

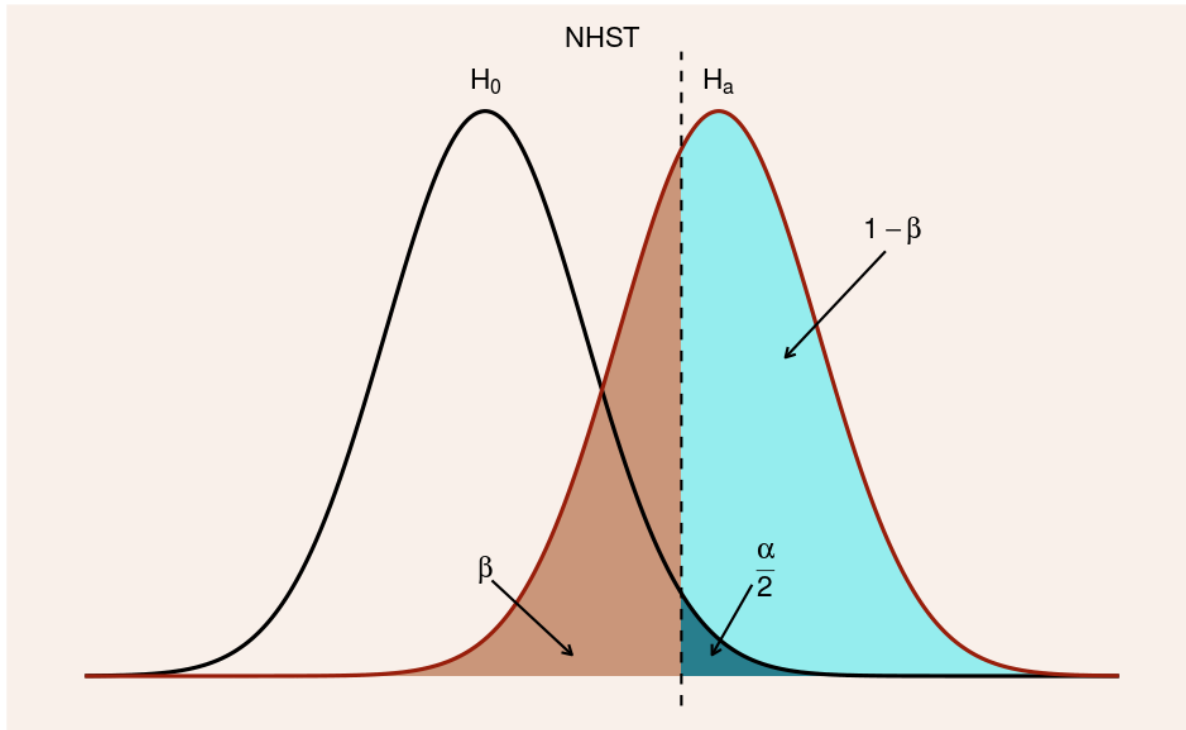
# Data Analysis in Software Engineering. Part 4

