UNIVERSIDADE FEDERAL
DE MINAS GERAIS

Design and Analysis of Experiments

08 - Testing Equivalence and Non-Inferiority

Version 2.11

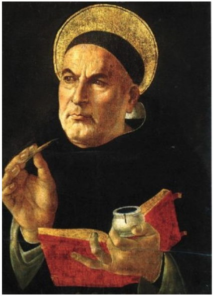
Felipe Campelo
http://www.cpdee.ufmg.br/~fcampelo

Graduate Program in Electrical Engineering

Belo Horizonte
April 2015

BY NC SA

"*Distinctions drawn by the mind*
are not necessarily equivalent
to distinctions in reality."



Thomas Aquinas
1225 - 1274
Italian philosopher and theologian.

Image: `http://www.bfriars.ox.ac.uk/hall/aquinas-ins/`

Testing equivalence
Introduction

The tests introduced in the preceding chapters deal with situations in which one is interested in detecting *differences* between a population parameter θ – e.g., a population mean μ or a difference between population means $(\mu_1 - \mu_2)$ – and its nominal value θ_0 under a null hypothesis;

Another useful class of experiments in engineering and science is one in which the experimenter is interesting in investigating *equivalence* (within a given margin of error), for instance:

- Conformity/compliance testing (industrial certification);
- Equivalence of effects (pharmaceutical industry);

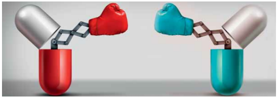


Image adapted from: `http://goo.gl/iXeCiY`

In principle, one could express this as a shift in focus from trying to establish whether a population parameter is different from a given reference to trying to determine whether it is equal to that reference.

In usual (two-sided) comparative studies, the alternative hypothesis (i.e., the one that presents novelty in relation to the current state of knowledge) is the one of difference between the parameters of interest - that is, unless there is strong evidence of differences, one cannot rule out the null hypothesis of equality;

In equivalence testing, the situation is reversed: the (approximate) equality of two parameters is the novelty one hopes to establish. Consequently, the burden of proof shifts to providing evidence that there is no difference.

The term *equivalent* is not used strictly, but to mean the absence of practical differences - that is, any differences that might exist fall within an *equivalence margin* or *limit of practical significance* δ^*.

Using this approach, the equivalence of two parameters can be established if a sample provides enough evidence that the true difference is smaller than δ^* units.

Testing Non-inferiority

Definition

A similar concept to equivalence testing is the definition of non-inferiority of a given treatment/ process/ method in relation to another (e.g., a standard solution).

In non-inferiority tests, one can declare that a given process is not worse than a standard one only if enough evidence is provided to conclude that the performance of the proposed process is no more than δ^* units worse than that of the standard.

In the case of non-inferiority tests, one can in principle use a regular test of differences with a one-sided alternative (which would be equivalent to setting $\delta^* = 0$), or define the null hypothesis in a way that includes δ^* in its formulation.

Comparison of studies

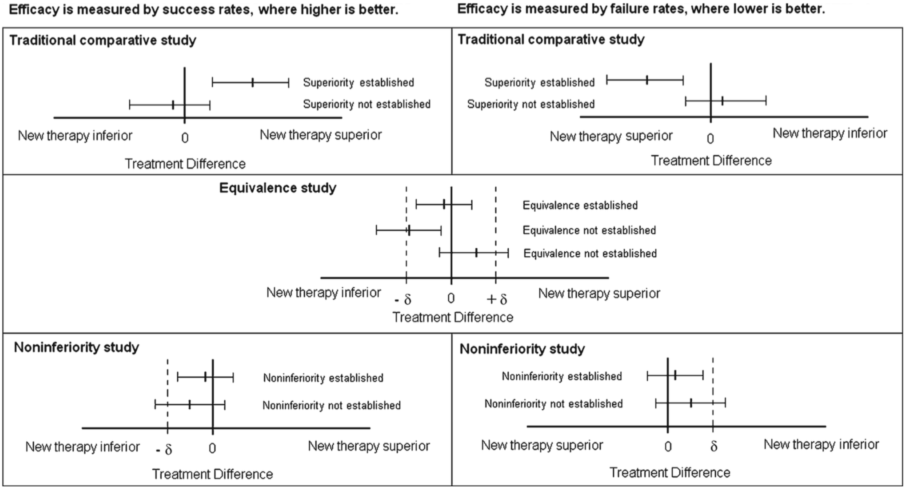


Image: Walker and Nowacki (2011), J. General Internal Medicine 26(2):192-196.

