So…What's with 95% Confidence?

Throughout the course, we often desire to achieve 95% confidence (or, if you're reading this later in the course, 5% significance). But ***why***? What's so special about this magical number 95?

Well, I like to think of it as the "Old Faithful" effect: there are plenty of geysers in Yellowstone National Park. So, why do so many folks visit Old Faithful? Well, you might say, "well, duh, dude…it goes off, like, every 90 seconds. Old ***Faithful***, remember?" To which I would inform you that, yes, Old Faithful is pretty regular (between ½ an hour to 2 hours between eruptions), but there are other geysers in Yellowstone that are even more regular. Plume Geyser, for example, goes off ***every*** 30 minutes, and Minute geyser goes off every 60 seconds. "Well, you might add…Old Faithful's BIG!" Sure, I'd tell you…Old Faithful gets to about 185 feet on its biggest bursts, but lots of others in Yellowstone (Giant, Excelsior, Beehive, etc.) do as well. Growing impatient, you might add, "Well, it's the closest to the ***road***?" To which I say, "Yes…but why did they build the road so close to ***that*** one?"

So, what's the deal with Old Faithful? Well, here's my 2 cents: it's the perfect blend of both frequency of eruption and size. That's what I think makes it popular. And that's what I'm going to use to explain 95% confidence.

When you set a confidence[1] level, as you remember from class, you're effectively deciding how many standard deviations away from average you're allowing "typical" data to fall. 95% confidence is roughly 2 standard deviations from average. In Sir Ronald Fisher's textbook "*Statistical Methods for Research Workers*", he notes that 2 standard deviation is sufficient and convenient to decide outlier ranges. 2 standard deviations, as you remember from your normal curve work, represents 95% of your data. Since we can expect a bell curve, via the mighty Central Limit Theorem, then we can expect that, 95% of the time, our sample will yield an average within our interval. I'll call this "19 out of 20" ***accuracy***.

"So", you might say, "why not just increase the standard deviations so you have better than a 95% chance?" Bravo! I respond. Going to 3 standard deviations gets up to about 99.7% confidence, and going out 6 ("6-sigma") gets us to around 99.99966%. However, as you increase the likelihood of trapping your parameter, you're also widening your confidence interval (CI) to a point where it might effectively become useless. Here's an example why…consider this Quinnipiac poll (9.1.11):

> "*President Barack Obama's overall job approval rating has sunk to an all-time low, as 52% of American voters disapprove … Quinnipiac University surveyed 2,730 registered voters with a margin of error of +/- 1.9 percentage points*."

So, with that sample size, and a margin of error of 1.9 percentage points, we can safely say that the majority of Americans disapprove (since $52\% \pm 1.9\%$, even at its low end, is still greater than 50%). However, we lose that certainty as our CI gets wider, because, since confidence is directly proportional to margin of error size, as one goes up, the other must as well. Moving to the "six sigma" level of confidence would mean being forced to accept a 5.7% margin of error. Thus, as our confidence gets larger and larger, we lose ***precision***.

So, from my research, it seems that 95% confidence is the perfect blend of accuracy and precision. Much like Old Faithful, 95% confidence isn't the most accurate, nor the most precise, but it's the best blend of both.

---

[1] This same argument can be made for significance levels, as well.