Evaluating Poll Results

Eliott Rosenberg

February 24, 2020

1 Theory

Suppose that a poll asks voters whether, if candidate A were running against candidate B, they would vote for A or B. Out of N respondents, $\hat{n}_1$ say A, $\hat{n}_2$ say B, and $\hat{n}_3 = N - \hat{n}_1 - \hat{n}_2$ say neither. Suppose that the null hypothesis is that the candidates are tied, i.e. $p_1 = p_2$, where p_i is true probability of a voter choosing option i. Suppose that $\hat{n}_1 > \hat{n}_2$, and define $\delta\hat{p} = (\hat{n}_1 - \hat{n}_2)/N$. What is the p-value corresponding to this outcome? The p-value is the probability of observing the actual outcome or something more extreme, given the null hypothesis. It is given by

$$\begin{aligned} p &= \int_0^1 dp_3 f(p_3) P(\delta\hat{P} > \delta\hat{p} | p_1 = p_2 = (1-p_3)/2) \\ &= \sum_{n_1,n_2} \int_0^1 dp_3 f(p_3) \left(\frac{1-p_3}{2}\right)^{n_1+n_2} p_3^{N-n_1-n_2} \frac{N!}{n_1! n_2! (N-n_1-n_2)!} \Theta(n_1+n_2 \le N)\Theta(n_1 - n_2 \ge N\delta p), \end{aligned} \tag{1}$$

where $f(p_3)$ is the probability density for p_3 in the null hypothesis. We can enforce the constraints by writing the sum as

$$\sum_{n_+ = N\delta p}^{N} \sum_{n_2=0}^{\lfloor \frac{1}{2}(n_+ - N\delta p)\rfloor}, \tag{2}$$

where $n_+ = n_1 + n_2$ and $\lfloor\cdot\rfloor$ is the floor function, which rounds down to the nearest integer.

1.1 Assuming no knowledge of p_3.

Next, let's consider $f(p_3)$. First, let's see what happens if we are maximally agnostic about p_3 and say that $f(p_3) = 1$. Then we can evaluate the p_3 integral. The result is

$$\boxed{p = \frac{1}{N+1} \sum_{n_+ = N\delta p}^{N} \sum_{n_2=0}^{\lfloor \frac{1}{2}(n_+ - N\delta p)\rfloor} 2^{-n_+} \binom{n_+}{n_2}}, \tag{3}$$

where $\binom{n_+}{n_2} = \frac{n_+!}{n_2!(n_+-n_2)!}$. Note that this is the average probability that $n_2 \le \frac{1}{2}(n_+ - N\delta p)$ (i.e. that $n_1 - n_2 \ge N\delta p$), where the average is over all n_+ between 0 and N, with each probability computed according to a binomial distribution with equal weights and n_+ as the total counts.

This expression is easily computed in Matlab as

```
nPlus = Ndp:N;
p = 1/(N+1) * sum( binocdf(floor((nPlus - Ndp)/2), nPlus, 0.5) );
```

See Figures 1-3 for numerical results. We can see that if the candidates' polling percentages differ by $\frac{1}{\sqrt{N}}$, this corresponds to a p-value of about 0.08. Similarly, if their percentages differ by $\frac{2}{\sqrt{N}}$, this corresponds to a p-value of about 0.006.

1.2 Assuming complete knowledge of p_3.

What if, instead, the null hypothesis is that p_3 is equal to a particular value (such as whatever it is in this particular survey)? Then we have

$$\boxed{p = \sum_{n_+=N\delta p}^{N} \sum_{n_2=0}^{\lfloor \frac{1}{2}(n_+ - N\delta p) \rfloor} \left(\frac{1-p_3}{2}\right)^{n_+} p_3^{N-n_+} \frac{N!}{n_2!(n_+ - n_2)!(N-n_+)!}}. \tag{4}$$

