

Title: Equivalence tests
Teacher: Michael Meyners
Software: no software exercises are planned
Duration: 1 day
Audience: Sensometricians / statisticians and sensory scientists who want to learn about equivalence tests, illustrated by applications from sensory sciences
Background: basic understanding of statistics is advantageous

Aims and Scope

Equivalence (also called similarity or parity) tests are becoming increasingly popular in many application areas, among which sensory sciences. They can and should be applied whenever the aim of the study is not to show differences, but to conclude similarity.

There has been quite some debate about pros and cons of different approaches, and this debate has reached the sensory and sensometrics community in recent years (e.g. a mini-symposium at the 2008 Sensometrics meeting and the respective special issue of FQAP, and several letters to the editor in FQAP 19 (3)). Parts of the sometimes heated debate are, in my opinion, due to mutual misunderstandings, and to different objectives and hence selection criteria for an “optimal” test.

The workshop intends to give an introduction into equivalence tests, starting with some general considerations on the statistical testing of hypotheses (depending on the prior knowledge of the participants). I will subsequently give an overview over the most common approaches, some of which are shown to be inappropriate (e.g. the power approach). Some valid and relatively simple methods will be introduced and their correspondence to confidence intervals clarified, while the mathematical details of “the emperor’s new tests” (Perlman & Wu, Stat Sci 14, 1999) will be skipped. Instead, I will discuss the pros and cons of different approaches, including the recent proposal of Ennis & Ennis using symmetric open intervals. Finally, equivalence tests for multi-parameter distributions (like, e.g., the Beta-binomial model) will be introduced.

A main objective of the workshop is to educate the participants to identify a problem which needs to be addressed by an equivalence approach rather than a difference test. Subsequently, instead of introducing one approach only and rejecting the others, I will discuss the pros and cons and the implications of the tests. The participants will therefore learn to make an appropriate choice for future problems; they should learn that none of the tests is generally superior to all others (and that the differences are often enough negligible).

It is assumed that the participants are familiar with t-tests, binomial tests and alike; the ideas of this workshop will be presented without sophisticated methods and in an interactive manner; they will be illustrated by examples from sensory applications.

Course outline (sessions will last 1:30-2h approximately)

Session 1: Statistical hypothesis testing, reversed hypotheses, and equivalence tests

- The general notion of statistical inference
- The role of the null and alternative hypothesis, significance level, and power
- How to derive evidence against H_0 to prove the alternative H_1
- “Absence of evidence is not evidence of absence”
- Margins of equivalence
- Equivalence vs. non-inferiority / non-superiority
- The power approach
- The approach of Anderson & Hauck (1983)

Session 2: The two one-sided tests approach (TOST)

- Splitting the test problem into two
- Justification of its validity
- Relationship to confidence intervals
- Least Equivalent Allowable Difference (LEAD)
- Implications: Pros and Cons

Session 3: “The emperor’s new tests”

- Power limitations of the TOST
- Unbounded rejection regions
- Approaches by Berger & Hsu (1996), Brown, Hwang & Munk (1997), and others
- Open symmetric intervals (Ennis & Ennis, 2009)
- Implications: Pros and Cons
- Equivalence tests in R&D and in judicial proceedings of advertising claims: Your aim determines your approach!
- An analysis of the heated discussion

Session 4: Multiparameter tests and equivalence tests in sensometrics

- Equivalence tests for distributions with more than one unknown parameter
- Example: The Beta-Binomial model
- A commented literature overview with focus on publications in Sensometrics
- Summary and conclusions: Some guidelines for choosing the right approach

Participants are encouraged to let me know in advance if they have specific applications they would like to be covered; I'm happy to take their examples as far as possible.