7 Deadly Statistical Sins Even the Experts Make

Do you know how to avoid them?

Get the facts >

Why Shrewd Experts "Fail to Reject the Null" Every Time

Eston Martz · 3 October, 2016

[Twitter] 0   [Facebook] 1   [LinkedIn] 0   [Comment] 0

I watched an old motorcycle flick from the 1960s the other night, and I was struck by the bikers' slang. They had a language all their own. Just like statisticians, whose manner of speaking often confounds those who aren't hep to the lingo of data analysis.

It got me thinking...what if there were an all-statistician biker gang? Call them the Nulls Angels. Imagine them in their colors, tearing across the countryside, analyzing data and asking the people they encounter on the road about whether they "fail to reject the null hypothesis."

If you point out how strange that phrase sounds, the Nulls Angels will *know* you're not cool...and not very aware of statistics.

Speaking purely as an editor, I acknowledge that "failing to reject the null hypothesis" *is* cringe-worthy. "Failing to reject" seems like an overly complicated equivalent to *accept*. At minimum, it's clunky phrasing.

But it turns out those rough-and-ready statisticians in the Nulls Angels have good reason to talk like that. From a *statistical* perspective, it's undeniably accurate—and replacing "failure to reject" with "accept" would just be wrong.

What *Is* the Null Hypothesis, Anyway?

Hypothesis tests include one- and two-sample t-tests, tests for association, tests for normality, and many more. (All of these tests are available under the **Stat** menu in Minitab statistical software. Or, if you want a little more statistical guidance, the Assistant can lead you through common hypothesis tests step-by-step.)

A hypothesis test examines two propositions: the null hypothesis (or H_0 for short), and the alternative (H_1). The *alternative* hypothesis is what we hope to support. We presume that the null hypothesis is true, unless the data provide sufficient evidence that it is not.

You've heard the phrase "Innocent until proven guilty." That means innocence is assumed until guilt is proven. In statistics, the null hypothesis is taken for granted until the alternative is proven true.

So Why Do We "Fail to Reject" the Null Hypothesis?

That brings up the issue of "proof."

The degree of statistical evidence we need in order to "prove" the alternative hypothesis is the confidence level. The confidence level is 1 minus our risk of committing a Type I error, which occurs when you incorrectly reject the null hypothesis when it's true. Statisticians call this risk alpha, and also refer to it as the significance level. The typical alpha of 0.05 corresponds to a 95% confidence level: we're accepting a 5% chance of rejecting the null even if it is true. (In life-or-death matters, we might lower the risk of a Type I error to 1% or less.)

Regardless of the alpha level we choose, any hypothesis test has only two possible outcomes:

1. **Reject the null hypothesis** and conclude that the alternative hypothesis is true at the 95% confidence level (or whatever level you've selected).

2. **Fail to reject the null hypothesis** and conclude that *not* enough evidence is available to suggest the null is false at the 95% confidence level.

We often use a p-value to decide if the data support the null hypothesis or not. If the test's p-value is less than our selected alpha level, we reject the null. Or, as statisticians say "When the p-value's low, the null must go."

This still doesn't explain *why* a statistician won't "accept the null hypothesis." Here's the bottom line: failing to reject the null hypothesis does not mean the null hypothesis *is* true. That's because a hypothesis test does not determine *which* hypothesis is true, or even which is most likely: it *only* assesses whether evidence exists to reject the null hypothesis.

"Null Until Proven Alternative"

Hark back to "innocent until proven guilty." As the data analyst, you are the judge. The hypothesis test is the trial, and the null hypothesis is the defendant. The alternative hypothesis is the prosecution, which needs to make its case *beyond a reasonable doubt* (say, with 95% certainty).

If the trial evidence does not show the defendant is guilty, neither has it proved that the defendant *is* innocent. However, based on the available evidence, you can't reject that *possibility*. So how would you announce your verdict?

"Not guilty."

That phrase is perfect: "Not guilty"doesn't say the defendant *is* innocent, because that has not been proven. It just says the prosecution couldn't convince the judge to abandon the assumption of innocence.

So "failure to reject the null" is the statistical equivalent of "not guilty." In a trial, the burden of proof falls to the prosecution. When analyzing data, the entire burden of proof falls to your sample data. "Not guilty" does not mean "innocent," and "failing to reject" the null hypothesis is quite distinct from "accepting" it.

So if a group of marauding statisticians in their Nulls Angels leathers ever asks, keep yourself in their good graces, and show that know "failing to reject the null" is not "accepting the null."

You Might Also Like:

Bewildering Things Statisticians Say: "Failure to Reject the Null Hypothesis"
Alphas, P-Values, and Confidence Intervals, Oh My!
Common Statistical Mistakes You Should Avoid
Equivalence Testing for Quality Analysis (Part I): What are You Trying to Prove?

Comments

Authors

Carly Barry
Patrick Runkel
Kevin Rudy
Jim Frost
Greg Fox
Eric Heckman
Dawn Keller
Eston Martz
Bruno Scibilia
Eduardo Santiago
Cody Steele