

Are the polls right?



© 1997–2009, Millennium Mathematics Project, University of Cambridge.

Permission is granted to print and copy this page on paper for non-commercial use. For other uses, including electronic redistribution, please contact us.

May 1997

Features

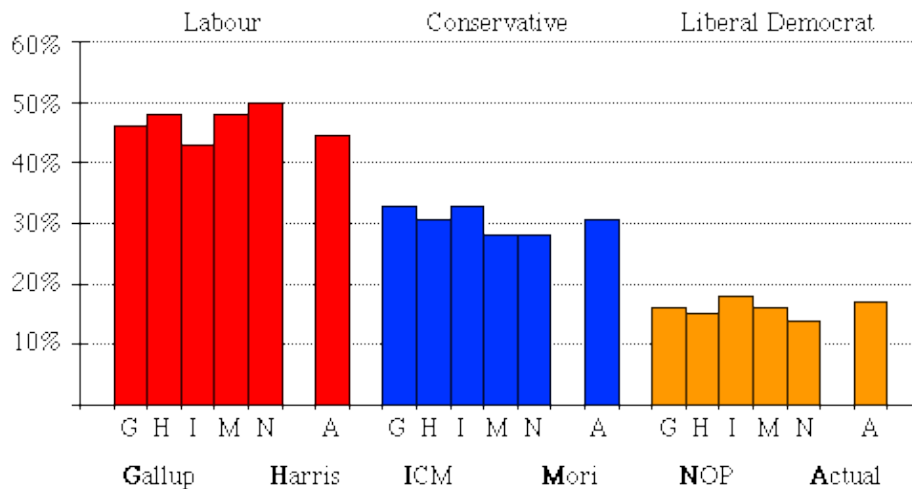
Are the polls right?



The British General Election (May 1997) is an example of how simple mathematical ideas help in understanding information that involves numbers.

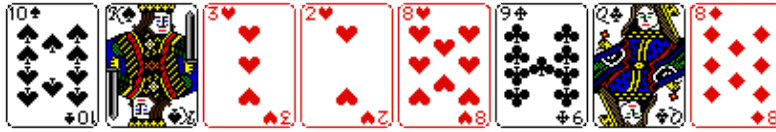
The predictions of the five main opinion polls, taken the day before the election, are shown in the following table and chart:

	Gallup	Harris	ICM	Mori	NOP	Actual
Labour	46	48	43	48	50	44.5
Conservative	33	31	33	28	28	31.0
Liberal Democrat	16	15	18	16	14	17.0



Newspapers usually publish the figures in the form of a table as shown, but this leaves out vital information: the "sampling error". The pollsters cannot ask every voter, so they ask a sample of them. This sample will not usually contain exactly the same proportion of political views as the nation as a whole, for the same reason that if you deal 8 cards from a shuffled pack you would not usually expect to see two from each suit. You are more likely to get something like this:

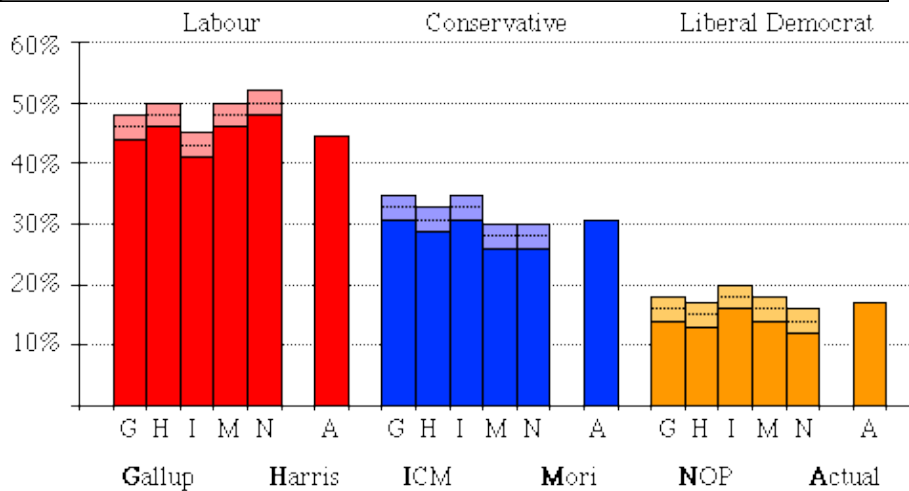
Are the polls right?



On this sample (8 cards out of a population of 52), we have overestimated the Hearts by 50% (viz 3 hearts instead of 2) and underestimated the Diamonds by 50% (viz 1 diamond instead of 2).

Assuming we know approximately what the percentage voting intentions are, then mathematics can tell us just what are the odds for any particular outcome of an opinion poll. From this, it is possible to work out the "likely error". For opinion polls using approximately 1000 people, the "likely error" is about 2% either way. We also need to know what "likely" means. The figure of 2% just quoted will hold in about 2/3 of the cases. Notice that there is about a 1 in 3 chance that the poll will be wrong by more than 2%, for any one figure. With this information, the final opinion polls can be presented more usefully as follows:

	Gallup	Harris	ICM	Mori	NOP	Actual
Labour	44-48	46-50	41-45	46-50	48-52	44.5
Conservative	31-35	29-33	31-35	26-30	26-30	31.0
Liberal Democrat	14-18	13-17	16-20	14-18	12-16	17.0



Each prediction is now given as a band of values.

Looking at these figures, we see that of the fifteen bands, only 9 were right. We expected 10 to be right (2/3 of 15) and 5 to be wrong. All in all, this is quite impressive because we have assumed all along that our only problem was random sampling error. In real life, there are other problems, such as tactical voting, reluctance of voters to tell the pollsters their true intentions; there may even be voters who try to mislead the pollsters.

By the way, we have made some simplifications in the above discussion. For example, the voting intentions for the three parties aren't completely independent. In the card sample example, an overestimate for one suit must imply an underestimate for another. Nevertheless, the picture is much as we have painted it.

What can we tell from all this? First of all, it is not necessary to know or use any complicated mathematics to understand in principle what is going on and make use of the information, but it is important that some mathematician knows how to estimate and justify those "likely errors".

Are the polls right?

Secondly, armed with the idea of sampling error and the figure of 2% for an opinion poll based on 1000 voters, we can read the final poll predictions as being broadly in agreement, although the figures in the form printed by the newspapers appear at first sight to be very scattered.

Thirdly, we see that the claimed accuracy of polls of "about 2%" is not nearly as good as it sounds for predicting the lead one party has over another. Newspapers really ought to improve their presentation of numerical results by showing the bands of likely values.

About a week before polling day one company, ICM, published a poll showing the Labour lead down to single figures, about 5%. This caused a lot of excitement – did it mark the beginning of a slide in Labour support? With our knowledge about sampling error prediction ranges, we would not have regarded this single poll as very good evidence for a slide. In fact, ICM's final poll (figures as above) shows only a 10% lead by Labour; as we have seen, we must expect the occasional poll to be out of line. Philip Cowley's article published in *The Independent* on Sunday "A rogue that rocked the parties" gives a good discussion of this point (*The Independent* on Sunday, 27 April 1997).

Further reading

- Pollsters seek answer to their errors – by Anthony King [Electronic Telegraph](#), Monday 7 April 1997
- Making sense of the polls' position – by John Curtice [The Guardian](#), Thursday 24 April 1997
- [Pollsters vie for the right results](#) – by John Rentoul, *The Independent*, Wednesday 30 April 1997



Plus is part of the family of activities in the Millennium Mathematics Project, which also includes the [NRICH](#) and [MOTIVATE](#) sites.