In Matlab language, the summand is

```
mnpdf([nPlus - n2,n2,N - nPlus],[(1-p3)/2,(1-p3)/2,p3]);
```

2 Application to a Poll

Now, let's apply Eqs. 3 and 4 to a real poll, namely the Quinnipac University "Swing State Poll" from Feb. 20, 2020.[1] The results are shown in Table 1. Let's adopt the standard that the p-value must be less than 0.05 for the result to be statistically significant. Then, in Michigan, the leads of Sanders and Bloomberg are statistically significant according to Eq. 3 but not Eq. 4. The leads of the other candidates in Michigan are not statistically significant. In Pennsylvania, the leads of Biden, Klobuchar, and Bloomberg are statistically significant according to either equation, but none of the other candidates' leads are significant. In Wisconsin, Trump's lead over any of the other candidates is statistically significant according to either equation.

Table 1: Application of Eqs. 3 and 4 to a Quinnipac University Poll

State	N	Democratic Candidate	$\delta\hat{p}$	p_3	p-value (Eq. 3)	p-value (Eq. 4)
Michigan	845	Sanders	0.05	0.09	0.027	0.067
		Bloomberg	0.05	0.11	0.027	0.065
		Biden	0.04	0.1	0.053	0.11
		Warren	0.02	0.12	0.18	0.27
		Buttigieg	0.01	0.11	0.33	0.39
		Klobuchar	0.01	0.11	0.33	0.39
Pennsylvania	849	Biden	0.08	0.08	0.0022	0.0078
		Klobuchar	0.07	0.09	0.0056	0.018
		Bloomberg	0.06	0.1	0.012	0.034
		Sanders	0.04	0.08	0.054	0.12
		Buttigieg	0.04	0.1	0.054	0.11
		Warren	0.03	0.09	0.11	0.19
Wisconsin	823	Klobuchar	-0.11	0.11	1.10E-04	4.10E-04
		Warren	-0.1	0.08	3.60E-04	0.0015
		Buttigieg	-0.08	0.1	0.0024	0.008
		Bloomberg	-0.08	0.1	0.0024	0.008
		Sanders	-0.07	0.07	0.0057	0.019
		Biden	-0.07	0.09	0.0057	0.018

[1] Available at `https://poll.qu.edu/2020-presidential-swing-state-polls`.

