The ‘Margin of Error’ for Differences in Polls

Charles H. Franklin
University of Wisconsin, Madison

October 27, 2002
(Revised, February 9, 2007)

The “margin of error” for a poll is routinely reported.[1] But frequently we want to know about the *difference* between two proportions (or percentages). Often the question concerns differences between two responses to the same question within a single poll. For example, what is the lead of one candidate over another in an election poll. The second common question is whether a proportion has changed from one poll to the next. For example, has presidential approval increased from one poll to the next. The margin of error for these *differences* is not the same as the margin of error for the poll, which is what virtually all polls routinely report. The margin of error for the poll is for a single proportion, not differences. This leads to considerable confusion among reporters and interpreters of polls.

This note explains the correct way to calculate the margin of error (and hence the “significance”) for differences of proportions in polls. There is also a “quick reference” section at the end that provides the formulas in a single spot.

1 The Margin of Error for a Single Proportion

The usual “margin of error” for a poll is the 95% confidence interval for an individual proportion.[2]

The formula for the variance of a proportion, p, is

$$\mathrm{Var}(p) = \frac{pq}{n-1}$$

where $q = 1 - p$.[3] We usually use just n in the denominator since n is rather large for

[1] In fact this is a misnomer. Each item in a poll has a margin of error which depends on the sample design and the distribution of responses to the item. We characterize the “margin of error for the poll” by arbitrarily picking one type of item, a dichotomous item equally split, as a conservative estimate of the margin of error for all items. This practice is a convenient rule of thumb but may overstate (or in some cases understate) the actual margin for a particular item.

[2] At least those not too close to 0% or 100% where the confidence interval becomes asymmetric. We usually ignore such technical issues.

[3] This assumes “simple random sampling”. Some polls use stratification of the sample to increase the precision of the sample, for example stratifying by region. Others may cluster the interviews, which decreases the precision. Clustering is rarely done for telephone polls but is universal practice for face-to-face interviews. The typical telephone poll may use some stratification, but the details of the sampling design are rarely made public. The simple random sampling assumption is used for almost all reported margin of error calculations.

virtually all survey samples so the difference in n and $n-1$ is trivial. The standard error for the proportion is therefore

$$\mathrm{se}(p) = \sqrt{\frac{pq}{n}}$$

The 95% confidence interval (usually called the "margin of error" of the poll) is $\pm 1.96 \times \mathrm{se}(p)$, using the normal distribution approximation for large samples.

The standard error depends on the proportion, p, and is at a maximum for $p = .5$, so a quick approximation of the widest confidence interval for a single proportion is

$$\begin{aligned} \mathrm{CI}(p) &= \pm 1.96 \times \sqrt{\frac{.5 \times .5}{n-1}} \\ &\approx \pm 2 \times \frac{.5}{\sqrt{n}} \\ &\approx \pm \frac{1}{\sqrt{n}} \end{aligned}$$

This is usually what is reported as the margin of error for a poll. For example, if $n = 400$, $1/\sqrt{400} = .050$, a MOE of $\pm 5\%$. For $n = 625$, the MOE is $1/\sqrt{625} = .040$ and for $n = 1111$, the MOE is $1/\sqrt{1111} = .030$.

For proportions different from .5, the MOE is somewhat smaller. For example, if $p = .6$, then the MOE is approximately $2 \times \sqrt{.6 \times .4/625} = .039$, a trivial difference. As the responses become more skewed the MOE can be noticeably smaller, as for example if $p = .2$, $2 \times \sqrt{.2 \times .8/625} = .032$, or about a 3 point MOE compared to a 4 point margin for the $p = .5$ case. Still, unless we are looking at a highly skewed variable, these differences in the MOE are usually small enough to be ignored. Calculations for different distributions, such as these, are almost never reported in media accounts of polls.

2 Differences within a single poll question

The problem is that in political horse-race polls there is usually less interest in the proportion supporting a single candidate than in the lead one candidate has over another, the difference between the two proportions. When reporters (and political scientists) try to explain this they usually run into trouble because they don't use the correct formula for the confidence interval of the difference of proportions.

A voter preference survey question will always have at least three categories and often more. There will be response categories for the Republican and Democratic candidates plus at least an undecided category. There may also be categories for candidates for third parties and perhaps an "other" category as well. In presidential or other primary contests, there will usually be a number of candidates, often considerably greater than two. When responses fall into more than two discrete categories the proportions have a multinomial distribution.

