

# Can we prove the null hypothesis (H0)?

## Statistics and logic

Suppose you want to compare two populations. Normally, you want to prove that they are significantly different. One of the most common ways to do so is to test if the difference in means is a result of condition or occurred by chance. Common tests to assess this are t-test or in ANOVA. You compute a test statistic and check how likely it is to happen under the assumption that there is no difference between these groups. If it is very unlikely (which is quantitatively expressed by  $\alpha$ ) you reject the assumption and conclude that the groups are different.

Idea is based on the classical logic and a so-called *modus tollens*. This rule can be stated like this

“ Assume some hypothesis  $H$ . You observe a consequence which contradicts  $H$ . Therefore you conclude that the assumption is incorrect and that the negation of  $H$  is true. ”

In statistics, however, we are dealing with a probabilistic version of this rule, which may lead to some misunderstandings. In a similar language it can be stated as

“ Assume  $H_0$  (no difference between groups). You observe data which is “unlikely” under the assumption  $H_0$ . You conclude that “probably”  $H_0$  assumption is incorrect and “probably” negation of  $H_0$  is true (meaning there is a difference between groups). ”

Now, imagine that you expect no difference between two groups. How to prove it?

Note that, in the framework of the rule stated above, it is not possible to prove that  $H_0$  is correct. We assume  $H_0$ , so under this assumption, it is evidently true, but we don't know if there exist data which can contradict  $H_0$ . To illustrate the problem consider a scenario from the famous book of Nassim Taleb, "The black swan".

“Assume that all swans are white. Investigate some (even very large) sample of swans. You observe that all of them are white. Conclude (INCORRECTLY) that all swans are white.”

If we want to be sure that all swans are white we need to check **all swans**. Only then, we may be sure that there is no counterexample.

## All means are different

Things get even harder in the context of the initial scenario with two populations. There is a finite number of subjects in each of them so, almost surely, real means in these populations are not equal. Let's consider another example.

Assign all people randomly into two groups. Measure their heights and take means in both groups. These means will be very close to each other but will be different at some decimal point! Most of the tests will show no difference even with large samples, but still  $H_0$  is false.

**By a similar argument in almost any experiment  $H_0$  is false!!!**

Ok, but if we can reject all  $H_0$ s in the first place, why do we care about these sort of tests anyway? The key here is the "significance". In fact, by the construction, our tests will detect the difference only if it is relatively big. In practice what we care about is not the existence of the difference but it's magnitude.

So, when we want to "prove the null hypothesis", we should focus on the magnitude of the difference between the means instead of considering their equality.

## Heuristic solution

In practice you can follow this procedure:

1. Define a margin you want to tolerate (e.g. difference in means not larger than 5%).

2. Figure out the design and, in particular, a sufficiently large sample which makes you confident that you avoid “no effect” occurring by chance.
3. If your confidence intervals of estimators of means are small and the difference is below the margin you conclude that   $H_0$  is very probable.

## References

- [1] Cohen J, *The Earth is Round* (  $p < 0.05$ ), *American Psychologist* **49**(12), 997-1003 (1994)
- [2] Streiner, D.L. *Unicorns Do Exist: A Tutorial on “Proving” the Null Hypothesis*, *Canadian Journal of Psychiatry* **48**, 756-761 (2003).

Posted on [January 2, 2015](#) by [Lukasz Kidzinski](#) in [Philosophy](#), [Statistics](#) Tagged [ANOVA](#), [null hypothesis](#), [t-test](#), [testing](#)

---

Copyright © datamine.it 2017. WordPress theme by [Ryan Hellyer](#).