A lesson in the errors of statistical thinking: Nate Silver on Trump

Yaneer Bar-Yam and Taeer Bar-Yam
New England Complex Systems Institute and MIT Media Laboratory, Cambridge, MA 02139, USA
(Dated: July 16, 2016)

Nate Silver is one of the most highly regarded statisticians of sports, politics and other domains [1]. During the 2016 presidential campaign, his early analysis of the chances of Donald Trump becoming Republican nominee stands out—he estimated only a 2% probability. Even though statistics are not about actualities but probabilities, subsequent events do not appear to be consistent with those predictions, as he later acknowledged [2–4]. He has explained the problem with the analysis as due to political factors [3], and in terms of the difficulty of analysis [4], but not why the model he used is essentially flawed. Here we point out fundamental problems with the statistical ideas he uses. Statistics begins from an assumption of independence, which is generally not valid. In this case, the assumptions lead to mathematical inconsistencies. This illustrates how statistics can lead to illogic even for sophisticated users. Indeed, perhaps it is more likely to mislead those who are sophisticated—a cautionary tale.

Silver's analysis [2] is based on a gauntlet of six "stages of doom" of nomination. He assigns each stage independently a 1 in 2 chance of being won, leading to less than $2\% = (1/2)^6$ chance of nomination. Like winning 6 coin tosses in a row.

There is an argument that makes Silver's result suspect. Some of the stages of Silver's analysis appear unique to Trump. However, each candidate faces difficulties, and every stage of the nomination is surely not guaranteed to any of them. While the specific terms that are used might not be the same, a similar analysis would hold for each one: gaining and keeping attention, withstanding scrutiny, achieving early state success, building organization, accumulating delegates, and achieving a majority for the convention. If anything, they faced greater challenges because Trump was ahead in polls.

Thus, a similar application of logic would lead us to the conclusion that each had a 2% chance of winning. This is unreasonable because someone has to win—the sum over probabilities has to be 1 (excepting a small chance of a non-candidate nominee). If each candidate had the same probability their chance would be no less than 6%, 1 in 17, which was the number of original candidates. Surely someone has to have a probability greater than 2%. This suggests that Silver's reasoning was not internally consistent.

In fact, Silver wrote the article because of concerns over Trump's lead in polls at the time. One might guess therefore he had more than 1 in 17 chance. The concerns suggest a crowdsourced estimate would be much higher.

There are other assumptions in Silver's gauntlet. The multiplication of factors of 1/2 is based upon assuming that losing any one stage is a prohibitive barrier to the nomination. This doesn't seem well justified. We could just as easily invent other independence assumptions: That each stage has independently 1/2 chance of success, including the nomination—50% instead of 2%. And why 1/2? Perhaps because it is often used in statistics examples.

The real problem with the estimation is dependencies. Overcoming one stage gives a higher probability of overcoming others. It is not quite the case that winning one guarantees winning the others. Still, it is known that factors that give rise to winning one contribute to winning the others, and the fact of winning one helps winning the others (momentum). The strength of dependence is not understood but this question can completely dominate the predictions of the model. Dependence is thus not a small effect, and so even in a rough approximation it must be considered.

Applying statistics is tricky. Despite the issues we raise here, Silver has made mathematics of real world problems a much more highly respected endeavor and he should be given much credit for doing so.

Statistical assumptions are used because they make calculations possible. But if the assumptions are wrong, so are the calculations. What should be done? Silver has written a thoughtful lessons learned [4] pointing to the importance of complexity, feedback loops and chaotic dynamics. Incorporating the mathematical frameworks that these processes refer to will advance analyses beyond statistics to enable better mathematical prediction. Being concerned about interdependence, like the concerns about Brexit causing problems for Europe, is not enough. We need to *understand* interdependence [5] in order to make correct assumptions, and derive correct conclusions.

[1] `http://fivethirtyeight.com`
[2] `http://fivethirtyeight.com/features/donald-trumps-six-stages-of-doom/`
[3] `http://fivethirtyeight.com/features/why-republican-voters-decided-on-trump/`
[4] `http://fivethirtyeight.com/features/how-i-acted-like-a-pundit-and-screwed-up-on-donald-trump/`
[5] Y. Bar-Yam, Dynamics of Complex Systems, Westview Press (1997) `http://necsi.edu/publications/dcs/`