The odds of the Democrats taking the House

What's the chance that the Democrats will win 15 of the 24 tossup House races and thereby take control of the House of Representatives? In the main text, we explained how to estimate this probability by means of the Central Limit Theorem, arguing that the probability should be approximately given by the normal distribution. Here I'll show you how to compute the probability more exactly. "Show," not "explain" – I'm afraid that this time around, rather than trying to stuff a whole freshman course in probability into two pages, I'm going to share some handy formulas without saying too much about why they work.

To start with: "The Democrats win 15 of 24 seats" is actually a combination of many, many more specific possibilities. For instance, one such possibility is

> "The Democrats win CT-2, FL-22, KY-4,NM-1,NC-11,OH-18,OH-15,PA-6,PA-7,VA-2,AZ-8,CO-7,IN-2,IN-8, and IN-9 but lose IA-1,TX-22,CT-4,CT-5,IL-6,MN-6,OH-1,PA-8, and PA-10."

What's the chance of this happening? Well, we've specified 24 things that have to happen for the above prediction to come true. Each one has a probability of 1/2 of occurring. So – and this is where we use the crucial hypothesis that the races are *independent* – the probability of all 24 races going as predicted is $(1/2)^{24}$, or 0.00000006. Not very likely! But then again, there are a *lot* of ways to pick 15 races for Dems to win out of the 24 available. The number of ways to choose 15 races out of 24 is written $\binom{24}{15}$ and pronounced "24 choose 15." So when I say there are a lot of ways, I mean more precisely that there are $\binom{24}{15} = 1,307,504$ ways. Obviously I didn't count them by hand; $\binom{24}{15}$ is an example of what's called a *binomial coefficient*, which can be computed by the following formula:

$$\binom{n}{k} = \frac{n \times n-1 \times \ldots n-k+1}{k \times k-1 \times \ldots \times 1}.$$

(Exactly *why* this formula computes the number of ways to choose k widgets from a set of n is a bit beyond the scope of this note – I encourage the reader who hasn't seen binomial coefficients before to work them out for some small values of n and k, and see for themself that the binomial coefficient does what I said it does.)

So the chance that Democrats will win *some* set of exactly 15 seats is

$$\binom{24}{15} \times \frac{1}{2}^{24} = .078.$$

But that's not enough – we have to take into account the possibility that the Democrats will win *at least* 15 seats. But that's easy enough – we just add up the chance that they'll win k seats for each k greater than or equal to 15, like so:

$$\binom{24}{15}\times\frac{1}{2}^{24}+\binom{24}{16}\times\frac{1}{2}^{24}+\binom{24}{17}\times\frac{1}{2}^{24}+\binom{24}{18}\times\frac{1}{2}^{24}+\binom{24}{19}\times\frac{1}{2}^{24}+\binom{24}{20}\times\frac{1}{2}^{24}$$
$$+\binom{24}{21}\times\frac{1}{2}^{24}+\binom{24}{22}\times\frac{1}{2}^{24}+\binom{24}{23}\times\frac{1}{2}^{24}+\binom{24}{24}\times\frac{1}{2}^{24}=.1537\ldots$$

which is the 15.4 percent figure given in the article.

It still seems like a lot of work to compute these binomial coefficients, even with calculator in hand. So let me give you a shortcut. The *binomial theorem* says that, for any positive whole number n,

$$(1+x)^n=\binom{n}{0}+\binom{n}{1}x+\binom{n}{2}x^2\ldots+\binom{n}{n-1}x^{n-1}+\binom{n}{n}x^n.$$

In particular, on any moderately au courant calculator you can compute the expression $(0.5+0.5x)^{24}$, which you will find is equal to

$$0.00000006+0.0000014x+\ldots+0.078x^{15}+0.044x^{16}+0.021x^{17}+0.008x^{18}+0.002x^{19}$$
$$+0.0006x^{20}+0.0001x^{21}+0.00002x^{22}+0.000001x^{23}+0.00000006x^{24}.$$

The coefficient of x^k is the probability that the Democrats win k seats; so adding those last ten coefficients together gives you the desired figure of 0.1537.

Not only is this method quicker to compute, it is easy to generalize to more complicated problems. For example, suppose we throw in the 11 seats rated by Rothenberg as leaning Republican. All are held by Republicans, and surely it's not too crazy to think the Dems might take one or two. Let's say these 11 seats have a 90 percent chance of staying Republican and a 10 percent chance of flipping. This corresponds to multiplying our polynomial above by a factor of $(0.9+0.1x)^{11}$. Similarly, the 4 seats (all currently Democratic) which are leaning towards Democrats contribute a factor of $(0.1+0.9x)^4$. Fire up the calculator again and compute $(0.5+0.5x)^{24}(0.9+0.1x)^{11}(0.1+0.9x)^4$; you get a monster polynomial which I don't care to type in in full – but in the middle of the expansion you should see

$$\ldots+0.102x^{19}+0.07x^{20}+0.04x^{21}+0.02x^{22}+0.01x^{23}+\ldots$$

which is to say that the Dems would have a 10 percent chance of winning 19 of the seats (the amount needed for a majority, since we've put 4 Democratic seats in play), a 7 percent chance of getting 20, a 4 percent chance of getting 21, and so on. On this reckoning, the chance of the Democrats capturing the majority goes up to 25 percent. If you want to see the effect of changing that 10 percent chance of flipping, play with the formula yourself!

By the way, it is not supposed to be obvious why the algebraic computation computes the probability I say it does. But to get a sense of how it works, it's useful to imagine there were just two seats at issue; one which the Democrats win with probability p and lose with probability $1-p$, and one which Democrats win with probability q and lose with probability $1-q$. Then by analogy with the algebra above we ought to compute

$$((1-p)+px)\times((1-q)+qx)=(1-p)(1-q)+[p(1-q)+q(1-p)]x+pqx^2$$

The algebra then suggests that the chance of the Democrats winning exactly 1 of the 2 seats is $p(1-q)+q(1-p)$. As an exercise, convince yourself that this is correct!