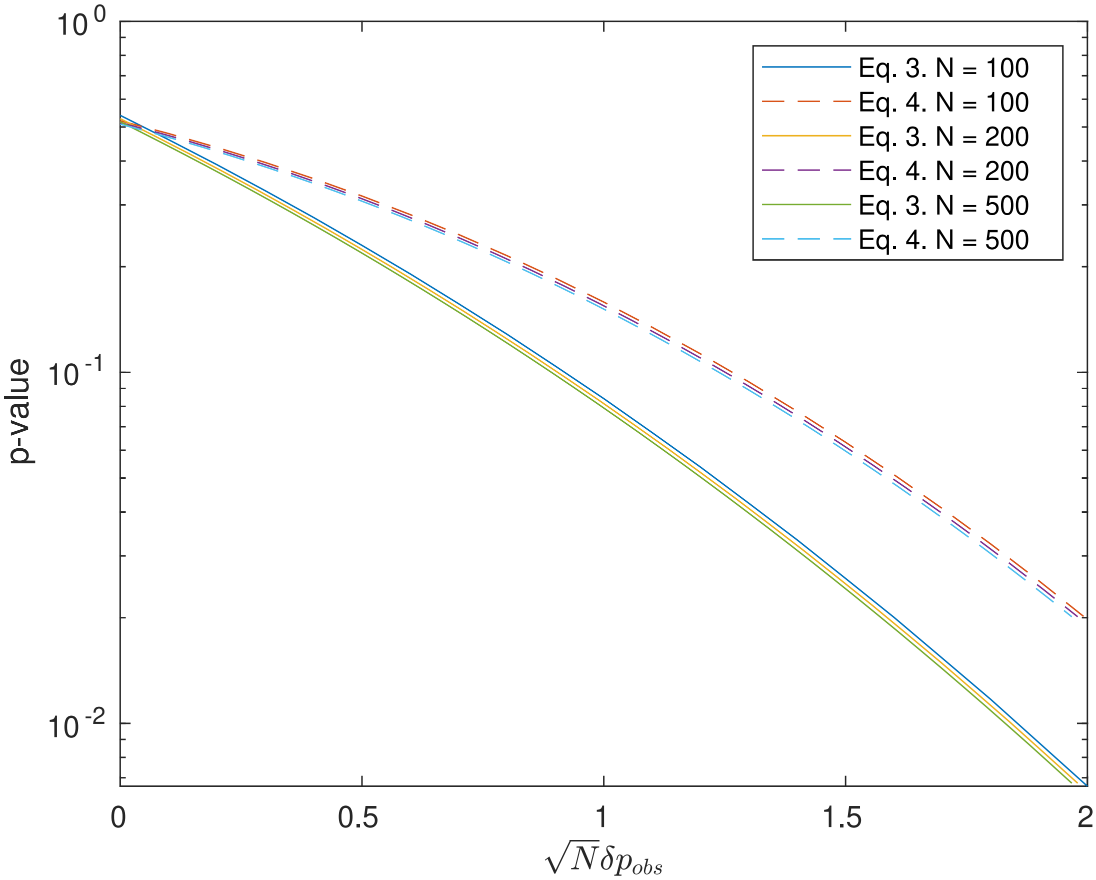

Figure 1: Plot of Eqs. 3 and 4, demonstrating that the functional dependence on N and $\delta\hat{p}$ is primarily just through the product $\sqrt{N}\delta\hat{p}$. In evaluating Eq. 4, we assumed $p_3 = 0.1$.