## Classical Hypothesis Testing

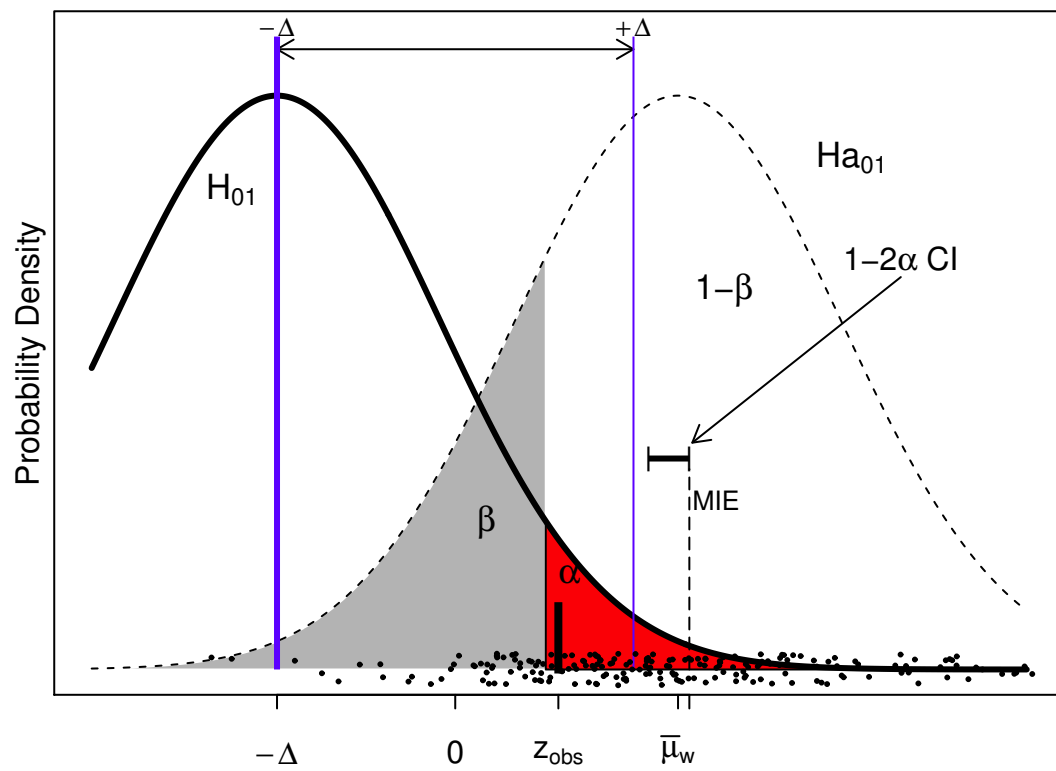
- By “classical” we mean the standard “frequentist” approach to hypothesis testing. The “frequentist” approach to probability sees it as the frequency of events in the long run. We repeat experiments over and over and we count the times that our object of interest appears in the sequence.
- The classical approach is usually called **null hypothesis significance testing** (NHST) because the process starts by setting a null hypothesis  $H_0$  which is the opposite about what we think is true.
- The rationale of the process is that the statistical hypothesis should be *falsifiable*, that is, we can find evidence that the hypothesis is not true. We try to find evidence against the null hypothesis in order to support our alternative hypothesis  $H_A$ .
- Usually, the null hypothesis is described as the situation of “no effect” and the alternative hypothesis describes the effect that we are looking for.
- After collecting data, taking an actual sample, we measure the distance of our parameter of interest from the hypothesized population parameter, and use the facts of the sampling distribution to determine the probability of obtaining such a sample *assuming the hypothesis is true*. This amounts to a test of the hypothesis.
- If the probability of our sample, given the null hypothesis is high, this provides evidence that the null hypothesis is true. Conversely, if the probability of the sample is low (given the hypothesis), this is evidence against the null hypothesis. The hypothesis being tested in this way is named the *null hypothesis*.
- The goal of the test is to determine if the null hypothesis can be rejected. A statistical test can either reject or fail to reject a null hypothesis, but never prove it true.
- We can make two types of errors: false positive (Type I) and false negative (Type II)
- Type I and Type II errors

		Null hypothesis ( $H_0$ ) is	
		Valid/True	Invalid/False
Judgement of Null Hypothesis ( $H_0$ )	Reject	Type I error False Positive	Correct inference True Positive
	Fail to reject	Correct inference True Negative	Type II error False negative

- Two-tailed NHST



- One-tailed NHST



- elementary example

```
data = c(52.7, 53.9, 41.7, 71.5, 47.6, 55.1, 62.2, 56.5, 33.4, 61.8, 54.3, 50.0, 45.3, 63.4, 53.9, 65.5)
t.test(data, mu=50, alternative = 'greater')
```

```
##
## One Sample t-test
##
## data: data
## t = 2.1562, df = 23, p-value = 0.02088
## alternative hypothesis: true mean is greater than 50
## 95 percent confidence interval:
## 50.88892 Inf
## sample estimates:
## mean of x
## 54.33333
```

- Quick introduction at <https://psychstatsworkshop.wordpress.com/2014/08/06/lesson-9-hypothesis-testing/>

## p-values

- p-value: the p-value of a statistical test is the probability, computed as suming that  $H_0$  is true, that the test statistic would take a value as extreme or more extreme than that actually observed.
- <http://www.nature.com/news/psychology-journal-bans-p-values-1.17001>
- <https://www.sciencenews.org/blog/context/p-value-ban-small-step-journal-giant-leap-science>

The screenshot shows the ScienceNews website. At the top is the 'nature' logo with the tagline 'International weekly journal of science'. Below it is a navigation bar with links: Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, and For Authors. A secondary navigation bar shows 'Archive', 'Volume 519', 'Issue 7541', 'Research Highlights: Social Selection', and 'Article'. Below this is a sub-header 'NATURE | RESEARCH HIGHLIGHTS: SOCIAL SELECTION' with social media icons. The main headline is 'Psychology journal bans *P* values' with a sub-headline 'Test for reliability of results 'too easy to pass', say editors.' The author is 'Chris Woolston' and the date is '26 February 2015 | Clarified:'. On the right, there's a 'Subscribe' button and a 'Search Science News...' field. Below the headline, there's a 'Context' section with a 'P value ban: small step for a journal, giant leap for science' article by Tom Siegfried, dated March 17, 2015. On the left, there's an 'Explore' section with 'LATEST' and 'MOST VIEWED' tabs. The 'LATEST' tab is selected, showing three news items: 'Polar bears "walking hibernation" not much of an energy saver' by Susan M. Leake (July 16, 2015), 'Good luck outsmarting a mosquito' by Susan M. Leake (July 16, 2015), and 'Defense hormones guide plant'.