Testing Equivalence
Quick-and-dirty approach

A simple way of thinking about testing equivalence of two means is to observe confidence intervals instead of p-values:

> "*Equivalence can be established at the α significance level if a $(1-2\alpha)$-confidence interval for the difference between the two means is contained within a interval $\pm\delta^*$.*"

The difference between testing for differences and for equivalence can be easily illustrated using this approach:

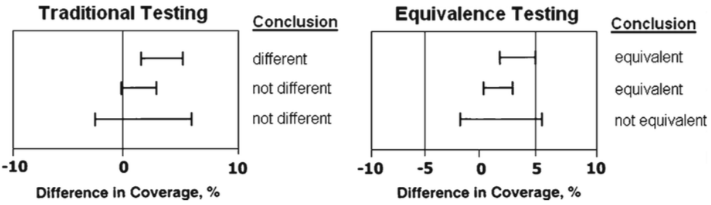


Image: Walker and Nowacki (2011), J. General Internal Medicine 26(2):192-196.

Methodology

Equivalence test for a single mean

An equivalence test for a single population mean can be expressed by the hypotheses:

$$\begin{cases} H_0: & |\mu - \mu_0| \geq \delta^* \\ H_1: & |\mu - \mu_0| < \delta^* \end{cases}$$

The most usual way of testing these hypotheses is the TOST (*two one-sided tests*) method. As the name suggests, two one-sided significance tests are constructed so that the desired statistical properties can be achieved. Using our standard notation:

$$\begin{cases} H_0^1: & \mu - \mu_0 = -\delta^* \\ H_1^1: & \mu - \mu_0 > -\delta^* \end{cases} \qquad \begin{cases} H_0^2: & \mu - \mu_0 = \delta^* \\ H_1^2: & \mu - \mu_0 < \delta^* \end{cases}$$

If both tests reject their respective H_0, then equivalence (within the equivalence margin δ^*) can be declared with significance level α.

Methodology

Equivalence test for a single mean

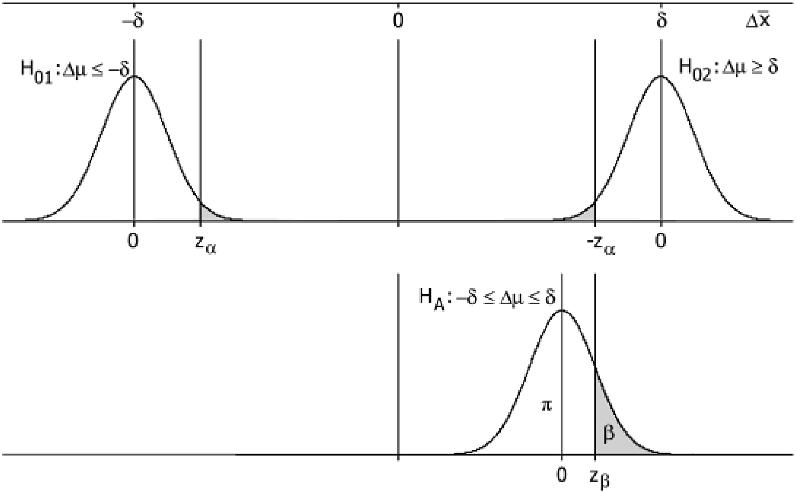


Image: Matthews (2010), Sample Size Calculations, MMB. pg. 46

Equivalence of a single mean

Sample size

Sample sizes for testing equivalence of a single mean can be derived using essentially the same considerations used for the usual tests. In the case of a single sample:

$$n \geq \left(\frac{\left(t_{(1-\alpha)} + t_{(1-\beta)}\right) \hat{\sigma}}{\delta^* - (\mu - \mu_0)} \right)^2$$

As in the previous cases, iteration is needed to solve for n (since the quantiles of the t distribution depend on n). Use $t_x = z_x$ for the first iteration.

Bibliography

Required reading

1. E. Walker, A.S. Nowacki, *Understanding Equivalence and Noninferiority Testing*, Journal of General Internal Medicine 26(2):192-196, 2011.

Recommended reading

1. P. Mathews, *Sample Size Calculations: Practical Methods for Engineers and Scientists*, Ch. 2.4, 1st ed., MMB, 2010.

About this material

Conditions of use and referencing

This work is licensed under the Creative Commons CC BY-NC-SA 4.0 license (Attribution Non-Commercial Share Alike International License version 4.0).

http://creativecommons.org/licenses/by-nc-sa/4.0/

Please reference this work as:

Felipe Campelo (2015), *Lecture Notes on Design and Analysis of Experiments*. Online: https://github.com/fcampelo/Design-and-Analysis-of-Experiments Version 2.11, Chapter 8; Creative Commons BY-NC-SA 4.0.

```
@Misc{Campelo2015-01,
  title={Lecture Notes on Design and Analysis of Experiments},
  author={Felipe Campelo},
  howPublished={\url{https://github.com/fcampelo/Design-and-Analysis-of-Experiments}},
  year={2015},
  note={Version 2.11, Chapter 8; Creative Commons BY-NC-SA 4.0.},
}
```

CC
SOME RIGHTS RESERVED