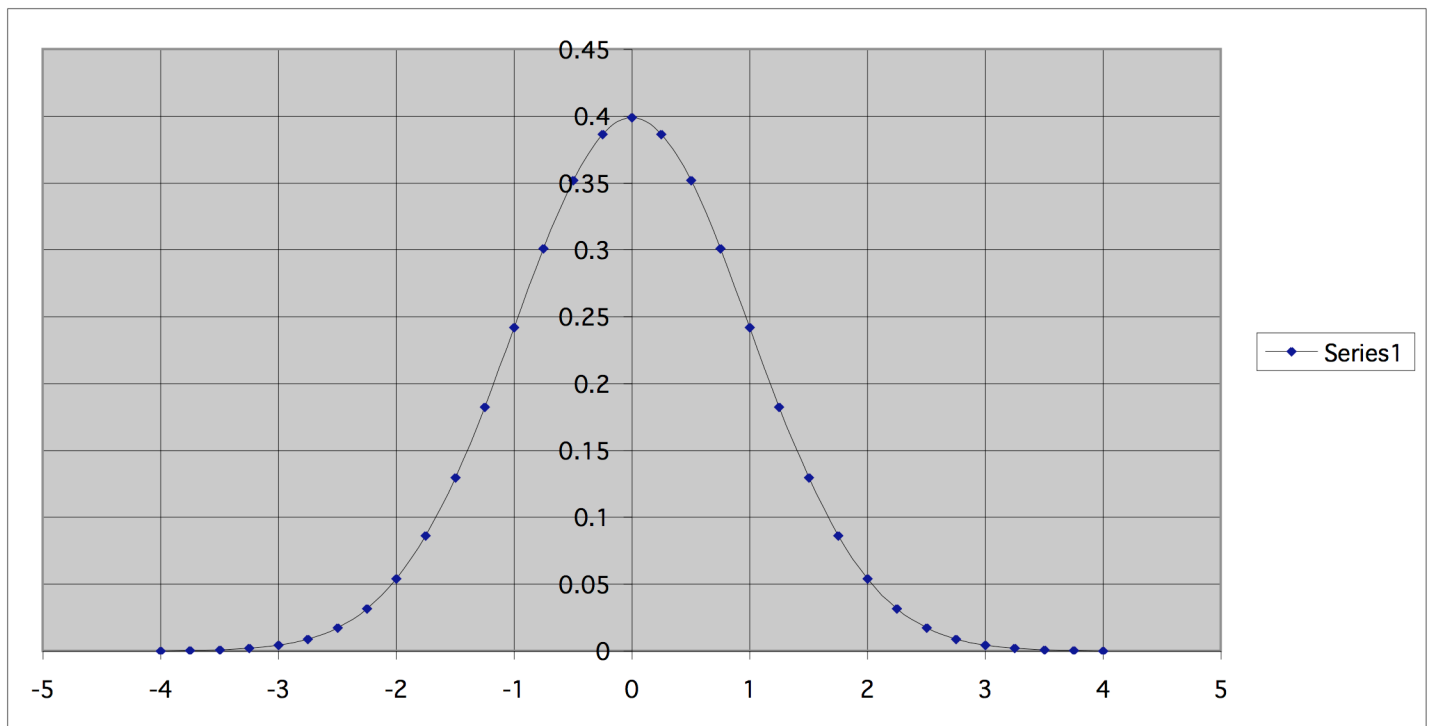
$$N = 49^2 = 2,401$$

So, I need to poll about 2 400 people in order to get a 2% accuracy 19 times out of 20. On the next page is a small table that gives you the number "N" and what the 95% and 99% margin of errors are. Following it, I will show you how the formula was worked out.

By looking at the following little table, one can see how the pollsters can get a fairly good accuracy of a population by only polling a small segment of them. (assuming that they choose their sample space properly).

Number Polled	95% Margin of error	99% Margin of error
200	6.93%	9.12%
400	4.90%	6.45%
600	4.00%	5.27%
800	3.46%	4.56%
1000	3.10%	4.08%
1200	2.83%	3.72%
1400	2.62%	3.45%
1600	2.45%	3.23%
1800	2.31%	3.04%
2000	2.19%	2.88%
2200	2.09%	2.75%
2400	2.00%	2.63%

Now, here is how the formulas were worked out. On the graph below, we see the “Normal” Curve. The numbers across the bottom are the number of Standard Deviations that occur to the right (positive) or to the left (negative) of the mean. On a large population, there are 68.3% of the population is between -1 and $+1$ standard deviations, and 95.4% of the population is between -2 and $+2$ standard deviations.



If you want the middle 95%, then you have 2 ½% on either side. This corresponds to 1.959963986 standard deviations. If I take half of this, I get 0.979981993 or 98%.

If you want the middle 99%, then you have ½% on either side. This corresponds to 2.575829303 standard deviations. If I take half of this, I get 1.287914652 or 129%.