Forbes http://onforb.es/1IQOJMi

Jerry Bowyer Contributor

Opinions expressed by Forbes Contributors are their own.

OPINION 1/06/2016 @ 10:21AM | 2,092 views

Beer Vs. Eugenics: The Good And The Bad Uses Of Statistics

Like me, you've probably noticed that people tend to think of math as morally neutral. There's the world of ethics, values, faith, meaning and philosophy. Then there's the world of math and science. CP Snow called this artificial division of head and heart 'the two cultures'. Ideologies, philosophies, and religions clash, and then math steps in as the neutral referee. In a debate with Congress over the budget, even President Obama—a man of faith and letters—said that it's not about ideology, 'it's just math'.

But it doesn't take too much of a dip into the history of mathematics and especially the history of statistics to see that mathematical scientists are as agenda-driven as any other intellectuals and that their math tilts toward and then is used to buttress an agenda.

Professors Stephen Ziliak and Deirdre McCloskey have done the historical and theoretical spadework needed to expose the tilt in the foundation of modern statistical theory. They wrote the critically acclaimed and University of Michigan Press best-selling book *The Cult of Statistical Significance* and McCloskey edited and Ziliak contributed to the soon to be published *Oxford Handbook of Professional Economic Ethics* (Oxford University Press), not only to expose that tilt, but also partly to show that economists and other mathematical statisticians—such as drug and medical researchers—have an ethical obligation to reject the notion that research methods are ethically neutral.

With his Dublin Guinness brewery, W.S. Gosset was in the business of

making a consistently good beer at minimum efficient cost. (AP Photo/Jozef Klamar)

They aren't and if the ethics of research are forgotten, great harm is done.

Take R.A. Fisher and the notion of statistical significance, please. Too few researchers realize the birth of modern notions of statistical science was a battle between two titans, espousing two importantly different approaches to the numbers game. The battle was between W.S. Gosset, brewer, businessman, and experimental scientist, and the younger R.A. Fisher, an ambitious academic. The breakthrough insights and statistical techniques came from Gosset, but hardly anyone, even in the world of statistics, knows his name. Why? In part because Gosset was protecting his company's scientific property —Guinness's brewery in Dublin—and so he wrote under a pseudonym. He called himself, 'Student'. And anyone in stats knows that name, and they know what he created, Student's t, a test which can help study statistical precision with small samples. Why small samples? Because experiments cost time and money and Gosset was in the business of making a consistently good beer at minimum efficient cost.

Fisher, on the other hand, took an aristocratic approach. He believed that science was above the human swirl around money and quality control. He took Gosset's insight (so who was really the 'student' and who was the master?), changed it a little bit, plucked out the practical stuff, and published his slightly modified version of Student's test in a widely used textbook. In the process Fisher, not Gosset, got credit for inventing modern statistics. Gosset was busy making beer for Guinness and hampered by the inability to use his name. Fisher worked the world of academic politics and made a name for himself.

Though the master, Gosset cared little for credit; he cared a great deal for truth. And after Fisher 'borrowed' his insights, Gosset spent the rest of his life trying to stop Fisher from using them the wrong way. For Gosset the important thing was what Stephen Ziliak and Deirdre McCloskey call 'oomph'. Oomph is the magnitude of the change created by changing an input. How much better beer do you get by switching varieties and amounts of hops? Using a starchier barley for the beer's malt, how much sweeter is the beer? Most important of all, how does all this change the alcohol content and customer demand? Gosset and Guinness stakeholders wanted to know...how much benefit for how much cost? Fisher and his acolytes had no interest in that. They wanted to know...can I prove in a journal article that the statistics prove my theory right and my opponent's theory wrong? Improving the world? Not our job, except in so far as making and proving theories in the realm of pure science somehow in some unforeseen way improves lives downstream of the academy.

So if it's about winning the argument, there is no loss function, there is no cost function, there is no oomph function, just academic debate. One of the many problems with this is that it can be used to support wrong-headed notions and even more wrong-headed solutions. Most statisticians would be surprised to learn that the founding father of their science, the one who created the P value (or actually slightly modified Student's t into a copyrightable version called the P value) was not a professor of statistics. He wasn't even in the math department when he did all this. He was a professor of eugenics. His first published essay along these lines was entitled 'Racial Repair'. It wasn't just Fisher, his predecessors, Francis Galton and Karl Pearson, were also thoroughgoing eugenicists. Galton in fact is widely considered the father of eugenics and Fisher occupied the Galton Chair of Eugenics. This isn't a matter of coincidence, the persons credited with being the founders of modern statistical analysis just happening to be some of the leading 'lights' in the eugenics movement: these two strands of thoughts were thoroughly and systematically intertwined with one another. (It's worth pointing out that Fisher picked a fight with a more talented mathematician, Jerzy Neyman, who happened to be a Catholic progressive and advocate of the civil rights movement. Ziliak and McCloskey note that Neyman tried to hire David Blackwell, an eminent statistician and black, at Berkeley but was stopped by racist colleagues.) But burgeoning academic disciplines of biometrics and social statistics were the theoretical branch whose application was eugenics. You can't manage what you can't measure, goes the old saying. Well the eugenics world set out to measure racial differences in IQ, head shape and size, social habits, etc. And having measured the races, set out to 'repair' the human race.

And there was no cost function to limit them. No one to say, 'hold on a minute, what's the cost of all this, not just money, but in abuse of human dignity, in violation of human rights?' With no loss function, there's no math to say 'hold on a minute, what if we're wrong, how much damage, how much loss to humanity will there be to our attempt to grab the reigns of human evolution?'. With no oomph function, there was no math which could rise up and ask 'even if the theory is right, precisely how much benefit will the human race get for all of this sterilizing and breeding?"

And an even greater tragedy than the issues of missing cost and loss and oomph functions, is that in the end the one thing which statistical theory is supposed to do, to show that the academic theory being proposed is true, it cannot actually do. It can't do it because random sampling just creates new biases, because experimenting on the whole human race would be an unattainably expensive proposition even if it were otherwise advisable and most of all because the standard test of statistical significance (for more on this read the transcript below, or better yet, read Ziliak and

McCloskey's new book) doesn't actually test the proposed theory. It tests ONE alternate theory, the null hypothesis, the 'nothing to see here' theory, and then fallaciously assumes that rebutting the null hypothesis somehow magically supports the author's favorite theory and not one of the vast number of other alternatives to the null.

In other words, not only is it morally wrong to try to improve mankind by suppressing some races and subsidizing other races, not only is it too 'costly' to do in terms of lost human potential, lost creativity, lost dignity and lost human rights. The standard tests of statistical significance can't prove that with all that wickedness and inefficiency and uncertainty and risk of things going wrong you really would even be improving the human family at all, because the low P value does not prove that black IQs are any different than white ones. And on that particular problem, we've made no real progress from Fisher to the Bell Curve.

I'm indebted to Ziliak and McCloskey for these insights. For more detail on Ziliak's view, please click on the link here or read a partial transcript of our interview here.