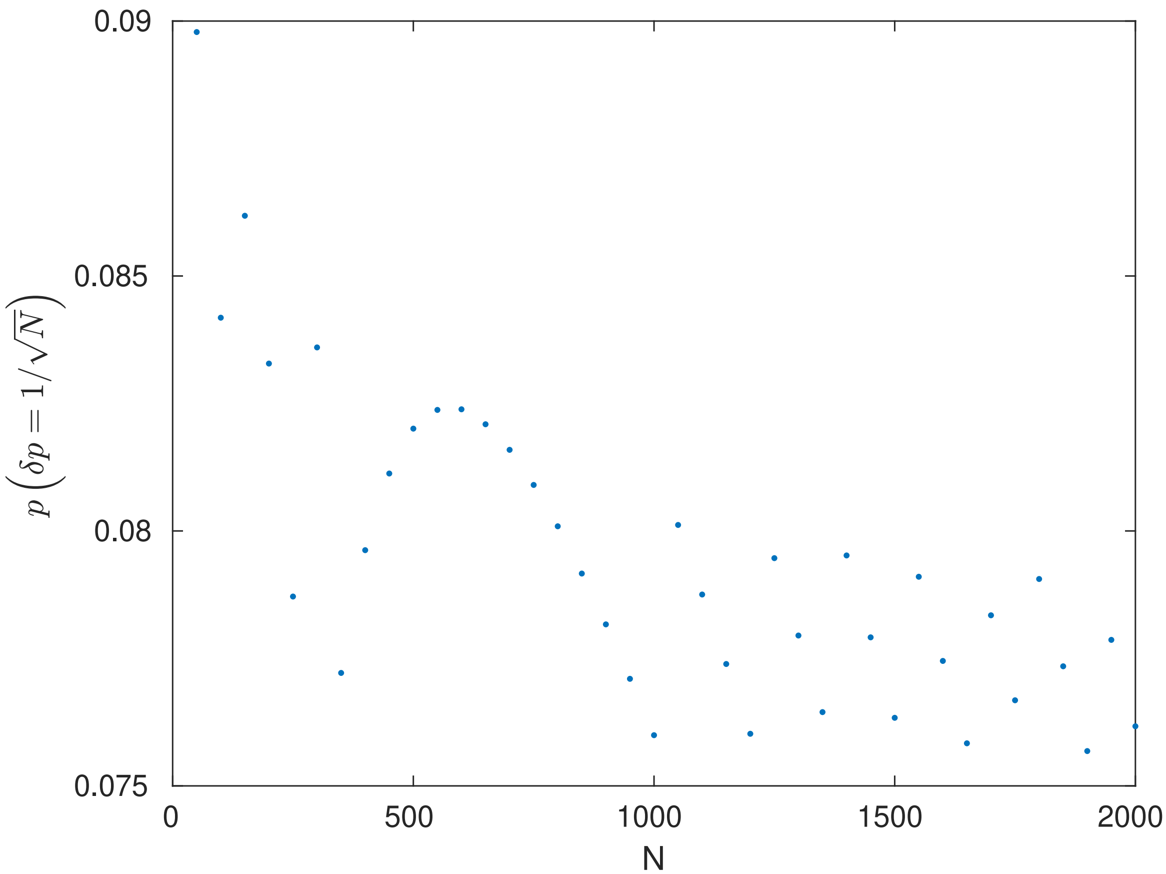

Figure 2: Plot of Eq. 3 evaluated at $N\delta\hat{p} = \text{round}(\sqrt{N})$.