In horse-race polls, we want to know the difference in proportions supporting the top two candidates (and perhaps between other pairs) and we need the confidence interval for this difference to tell if the lead is statistically significant (or "outside the margin of error"). Often pollsters, journalists and political scientists calculate this as twice the reported margin of error of the poll. This is done following the logic that if one candidate is at 55% and the other at 45% and the poll has a ±5% margin of error, then the first candidate *could* be as low as $55-5 = 50\%$ and the second *could* be as high as $45+5 = 50\%$. In this case reporters would say the race was a statistical dead heat because the gap between the candidates ($55 - 45 = 10\%$) is not more than two times the margin of error of the poll (5%). While this is the correct conclusion when there are only two possible survey responses, it is not correct when there are more than two possible responses, which is in fact virtually always the case. How much difference this makes depends on how many responses are outside the two categories of interest.

The correct formula for the variance of the difference of two multinomial proportions for candidates 1 and 2, p_1 and p_2, is (adapted from Kish, *Survey Sampling*, 1965, p. 498-501)

$$\mathrm{Var}(p_1 - p_2) = \frac{(p_1 + p_2) - (p_1 - p_2)^2}{n - 1}$$

The 95% confidence interval ("margin of error") for the difference of proportions is therefore

$$\begin{aligned}\mathrm{CI}(p_1 - p_2) &= 1.96\sqrt{\frac{(p_1 + p_2) - (p_1 - p_2)^2}{n - 1}} \\ &\approx 2\sqrt{\frac{(p_1 + p_2) - (p_1 - p_2)^2}{n}}\end{aligned}$$

Note that it doesn't matter what candidates $3 \ldots k$ have. We only need the proportions for the pair of candidates we care about in the formula. If there is considerable support for these other candidates then $p_1 + p_2$ will be a good deal less than 1.0, and this will shrink the standard error for the difference between p_1 and p_2, as we'll see below.

Whenever we compare proportions of candidate support within a single survey, this is the formula we should use. For low amounts of undecided or third party support the results will be close to the "twice the margin of error" formula, but the correct margin of error will be less than this as the proportion of "other" responses increases.

To see this effect Figure 1 plots the margin of error against the proportion of cases in the two categories of interest. For the example, I've picked a difference between p_1 and p_2 of eight points but let the total in $p_1 + p_2$ vary to illustrate how much the size of other categories matters for the MOE of the difference.

If $n = 600$, the margin of error for the sample is $1/\sqrt{600} = 0.0408$. Using the "twice the MOE of the sample" rule, we would find that the margin of error of the difference would be 0.0816. (If we were a bit pickier we would multiply by 1.96 instead of 2.0, and get .07997, and round that to .08, which is what is shown in the plot.) This is the margin of error for the difference if every respondent chose candidate 1 or candidate 2 and there were no "other" responses. In such a case, our hypothetical difference of 8 points between candidates 1 and 2 would be just barely statistically significant. But as the number of "other" responses grows (because of a single additional option, or because there are 11 other presidential candidates in the field) the margin of error declines (reading *right to left* in the figure.)

