FiveThirtyEight's April 23, 2021 Riddler

Emma Knight

April 23, 2021

This week's riddler, from Aaron Wilkowski, is about voting paterns:

Question 1. *Riddler Nation's neighbor to the west, Enigmerica, is holding an election between two candidates, Amy and Beth. Assume every person in Enigmerica votes randomly and independently, and that the number of voters is very, very large. Moreover, due to health precautions, 20 percent of the population decides to vote early by mail.*

On election night, the results of the 80 percent who voted on Election Day are reported out. Over the next several days, the remaining 20 percent of the votes are then tallied.

What is the probability that the candidate who had fewer votes tallied on election night ultimately wins the race?

To generalize this, let's assume p people vote in person and m people vote by mail, and that both p and $m \gg 0$. The case in the riddler is when $p/m = 4$.

For each voter v, define a randomvariable X_v by saying $X_v = 1$ if v votes for Amy and -1 if v votes for Beth. Additionally, define X_p to be the sum of X_v over all the voters that vote in person, and X_m to be the sum of X_v over all the voters that vote by mail.

Since X_v has mean 0 and variance 1 for all v, and that all of the X_vs are independent, one has that X_p has mean 0 and variance p (and X_m has mean 0 and variance m). Since p and m are large, we may assume that X_p and X_m are normal distributions, with PDFs given by $P_p(x) = \dfrac{1}{\sqrt{2p\pi}} \exp\left(\dfrac{-x^2}{2p}\right)$ and $P_m(y) = \dfrac{1}{\sqrt{2m\pi}} \exp\left(\dfrac{-y^2}{2m}\right)$.

In order for the mail-in votes to overturn the in-person votes, one needs $|X_m| > |X_p|$ and them to have different signs. Converting this into a region in the xy-plane, one looks at the two regions where $y > -x > 0$ and $y < -x < 0$. Call this overall region R. We then need to compute $I = \int\int_R P_p(x)P_m(y)\,dA$.

Let $X = x/\sqrt{p}$ and $Y = y/\sqrt{m}$. Using change of variables, one gets that we need to compute $I' = \int\int_{R'} \frac{1}{2\pi} \exp\left(\frac{-X^2-Y^2}{2}\right) dA$, where R' is the union of the regions where $Y > -\sqrt{\frac{p}{m}}X > 0$ and

$Y < -\sqrt{\frac{p}{m}}X < 0$. The integrand is radially symmetric and integrates to 1 over all of $\mathbb{R}^2$. R' is two slices of the plane, so all we need to know is what portion of $\mathbb{R}^2$ lies in those two slices, aka, the sum of the angles of these two slices divided by 2π. Drawing a triangle, it is not too hard to convince yourself that the angle of each of the slices is $\arctan\left(\sqrt{m/p}\right)$, so the probability of the mail-in ballots overturning the in-person ballots is $\dfrac{\arctan\left(\sqrt{m/p}\right)}{\pi}$.

A couple of sanity checks: if $p \gg m$, then this says that the odds of the mail-in ballots overturning the in-person ballots is close to 0, and if $m \gg p$, this says that the odds of the mail-in ballots overturning the in-person ballots is close to $1/2$. Both of these make sense: if there are many more in-person ballots, then the mail-in ballots are basically irrelevant. If there are many more mail-in ballots, then they overturn the in-person result basically as long as they disagree with the in-person result.

In the case in the original riddler, we have $p/m = 4$, so we get a $\dfrac{\arctan 1/2}{\pi}$ chance of overturning the result, which is approximately 14.7%.