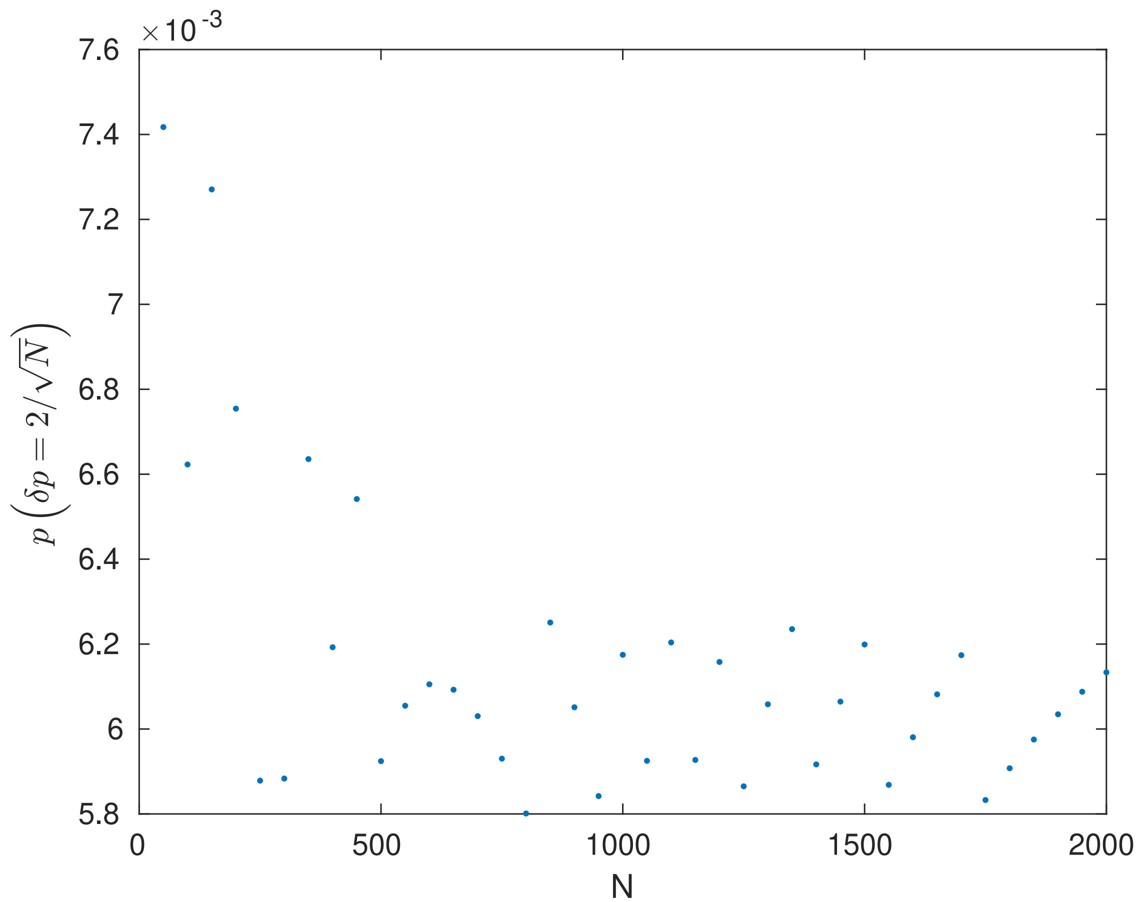

Figure 3: Plot of Eq. 3 evaluated at $N\delta\hat{p} = \mathrm{round}(2\sqrt{N})$.

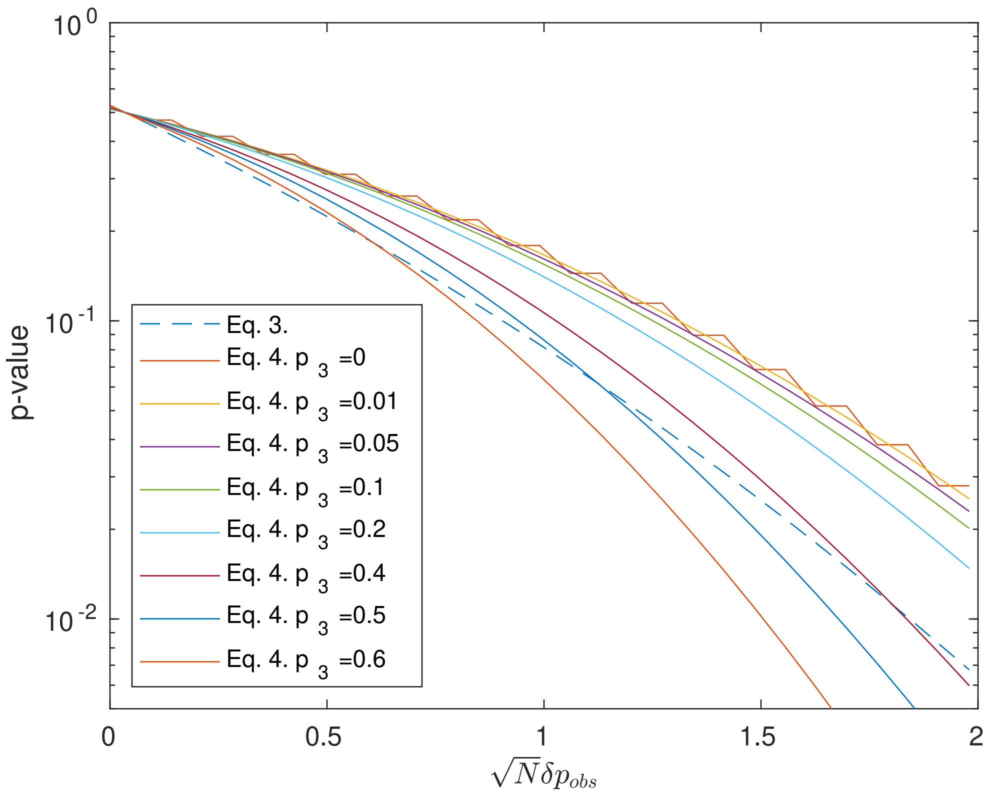

Figure 4: A demonstration of the dependence of Eq. 4 on the population value of p_3 contained in the null hypothesis. Computed for $N = 200$.