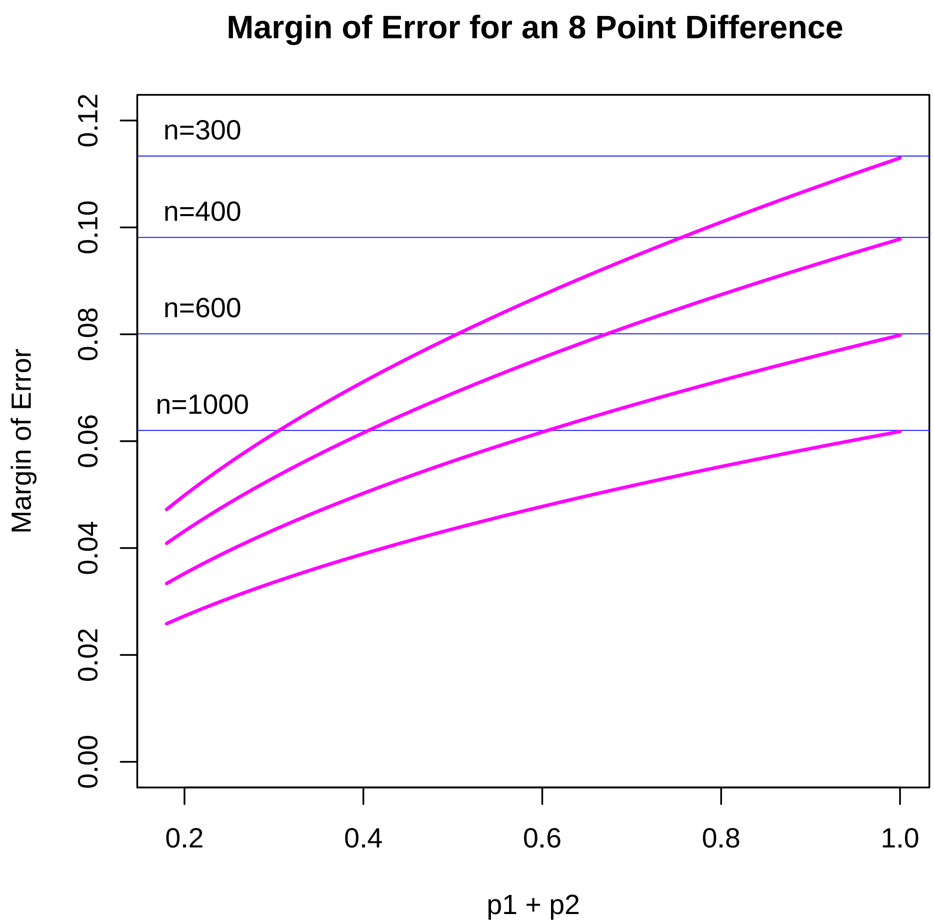

Figure 1: Margin of error depends on $p_1 + p_2$. The curved lines plot the margin of error as the proportion of the sample in $p_1 + p_2$ varies from 20% to 100% of the sample, for various sample sizes. The horizontal lines show the margin of error when the two categories are 100% of the sample, the "twice the margin of error of the poll" rule of thumb. The purple lines converge to this value as the proportion of the sample in the two categories of interest rises to 1.0. When $p_1 + p_2$ is less than 1.0, the correct margin of error is less, and sometimes substantially less, than twice the MOE of the sample. The precise margin of error also depends on the difference, $p_1 - p_2$, in the sample. Here the results are illustrated for an eight point difference.

3 Difference between two polls

A different issue is posed by the difference of proportions between two *independent* polls. Given two polls we want to know if opinion changed by a statistically significant amount from one poll to the next. Unlike the case within a single poll, here the percent support for a candidate in one poll is independent of that support in the other poll. (Because we draw independent random samples for each poll, this will be the case.)

Now the difference of interest is $p_2 - p_1$ where the subscripts 1 and 2 indicate *polls* 1 and 2, and we are measuring support for the *same* candidate in *both* polls. The variance of the difference with independent samples is

$$\begin{aligned} \mathrm{Var}(p_2 - p_1) &= \mathrm{Var}(p_1) + \mathrm{Var}(p_2) \\ &= \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2} \end{aligned}$$

where $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$, so the margin of error becomes

$$\mathrm{CI}(p_2 - p_1) = 1.96\sqrt{\left(\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right)}$$

This is not twice the margin of error for either poll (and the MOE for each poll may differ.) If the margin of error is the same for both polls, then the MOE for their difference is 1.41 times the MOE for the polls, not 2 times it.

Why the difference? Within a single poll, if support for the Democrat goes down, support for the Republican *must* go up if there are no undecided or third party voters. (If there are third alternatives then the proportions are likely to still be correlated, just not perfectly.) The formula for the difference of multinomial proportions given above takes this nonindependence between the proportions into account. In contrast when we compare across two independent samples, the proportion support estimated for the candidate in each poll is statistically independent because the samples are drawn independently. Hence a different formula for the independent samples case.

4 Examples

A Time magazine poll conducted January 22-23, 2007, surveyed 441 registered voters who are Republicans or lean Republican. The margin of error for the poll would be

$$\begin{aligned} MOE &= 1.96 \times \sqrt{\frac{.5 \times .5}{441 - 1}} \\ &= .0467 \\ &\approx .05 \text{ or } 5\% \end{aligned}$$

The Time poll found support for Senator John McCain at 30%. Former Mayor Rudy Giuliani received 26% support. The McCain lead over Giuliani is 4 points, but what is the margin of error? Using the “twice the MOE of the poll” rule, this would fall clearly short of $2 \times .0467 = .0934$ or just over 9 point MOE for the difference.

Using the correct formula for the MOE of a difference, we get

$$\begin{aligned}
\text{MOE(McCain} - \text{Giuliani)} &= 1.96 \times \sqrt{\frac{p_1 + p_2 - (p_1 - p_2)^2}{441 - 1}} \\
&= 1.96 \times \sqrt{\frac{.30 + .26 - (.30 - .26)^2}{440}} \\
&= .0698 \\
&\approx .07 \text{ or } 7\%
\end{aligned}$$

So the margin of error of the difference is 7 points and the difference of 4 points is not outside the MOE, and we conclude that the McCain lead is not statistically significant in this case.

An example where this distinction does matter is provided by a Public Policy Polling (PPP) survey of North Carolina Republican primary voters, taken February 5-6, 2007.[4] This poll of 735 respondents found support for Giuliani at 31% with former Speaker of the House Newt Gingrich at 25%. The margin of error for this poll is .036, using the standard calculation for a single question. If we use the "twice the MOE" rule, the MOE for the difference between Giuliani and Gingrich would be $2 \times .036 = .072$ or 7.2%. In that case, the six point Giuliani lead would not reach statistical significance. If we apply the correct formula for differences of multinomial proportions, using $p_1 = .31$ and $p_2 = .25$, then the MOE is .054, or 5.4%, in which case the Giuliani lead *is* outside the margin of error and we would be justified in calling this a statistically significant lead.

Now consider the difference between two independent polls. In two Associated Press/Ipsos polls, approval of President George W. Bush's handing of his job was at 36% in a January 16-18, 2007 poll of 1002 adults, and at 32% in a February 5-7, 2007 poll of 1000 adults. Is this difference statistically significant?

The MOE for the polls is .031, or 3.1% for both polls. Since the polls were conducted independently, the MOE for the difference in approval is

$$\begin{aligned}
MOE(p_2 - p_1) &= 1.96 \times \sqrt{frac.36 \times .641002 - 1 + \frac{.32 \times .68}{1000 - 1}} \\
&= .0415 \text{ or } 4.15\%
\end{aligned}$$

The change of 4 percentage points is slightly less than this margin of error, so the change in approval is not quite statistically significant.

5 Conclusions

The margin of error for a poll is not a simple guide to the margin of error for differences either within the poll or across independent polls. The multiple uses of the phrase "margin of error" compounds the confusion. It is not easy to accurately convey statistical issues in reports intended for a mass audience (or even many political professionals and

[4]This poll was conducted using "Interactive Voice Response" (IVR) technology in which a recorded voice asks the questions and respondents push the phone keypad to register their responses. I ignore the issue of IVR technology compared to standard polls with live interviewers here.

journalists) without becoming arcane. Nonetheless, the frequent confusion over when a difference is "significant" and when it is not needs to be addressed. The most direct way would be for authors of reports to use the correct calculation and report the result without bothering the reader with the details. For audiences accustomed to footnotes, perhaps the details could go there.

And for wordsmiths and others unaccustomed to making statistical calculations, a quick consultation with those who have a calculator ready at hand would solve a number of confusions.

6 Quick Reference

All percentages should be expressed as proportions, so 8% = .08, 30% = .30 and so on. The terms p_1 and p_2 are the proportions supporting candidate 1 and 2, or the proportions for a candidate in polls 1 and 2 and n, n_1 and n_2 are the number of respondents to the poll or to polls 1 and 2.

6.1 Margin of error of the poll

$$\mathrm{MOE}(p) = 1.96 \times \sqrt{\frac{p \times (1-p)}{n-1}}$$

6.2 Margin of error for difference in a single question

$$\mathrm{MOE}(p_1 - p_2) = 1.96 \times \sqrt{\frac{(p_1 + p_2) - (p_1 - p_2)^2}{n-1}}$$

6.3 Margin of error for difference between two polls

$$\mathrm{MOE}(p_2 - p_1) = 1.96 \times \sqrt{\left(\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right)}$$

7 References

Kish, Leslie. 1965. *Survey Sampling* New York: John Wiley & Sons.

Scott, Alastair J. and George A. F. Seber. 1983. "Difference of Proportions from the Same Survey." *The American Statistician* 37:319-320.